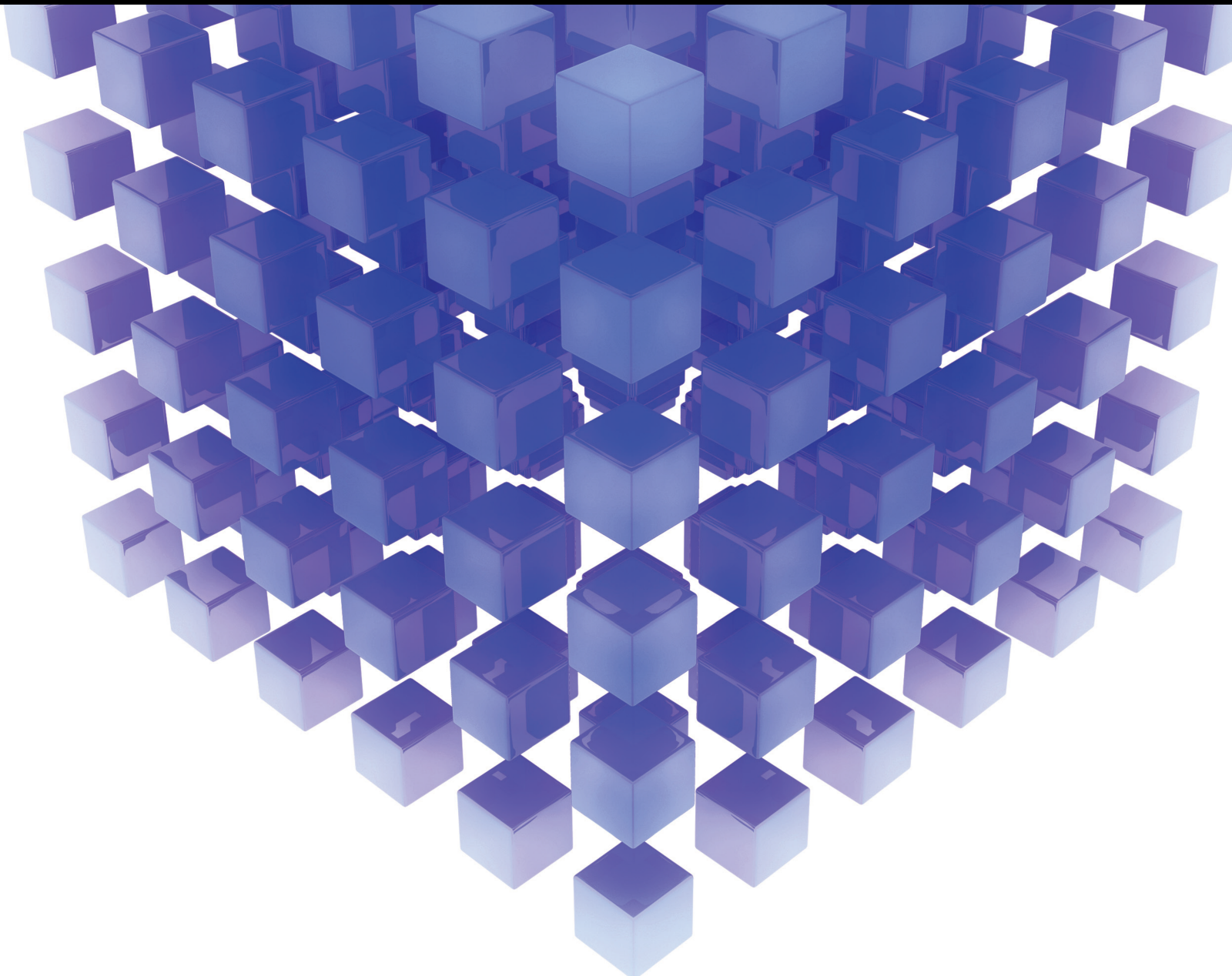


Mathematical Problems in Engineering

# Graph-Theoretic Techniques for the Study of Structures or Networks in Engineering

Lead Guest Editor: Jia-Bao Liu

Guest Editors: Shaohui Wang and Muhammad Javaid





---

# **Graph-Theoretic Techniques for the Study of Structures or Networks in Engineering**



Mathematical Problems in Engineering

---

**Graph-Theoretic Techniques for the  
Study of Structures or Networks in  
Engineering**

Lead Guest Editor: Jia-Bao Liu

Guest Editors: Shaohui Wang and Muhammad  
Javaid




---

Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Mathematical Problems in Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Chief Editor

Guangming Xie , China

## Academic Editors

Kumaravel A , India  
Waqas Abbasi, Pakistan  
Mohamed Abd El Aziz , Egypt  
Mahmoud Abdel-Aty , Egypt  
Mohammed S. Abdo, Yemen  
Mohammad Yaghoub Abdollahzadeh  
Jamalabadi , Republic of Korea  
Rahib Abiyev , Turkey  
Leonardo Acho , Spain  
Daniela Adnessi , Italy  
Arooj Adeel , Pakistan  
Waleed Adel , Egypt  
Ramesh Agarwal , USA  
Francesco Aggogeri , Italy  
Ricardo Aguilar-Lopez , Mexico  
Afaq Ahmad , Pakistan  
Naveed Ahmed , Pakistan  
Elias Aifantis , USA  
Akif Akgul , Turkey  
Tareq Al-shami , Yemen  
Guido Ala, Italy  
Andrea Alaimo , Italy  
Reza Alam, USA  
Osamah Albahri , Malaysia  
Nicholas Alexander , United Kingdom  
Salvatore Alfonzetti, Italy  
Ghous Ali , Pakistan  
Nouman Ali , Pakistan  
Mohammad D. Aliyu , Canada  
Juan A. Almendral , Spain  
A.K. Alomari, Jordan  
José Domingo Álvarez , Spain  
Cláudio Alves , Portugal  
Juan P. Amezcua-Sanchez, Mexico  
Mukherjee Amitava, India  
Lionel Amodeo, France  
Sebastian Anita, Romania  
Costanza Arico , Italy  
Sabri Arik, Turkey  
Fausto Arpino , Italy  
Rashad Asharabi , Saudi Arabia  
Farhad Aslani , Australia  
Mohsen Asle Zaem , USA

Andrea Avanzini , Italy  
Richard I. Avery , USA  
Viktor Avrutin , Germany  
Mohammed A. Awadallah , Malaysia  
Francesco Aymerich , Italy  
Sajad Azizi , Belgium  
Michele Bacciocchi , Italy  
Seungik Baek , USA  
Khaled Bahlali, France  
M.V.A Raju Bahubalendruni, India  
Pedro Balaguer , Spain  
P. Balasubramaniam, India  
Stefan Balint , Romania  
Ines Tejado Balsera , Spain  
Alfonso Banos , Spain  
Jerzy Baranowski , Poland  
Tudor Barbu , Romania  
Andrzej Bartoszewicz , Poland  
Sergio Baselga , Spain  
S. Caglar Baslamisli , Turkey  
David Bassir , France  
Chiara Bedon , Italy  
Azeddine Beghdadi, France  
Andriette Bekker , South Africa  
Francisco Beltran-Carbajal , Mexico  
Abdellatif Ben Makhlof , Saudi Arabia  
Denis Benasciutti , Italy  
Ivano Benedetti , Italy  
Rosa M. Benito , Spain  
Elena Benvenuti , Italy  
Giovanni Berselli, Italy  
Michele Betti , Italy  
Pietro Bia , Italy  
Carlo Bianca , France  
Simone Bianco , Italy  
Vincenzo Bianco, Italy  
Vittorio Bianco, Italy  
David Bigaud , France  
Sardar Muhammad Bilal , Pakistan  
Antonio Bilotta , Italy  
Sylvio R. Bistafa, Brazil  
Chiara Boccaletti , Italy  
Rodolfo Bontempo , Italy  
Alberto Borboni , Italy  
Marco Bortolini, Italy

Paolo Boscariol, Italy  
Daniela Boso , Italy  
Guillermo Botella-Juan, Spain  
Abdesselem Boulkroune , Algeria  
Boulaïd Boulkroune, Belgium  
Fabio Bovenga , Italy  
Francesco Braghin , Italy  
Ricardo Branco, Portugal  
Julien Bruchon , France  
Matteo Bruggi , Italy  
Michele Brun , Italy  
Maria Elena Bruni, Italy  
Maria Angela Butturi , Italy  
Bartłomiej Błachowski , Poland  
Dhanamjayulu C , India  
Raquel Caballero-Águila , Spain  
Filippo Cacace , Italy  
Salvatore Caddemi , Italy  
Zuowei Cai , China  
Roberto Caldelli , Italy  
Francesco Cannizzaro , Italy  
Maosen Cao , China  
Ana Carpio, Spain  
Rodrigo Carvajal , Chile  
Caterina Casavola, Italy  
Sara Casciati, Italy  
Federica Caselli , Italy  
Carmen Castillo , Spain  
Inmaculada T. Castro , Spain  
Miguel Castro , Portugal  
Giuseppe Catalanotti , United Kingdom  
Alberto Cavallo , Italy  
Gabriele Cazzulani , Italy  
Fatih Vehbi Celebi, Turkey  
Miguel Cerrolaza , Venezuela  
Gregory Chagnon , France  
Ching-Ter Chang , Taiwan  
Kuei-Lun Chang , Taiwan  
Qing Chang , USA  
Xiaoheng Chang , China  
Prasenjit Chatterjee , Lithuania  
Kacem Chehdi, France  
Peter N. Cheimets, USA  
Chih-Chiang Chen , Taiwan  
He Chen , China

Kebing Chen , China  
Mengxin Chen , China  
Shyi-Ming Chen , Taiwan  
Xizhong Chen , Ireland  
Xue-Bo Chen , China  
Zhiwen Chen , China  
Qiang Cheng, USA  
Zeyang Cheng, China  
Luca Chiapponi , Italy  
Francisco Chicano , Spain  
Tirivanhu Chinyoka , South Africa  
Adrian Chmielewski , Poland  
Seongim Choi , USA  
Gautam Choubey , India  
Hung-Yuan Chung , Taiwan  
Yusheng Ci, China  
Simone Cinquemani , Italy  
Roberto G. Citarella , Italy  
Joaquim Ciurana , Spain  
John D. Clayton , USA  
Piero Colajanni , Italy  
Giuseppina Colicchio, Italy  
Vassilios Constantoudis , Greece  
Enrico Conte, Italy  
Alessandro Contento , USA  
Mario Cools , Belgium  
Gino Cortellessa, Italy  
Carlo Cosentino , Italy  
Paolo Crippa , Italy  
Erik Cuevas , Mexico  
Guozeng Cui , China  
Mehmet Cunkas , Turkey  
Giuseppe D'Aniello , Italy  
Peter Dabnichki, Australia  
Weizhong Dai , USA  
Zhifeng Dai , China  
Purushothaman Damodaran , USA  
Sergey Dashkovskiy, Germany  
Adiel T. De Almeida-Filho , Brazil  
Fabio De Angelis , Italy  
Samuele De Bartolo , Italy  
Stefano De Miranda , Italy  
Filippo De Monte , Italy



































José António Fonseca De Oliveira  
Correia , Portugal  
Jose Renato De Sousa , Brazil  
Michael Defoort, France  
Alessandro Della Corte, Italy  
Laurent Dewasme , Belgium  
Sanku Dey , India  
Gianpaolo Di Bona , Italy  
Roberta Di Pace , Italy  
Francesca Di Puccio , Italy  
Ramón I. Diego , Spain  
Yannis Dimakopoulos , Greece  
Hasan Dinçer , Turkey  
José M. Domínguez , Spain  
Georgios Dounias, Greece  
Bo Du , China  
Emil Dumic, Croatia  
Madalina Dumitriu , United Kingdom  
Premraj Durairaj , India  
Saeed Eftekhar Azam, USA  
Said El Kafhali , Morocco  
Antonio Elipe , Spain  
R. Emre Erkmen, Canada  
John Escobar , Colombia  
Leandro F. F. Miguel , Brazil  
FRANCESCO FOTI , Italy  
Andrea L. Facci , Italy  
Shahla Faisal , Pakistan  
Giovanni Falsone , Italy  
Hua Fan, China  
Jianguang Fang, Australia  
Nicholas Fantuzzi , Italy  
Muhammad Shahid Farid , Pakistan  
Hamed Faruqi, Iran  
Yann Favennec, France  
Fiorenzo A. Fazzolari , United Kingdom  
Giuseppe Fedele , Italy  
Roberto Fedele , Italy  
Baowei Feng , China  
Mohammad Ferdows , Bangladesh  
Arturo J. Fernández , Spain  
Jesus M. Fernandez Oro, Spain  
Francesco Ferrise, Italy  
Eric Feulvarch , France  
Thierry Floquet, France

Eric Florentin , France  
Gerardo Flores, Mexico  
Antonio Forcina , Italy  
Alessandro Formisano, Italy  
Francesco Franco , Italy  
Elisa Francomano , Italy  
Juan Frausto-Solis, Mexico  
Shujun Fu , China  
Juan C. G. Prada , Spain  
HECTOR GOMEZ , Chile  
Matteo Gaeta , Italy  
Mauro Gaggero , Italy  
Zoran Gajic , USA  
Jaime Gallardo-Alvarado , Mexico  
Mosè Gallo , Italy  
Akemi Gálvez , Spain  
Maria L. Gandarias , Spain  
Hao Gao , Hong Kong  
Xingbao Gao , China  
Yan Gao , China  
Zhiwei Gao , United Kingdom  
Giovanni Garcea , Italy  
José García , Chile  
Harish Garg , India  
Alessandro Gasparetto , Italy  
Stylianios Georgantzinou, Greece  
Fotios Georgiades , India  
Parviz Ghadimi , Iran  
Ştefan Cristian Gherghina , Romania  
Georgios I. Giannopoulos , Greece  
Agathoklis Giaralis , United Kingdom  
Anna M. Gil-Lafuente , Spain  
Ivan Giorgio , Italy  
Gaetano Giunta , Luxembourg  
Jefferson L.M.A. Gomes , United Kingdom  
Emilio Gómez-Déniz , Spain  
Antonio M. Gonçalves de Lima , Brazil  
Qunxi Gong , China  
Chris Goodrich, USA  
Rama S. R. Gorla, USA  
Veena Goswami , India  
Xunjie Gou , Spain  
Jakub Grabski , Poland



Antoine Grall , France  
George A. Gravvanis , Greece  
Fabrizio Greco , Italy  
David Greiner , Spain  
Jason Gu , Canada  
Federico Guarracino , Italy  
Michele Guida , Italy  
Muhammet Gul , Turkey  
Dong-Sheng Guo , China  
Hu Guo , China  
Zhaoxia Guo, China  
Yusuf Gurefe, Turkey  
Salim HEDDAM , Algeria  
ABID HUSSANAN, China  
Quang Phuc Ha, Australia  
Li Haitao , China  
Petr Hájek , Czech Republic  
Mohamed Hamdy , Egypt  
Muhammad Hamid , United Kingdom  
Renke Han , United Kingdom  
Weimin Han , USA  
Xingsi Han, China  
Zhen-Lai Han , China  
Thomas Hanne , Switzerland  
Xinan Hao , China  
Mohammad A. Hariri-Ardebili , USA  
Khalid Hattaf , Morocco  
Defeng He , China  
Xiao-Qiao He, China  
Yanchao He, China  
Yu-Ling He , China  
Ramdane Hedjar , Saudi Arabia  
Jude Hemanth , India  
Reza Hemmati, Iran  
Nicolae Herisanu , Romania  
Alfredo G. Hernández-Díaz , Spain  
M.I. Herreros , Spain  
Eckhard Hitzer , Japan  
Paul Honeine , France  
Jaromir Horacek , Czech Republic  
Lei Hou , China  
Yingkun Hou , China  
Yu-Chen Hu , Taiwan  
Yunfeng Hu, China  
Can Huang , China  
Gordon Huang , Canada  
Linsheng Huo , China  
Sajid Hussain, Canada  
Asier Ibeas , Spain  
Orest V. Iftime , The Netherlands  
Przemyslaw Ignaciuk , Poland  
Giacomo Innocenti , Italy  
Emilio Insfran Pelozo , Spain  
Azeem Irshad, Pakistan  
Alessio Ishizaka, France  
Benjamin Ivorra , Spain  
Breno Jacob , Brazil  
Reema Jain , India  
Tushar Jain , India  
Amin Jajarmi , Iran  
Chiranjibe Jana , India  
Łukasz Jankowski , Poland  
Samuel N. Jator , USA  
Juan Carlos Jáuregui-Correa , Mexico  
Kandasamy Jayakrishna, India  
Reza Jazar, Australia  
Khalide Jbilou, France  
Isabel S. Jesus , Portugal  
Chao Ji , China  
Qing-Chao Jiang , China  
Peng-fei Jiao , China  
Ricardo Fabricio Escobar Jiménez , Mexico  
Emilio Jiménez Macías , Spain  
Maolin Jin, Republic of Korea  
Zhuo Jin, Australia  
Ramash Kumar K , India  
BHABEN KALITA , USA  
MOHAMMAD REZA KHEDMATI , Iran  
Viacheslav Kalashnikov , Mexico  
Mathiyalagan Kalidass , India  
Tamas Kalmar-Nagy , Hungary  
Rajesh Kaluri , India  
Jyotheeswara Reddy Kalvakurthi, India  
Zhao Kang , China  
Ramani Kannan , Malaysia  
Tomasz Kapitaniak , Poland  
Julius Kaplunov, United Kingdom  
Konstantinos Karamanos, Belgium  
Michal Kawulok, Poland

Irfan Kaymaz , Turkey  
Vahid Kayvanfar , Qatar  
Krzysztof Kecik , Poland  
Mohamed Khader , Egypt  
Chaudry M. Khalique , South Africa  
Mukhtaj Khan , Pakistan  
Shahid Khan , Pakistan  
Nam-Il Kim, Republic of Korea  
Philipp V. Kiryukhantsev-Korneev ,  
Russia  
P.V.V Kishore , India  
Jan Koci , Czech Republic  
Ioannis Kostavelis , Greece  
Sotiris B. Kotsiantis , Greece  
Frederic Kratz , France  
Vamsi Krishna , India  
Edyta Kucharska, Poland  
Krzysztof S. Kulpa , Poland  
Kamal Kumar, India  
Prof. Ashwani Kumar , India  
Michal Kunicki , Poland  
Cedrick A. K. Kwuimy , USA  
Kyandoghere Kyamakya, Austria  
Ivan Kyrchei , Ukraine  
Márcio J. Lacerda , Brazil  
Eduardo Lalla , The Netherlands  
Giovanni Lancioni , Italy  
Jaroslaw Latalski , Poland  
Hervé Laurent , France  
Agostino Lauria , Italy  
Aimé Lay-Ekuakille , Italy  
Nicolas J. Leconte , France  
Kun-Chou Lee , Taiwan  
Dimitri Lefebvre , France  
Eric Lefevre , France  
Marek Lefik, Poland  
Yaguo Lei , China  
Kauko Leiviskä , Finland  
Ervin Lenzi , Brazil  
ChenFeng Li , China  
Jian Li , USA  
Jun Li , China  
Yueyang Li , China  
Zhao Li , China






























Zhen Li , China  
En-Qiang Lin, USA  
Jian Lin , China  
Qibin Lin, China  
Yao-Jin Lin, China  
Zhiyun Lin , China  
Bin Liu , China  
Bo Liu , China  
Heng Liu , China  
Jianxu Liu , Thailand  
Lei Liu , China  
Sixin Liu , China  
Wanquan Liu , China  
Yu Liu , China  
Yuanchang Liu , United Kingdom  
Bonifacio Llamazares , Spain  
Alessandro Lo Schiavo , Italy  
Jean Jacques Loiseau , France  
Francesco Lolli , Italy  
Paolo Lonetti , Italy  
António M. Lopes , Portugal  
Sebastian López, Spain  
Luis M. López-Ochoa , Spain  
Vassilios C. Loukopoulos, Greece  
Gabriele Maria Lozito , Italy  
Zhiguo Luo , China  
Gabriel Luque , Spain  
Valentin Lychagin, Norway  
YUE MEI, China  
Junwei Ma , China  
Xuanlong Ma , China  
Antonio Madeo , Italy  
Alessandro Magnani , Belgium  
Toqeer Mahmood , Pakistan  
Fazal M. Mahomed , South Africa  
Arunava Majumder , India  
Sarfranz Nawaz Malik, Pakistan  
Paolo Manfredi , Italy  
Adnan Maqsood , Pakistan  
Muazzam Maqsood, Pakistan  
Giuseppe Carlo Marano , Italy  
Damijan Markovic, France  
Filipe J. Marques , Portugal  
Luca Martinelli , Italy  
Denizar Cruz Martins, Brazil

Francisco J. Martos , Spain  
Elio Masciari , Italy  
Paolo Massioni , France  
Alessandro Mauro , Italy  
Jonathan Mayo-Maldonado , Mexico  
Pier Luigi Mazzeo , Italy  
Laura Mazzola, Italy  
Driss Mehdi , France  
Zahid Mehmood , Pakistan  
Roderick Melnik , Canada  
Xiangyu Meng , USA  
Jose Merodio , Spain  
Alessio Merola , Italy  
Mahmoud Mesbah , Iran  
Luciano Mescia , Italy  
Laurent Mevel , France  
Constantine Michailides , Cyprus  
Mariusz Michta , Poland  
Prankul Middha, Norway  
Aki Mikkola , Finland  
Giovanni Minafò , Italy  
Edmondo Minisci , United Kingdom  
Hiroyuki Mino , Japan  
Dimitrios Mitsotakis , New Zealand  
Ardashir Mohammadzadeh , Iran  
Francisco J. Montáns , Spain  
Francesco Montefusco , Italy  
Gisele Mophou , France  
Rafael Morales , Spain  
Marco Morandini , Italy  
Javier Moreno-Valenzuela , Mexico  
Simone Morganti , Italy  
Caroline Mota , Brazil  
Aziz Moukrim , France  
Shen Mouquan , China  
Dimitris Mourtzis , Greece  
Emiliano Mucchi , Italy  
Taseer Muhammad, Saudi Arabia  
Ghulam Muhiuddin, Saudi Arabia  
Amitava Mukherjee , India  
Josefa Mula , Spain  
Jose J. Muñoz , Spain  
Giuseppe Muscolino, Italy  
Marco Mussetta , Italy

Hariharan Muthusamy, India  
Alessandro Naddeo , Italy  
Raj Nandkeolyar, India  
Keivan Navaie , United Kingdom  
Soumya Nayak, India  
Adrian Neagu , USA  
Erivelton Geraldo Nepomuceno , Brazil  
AMA Neves, Portugal  
Ha Quang Thinh Ngo , Vietnam  
Nhon Nguyen-Thanh, Singapore  
Papakostas Nikolaos , Ireland  
Jelena Nikolic , Serbia  
Tatsushi Nishi, Japan  
Shanzhou Niu , China  
Ben T. Nohara , Japan  
Mohammed Nouari , France  
Mustapha Nourelfath, Canada  
Kazem Nouri , Iran  
Ciro Núñez-Gutiérrez , Mexico  
Włodzimierz Ogryczak, Poland  
Roger Ohayon, France  
Krzysztof Okarma , Poland  
Mitsuhiro Okayasu, Japan  
Murat Olgun , Turkey  
Diego Oliva, Mexico  
Alberto Olivares , Spain  
Enrique Onieva , Spain  
Calogero Orlando , Italy  
Susana Ortega-Cisneros , Mexico  
Sergio Ortobelli, Italy  
Naohisa Otsuka , Japan  
Sid Ahmed Ould Ahmed Mahmoud , Saudi Arabia  
Taoreed Owolabi , Nigeria  
EUGENIA PETROPOULOU , Greece  
Arturo Pagano, Italy  
Madhumangal Pal, India  
Pasquale Palumbo , Italy  
Dragan Pamučar, Serbia  
Weifeng Pan , China  
Chandan Pandey, India  
Rui Pang, United Kingdom  
Jürgen Pannek , Germany  
Elena Panteley, France  
Achille Paolone, Italy

George A. Papakostas , Greece  
Xosé M. Pardo , Spain  
You-Jin Park, Taiwan  
Manuel Pastor, Spain  
Pubudu N. Pathirana , Australia  
Surajit Kumar Paul , India  
Luis Payá , Spain  
Igor Pažanin , Croatia  
Libor Pekař , Czech Republic  
Francesco Pellicano , Italy  
Marcello Pellicciari , Italy  
Jian Peng , China  
Mingshu Peng, China  
Xiang Peng , China  
Xindong Peng, China  
Yuexing Peng, China  
Marzio Pennisi , Italy  
Maria Patrizia Pera , Italy  
Matjaz Perc , Slovenia  
A. M. Bastos Pereira , Portugal  
Wesley Peres, Brazil  
F. Javier Pérez-Pinal , Mexico  
Michele Perrella, Italy  
Francesco Pesavento , Italy  
Francesco Petrini , Italy  
Hoang Vu Phan, Republic of Korea  
Lukasz Pieczonka , Poland  
Dario Piga , Switzerland  
Marco Pizzarelli , Italy  
Javier Plaza , Spain  
Goutam Pohit , India  
Dragan Poljak , Croatia  
Jorge Pomares , Spain  
Hiram Ponce , Mexico  
Sébastien Poncet , Canada  
Volodymyr Ponomaryov , Mexico  
Jean-Christophe Ponsart , France  
Mauro Pontani , Italy  
Sivakumar Poruran, India  
Francesc Pozo , Spain  
Aditya Rio Prabowo , Indonesia  
Anchasa Pramuanjaroenkij , Thailand  
Leonardo Primavera , Italy  
B Rajanarayan Prusty, India

Krzysztof Puszynski , Poland  
Chuan Qin , China  
Dongdong Qin, China  
Jianlong Qiu , China  
Giuseppe Quaranta , Italy  
DR. RITU RAJ , India  
Vitomir Racic , Italy  
Carlo Rainieri , Italy  
Kumbakonam Ramamani Rajagopal, USA  
Ali Ramazani , USA  
Angel Manuel Ramos , Spain  
Higinio Ramos , Spain  
Muhammad Afzal Rana , Pakistan  
Muhammad Rashid, Saudi Arabia  
Manoj Rastogi, India  
Alessandro Rasulo , Italy  
S.S. Ravindran , USA  
Abdolrahman Razani , Iran  
Alessandro Reali , Italy  
Jose A. Reinoso , Spain  
Oscar Reinoso , Spain  
Haijun Ren , China  
Carlo Renno , Italy  
Fabrizio Renno , Italy  
Shahram Rezapour , Iran  
Ricardo Rianza , Spain  
Francesco Riganti-Fulginei , Italy  
Gerasimos Rigatos , Greece  
Francesco Ripamonti , Italy  
Jorge Rivera , Mexico  
Eugenio Roanes-Lozano , Spain  
Ana Maria A. C. Rocha , Portugal  
Luigi Rodino , Italy  
Francisco Rodríguez , Spain  
Rosana Rodríguez López, Spain  
Francisco Rossomando , Argentina  
Jose de Jesus Rubio , Mexico  
Weiguo Rui , China  
Rubén Ruiz , Spain  
Ivan D. Rukhlenko , Australia  
Dr. Eswaramoorthi S. , India  
Weichao SHI , United Kingdom  
Chaman Lal Sabharwal , USA  
Andrés Sáez , Spain

Bekir Sahin, Turkey  
Laxminarayan Sahoo , India  
John S. Sakellariou , Greece  
Michael Sakellariou , Greece  
Salvatore Salamone, USA  
Jose Vicente Salcedo , Spain  
Alejandro Salcido , Mexico  
Alejandro Salcido, Mexico  
Nunzio Salerno , Italy  
Rohit Salgotra , India  
Miguel A. Salido , Spain  
Sinan Salih , Iraq  
Alessandro Salvini , Italy  
Abdus Samad , India  
Sovan Samanta, India  
Nikolaos Samaras , Greece  
Ramon Sancibrian , Spain  
Giuseppe Sanfilippo , Italy  
Omar-Jacobo Santos, Mexico  
J Santos-Reyes , Mexico  
José A. Sanz-Herrera , Spain  
Musavarah Sarwar, Pakistan  
Shahzad Sarwar, Saudi Arabia  
Marcelo A. Savi , Brazil  
Andrey V. Savkin, Australia  
Tadeusz Sawik , Poland  
Roberta Sburlati, Italy  
Gustavo Scaglia , Argentina  
Thomas Schuster , Germany  
Hamid M. Sedighi , Iran  
Mijanur Rahaman Seikh, India  
Tapan Senapati , China  
Lotfi Senhadji , France  
Junwon Seo, USA  
Michele Serpilli, Italy  
Silvestar Šesnić , Croatia  
Gerardo Severino, Italy  
Ruben Sevilla , United Kingdom  
Stefano Sfarra , Italy  
Dr. Ismail Shah , Pakistan  
Leonid Shaikhet , Israel  
Vimal Shanmuganathan , India  
Prayas Sharma, India  
Bo Shen , Germany  
Hang Shen, China

Xin Pu Shen, China  
Dimitri O. Shepelsky, Ukraine  
Jian Shi , China  
Amin Shokrollahi, Australia  
Suzanne M. Shontz , USA  
Babak Shotorban , USA  
Zhan Shu , Canada  
Angelo Sifaleras , Greece  
Nuno Simões , Portugal  
Mehakpreet Singh , Ireland  
Piyush Pratap Singh , India  
Rajiv Singh, India  
Seralathan Sivamani , India  
S. Sivasankaran , Malaysia  
Christos H. Skiadas, Greece  
Konstantina Skouri , Greece  
Neale R. Smith , Mexico  
Bogdan Smolka, Poland  
Delfim Soares Jr. , Brazil  
Alba Sofi , Italy  
Francesco Soldovieri , Italy  
Raffaele Solimene , Italy  
Yang Song , Norway  
Jussi Sopanen , Finland  
Marco Spadini , Italy  
Paolo Spagnolo , Italy  
Ruben Specogna , Italy  
Vasilios Spitas , Greece  
Ivanka Stamova , USA  
Rafał Stanisławski , Poland  
Miladin Stefanović , Serbia  
Salvatore Strano , Italy  
Yakov Strelniker, Israel  
Kangkang Sun , China  
Qiuqin Sun , China  
Shuaishuai Sun, Australia  
Yanchao Sun , China  
Zong-Yao Sun , China  
Kumarasamy Suresh , India  
Sergey A. Suslov , Australia  
D.L. Suthar, Ethiopia  
D.L. Suthar , Ethiopia  
Andrzej Swierniak, Poland  
Andras Szekrenyes , Hungary  
Kumar K. Tamma, USA



Yong (Aaron) Tan, United Kingdom  
Marco Antonio Taneco-Hernández , Mexico  
Lu Tang , China  
Tianyou Tao, China  
Hafez Tari , USA  
Alessandro Tasora , Italy  
Sergio Teggi , Italy  
Adriana del Carmen Téllez-Anguiano , Mexico  
Ana C. Teodoro , Portugal  
Efstathios E. Theotokoglou , Greece  
Jing-Feng Tian, China  
Alexander Timokha , Norway  
Stefania Tomasiello , Italy  
Gisella Tomasini , Italy  
Isabella Torricollo , Italy  
Francesco Tornabene , Italy  
Mariano Torrisi , Italy  
Thang nguyen Trung, Vietnam  
George Tsiatas , Greece  
Le Anh Tuan , Vietnam  
Nerio Tullini , Italy  
Emilio Turco , Italy  
Ilhan Tuzcu , USA  
Efstratios Tzirtzilakis , Greece  
FRANCISCO UREÑA , Spain  
Filippo Ubertini , Italy  
Mohammad Uddin , Australia  
Mohammad Safi Ullah , Bangladesh  
Serdar Ulubeyli , Turkey  
Mati Ur Rahman , Pakistan  
Panayiotis Vafeas , Greece  
Giuseppe Vairo , Italy  
Jesus Valdez-Resendiz , Mexico  
Eusebio Valero, Spain  
Stefano Valvano , Italy  
Carlos-Renato Vázquez , Mexico  
Martin Velasco Villa , Mexico  
Franck J. Vernerey, USA  
Georgios Veronis , USA  
Vincenzo Vespri , Italy  
Renato Vidoni , Italy  
Venkatesh Vijayaraghavan, Australia

Anna Vila, Spain  
Francisco R. Villatoro , Spain  
Francesca Vipiana , Italy  
Stanislav Vitek , Czech Republic  
Jan Vorel , Czech Republic  
Michael Vynnycky , Sweden  
Mohammad W. Alomari, Jordan  
Roman Wan-Wendner , Austria  
Bingchang Wang, China  
C. H. Wang , Taiwan  
Dagang Wang, China  
Guoqiang Wang , China  
Huaiyu Wang, China  
Hui Wang , China  
J.G. Wang, China  
Ji Wang , China  
Kang-Jia Wang , China  
Lei Wang , China  
Qiang Wang, China  
Qingling Wang , China  
Weiwei Wang , China  
Xinyu Wang , China  
Yong Wang , China  
Yung-Chung Wang , Taiwan  
Zhenbo Wang , USA  
Zhibo Wang, China  
Waldemar T. Wójcik, Poland  
Chi Wu , Australia  
Qihong Wu, China  
Yuqiang Wu, China  
Zhibin Wu , China  
Zhizheng Wu , China  
Michalis Xenos , Greece  
Hao Xiao , China  
Xiao Ping Xie , China  
Qingzheng Xu , China  
Binghan Xue , China  
Yi Xue , China  
Joseph J. Yame , France  
Chuanliang Yan , China  
Xinggang Yan , United Kingdom  
Hongtai Yang , China  
Jixiang Yang , China  
Mijia Yang, USA  
Ray-Yeng Yang, Taiwan


Zaoli Yang , China  
Jun Ye , China  
Min Ye , China  
Luis J. Yebra , Spain  
Peng-Yeng Yin , Taiwan  
Muhammad Haroon Yousaf , Pakistan  
Yuan Yuan, United Kingdom  
Qin Yuming, China  
Elena Zaitseva , Slovakia  
Arkadiusz Zak , Poland  
Mohammad Zakwan , India  
Ernesto Zambrano-Serrano , Mexico  
Francesco Zammori , Italy  
Jessica Zangari , Italy  
Rafal Zdunek , Poland  
Ibrahim Zeid, USA  
Nianyin Zeng , China  
Junyong Zhai , China  
Hao Zhang , China  
Haopeng Zhang , USA  
Jian Zhang , China  
Kai Zhang, China  
Lingfan Zhang , China  
Mingjie Zhang , Norway  
Qian Zhang , China  
Tianwei Zhang , China  
Tongqian Zhang , China  
Wenyu Zhang , China  
Xianming Zhang , Australia  
Xuping Zhang , Denmark  
Yinyan Zhang, China  
Yifan Zhao , United Kingdom  
Debao Zhou, USA  
Heng Zhou , China  
Jian G. Zhou , United Kingdom  
Junyong Zhou , China  
Xueqian Zhou , United Kingdom  
Zhe Zhou , China  
Wu-Le Zhu, China  
Gaetano Zizzo , Italy  
Mingcheng Zuo, China





## Contents



**Retracted: Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining**  
Mathematical Problems in Engineering  
Retraction (1 page), Article ID 9807263, Volume 2023 (2023)



**Retracted: Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries**  
Mathematical Problems in Engineering  
Retraction (1 page), Article ID 9863021, Volume 2023 (2023)



**Corrigendum to “Research on Economic Mathematical Analysis and Construction Model of Prefabricated Building Structure Based on Improved Neural Network Algorithm”**  
Xin Lin  
Corrigendum (1 page), Article ID 9831739, Volume 2021 (2021)




**Research on Economic Mathematical Analysis and Construction Model of Prefabricated Building Structure Based on Improved Neural Network Algorithm**  
Xin Lin   
Research Article (6 pages), Article ID 5362357, Volume 2021 (2021)

**On Partition Dimension of Some Cycle-Related Graphs**  
Changcheng Wei, Muhammad Faisal Nadeem , Hafiz Muhammad Afzal Siddiqui , Muhammad Azeem , Jia-Bao Liu , and Adnan Khalil  
Research Article (8 pages), Article ID 4046909, Volume 2021 (2021)


**Computing Analysis for First Zagreb Connection Index and Coindex of Resultant Graphs**  
Muhammad Javaid , Usman Ali, and Jia-Bao Liu   
Research Article (19 pages), Article ID 6019517, Volume 2021 (2021)

**A Network Flow Algorithm for Solving Generalized Assignment Problem**  
Yongwen Hu  and Qunpo Liu   
Research Article (8 pages), Article ID 5803092, Volume 2021 (2021)

**Three-Dimensional Numerical Calculation Model for Static Behavior Simulation of Cross-Laminated Timber Plates under Thermal Environment**  
Wenliang Hu , Wei Hou, Zhao Zhu , and Xuhui Huang  
Research Article (20 pages), Article ID 1538928, Volume 2021 (2021)




**On the Geodesic Identification of Vertices in Convex Plane Graphs**  
Fawaz E. Alsaadi, Muhammad Salman , Masood Ur Rehman , Abdul Rauf Khan, Jinde Cao , and Madini Obad Alassafi  
Research Article (13 pages), Article ID 7483291, Volume 2020 (2020)

**Algorithm Design of Early Warning Seatbelt Intelligent Adjustment System Based on Neural Network and Big Data Analysis**

Chunxu Zhou 



Research Article (7 pages), Article ID 7268963, Volume 2020 (2020)

**[Retracted] Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining**

Yun-Hua Xu, Lin-Fang Huang , Rong-Rong Guo , Xu-Yang Zhang, and Jia-Ming Zhu 


Research Article (10 pages), Article ID 9345901, Volume 2020 (2020)

**Construction of the Evaluation System of Sustainable Utilization of Large Stadiums Based on the AHP Method**

Li-Cai Zhu, Zhi Gao , Jia-Ming Zhu , and Di Zhang





Research Article (12 pages), Article ID 9396276, Volume 2020 (2020)

**Several Asymptotic Bounds on the Balaban Indices of Trees**

Bo Deng , Chengfu Ye, Weilin Liang, Yalan Li, and Xueli Su


Research Article (6 pages), Article ID 2081241, Volume 2020 (2020)

**[Retracted] Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries**

Ze-Jiong Zhou , Yao Wang , Miao-Miao Lu , and Jia-Ming Zhu 

Research Article (10 pages), Article ID 8723062, Volume 2020 (2020)

**The Influence of Government Intervention on the Performance of Independent Innovation under Financial Support Based on Data of Listed Companies in Strategic Emerging Industries**

Ze-Jiong Zhou, Peng-Yue Zhang, Miao-Miao Lu, and Zhi Gao 


Research Article (9 pages), Article ID 5063986, Volume 2020 (2020)

**Locating and Tracking Model for Language Radiation Transmission Based on Neural Network and FAHP**

SongGui Zhu, Hailang He , and Yuanyuan Zheng


Research Article (8 pages), Article ID 7625141, Volume 2020 (2020)

**Multiobjective Optimization Design of Toll Plaza**

Kai Wang , Peng Wang, Xin Chen, and Lu-Ting Zhao



Research Article (9 pages), Article ID 2324894, Volume 2020 (2020)

**The Approach of Hierarchical Linear Model to Exploring Individual and Team Creativity: A Perspective of Cultural Intelligence and Team Trust**

Chenhan Huang, Changqing He, and Xuesong Zhai 

Research Article (10 pages), Article ID 2025140, Volume 2020 (2020)



**The Development Law of the Freezing Temperature Field of a Calcareous Clay Layer**

Zhi Wang, Chuanxin Rong , Meng Du, Maoyan Ma, and Xiangyang Liu 


Research Article (12 pages), Article ID 3902054, Volume 2020 (2020)

# Contents


## **A Computing Model for Quantifying the Value of Structural Health Monitoring Information in Bridge Engineering**

Baoquan Cheng , Lijie Wang, Jianling Huang, Xu Shi, Xiaodong Hu, and Huihua Chen   
Research Article (7 pages), Article ID 8260909, Volume 2020 (2020)

## **Approximate Controllability for a Kind of Fractional Neutral Differential Equations with Damping**

Jun Du , Dongling Cui, Yeguo Sun, and Jin Xu  
Research Article (9 pages), Article ID 7592818, Volume 2020 (2020)



## **Locating-Total Domination Number of Cacti Graphs**

Jianxin Wei , Uzma Ahmad, Saira Hameed, and Javaria Hanif  
Research Article (10 pages), Article ID 6197065, Volume 2020 (2020)


## **Control of a Tilting Hexacopter under Wind Disturbance**

Hao Xu , Zhong Yang , Kaiwen Lu , Changliang Xu , and Qiuyan Zhang   
Research Article (17 pages), Article ID 9465153, Volume 2020 (2020)

## **The Cross-Entropy Method for the Winner Determination Problem in Combinatorial Auctions**

Hanmi Lin , Yongqiang Chen, Changping Liu, Jie Xie, and Baixiu Ni   
Research Article (7 pages), Article ID 5691719, Volume 2020 (2020)


## **Efficient Processing of Image Processing Applications on CPU/GPU**

Najia Naz, Abdul Haseeb Malik, Abu Bakar Khurshid, Furqan Aziz, Bader Alouffi, M. Irfan Uddin , and Ahmed ALGhamdi  
Research Article (14 pages), Article ID 4839876, Volume 2020 (2020)



## **Research on Management of Doctor-Patient Risk and Status of the Perceived Behaviors of Physician Trust in the Patient in China: New Perspective of Management of Doctor-Patient Risk**

Jiang Jie Sun , Zhi Bo Zheng, Xue Li Jiang, Wei Wei Hu, Jun Liu, Nan Zhen Ma, Meng Ying Li, Xiao-Zhe Yan, Cheng sen He, and Li Ping Zhang   
Research Article (8 pages), Article ID 2145029, Volume 2020 (2020)



## **Multimodal Semisupervised Deep Graph Learning for Automatic Precipitation Nowcasting**

Kaichao Miao, Wei Wang, Rui Hu, Lei Zhang, Yali Zhang, Xiang Wang, and Fudong Nian   
Research Article (9 pages), Article ID 4018042, Volume 2020 (2020)

## **Some New Bounds of Weighted Graph Entropies with GA and Gaurava Indices Edge Weights**


Tiejun Wu, Hafiz Mutee Ur Rehman , Yu-Ming Chu , Deeba Afzal, and Jianfeng Yu  
Research Article (9 pages), Article ID 5474501, Volume 2020 (2020)

## **Computing Edge Weights of Magic Labeling on Rooted Products of Graphs**

Jia-Bao Liu , Hafiz Usman Afzal, and Muhammad Javaid   
Research Article (16 pages), Article ID 2160104, Volume 2020 (2020)




### **Hermite–Hadamard Inequalities for Harmonic $(s, m)$ -Convex Functions**

Jian Zhong Xu, Umar Raza, Muhammad Waqas Javed, and Zaryab Hussain 

Research Article (7 pages), Article ID 1470837, Volume 2020 (2020)

### **Super-Efficiency Infeasibility in the Presence of Nonradial Measurement**

Dong Guo and Zheng-Qun Cai 

Research Article (7 pages), Article ID 6264852, Volume 2020 (2020)


### **Research on the Identification of Key Nodes in the Process of WeChat Epidemic Information**


#### **Dissemination: A Supernetwork Perspective**

Peng Wu  and Di Zhao

Research Article (10 pages), Article ID 6751686, Volume 2020 (2020)

### **Irregularity Measures for Benzene Ring Embedded in P-Type Surface**

Yun Liu, Aysha Siddiq, Yu-Ming Chu , Muhammad Azam, Muhammad Asim Raza Basra, and Abaid Ur

Rehman Virk 



Research Article (13 pages), Article ID 2462530, Volume 2020 (2020)

### **Computing Vertex-Based Eccentric Topological Descriptors of Zero-Divisor Graph Associated with Commutative Rings**

Abdullah Ali H. Ahmadini, Ali N. A. Koam , Ali Ahmad , Martin Bača , and Andrea Semaničová–Feňovčíková 



Research Article (6 pages), Article ID 2056902, Volume 2020 (2020)

### **Model Test and Numerical Simulation of the Mudcake Thickness Effect on the Bearing Capacity of Vertically Loaded Single Piles**

Ziguang Zhang , Jian Cui, Maoyan Ma , and Peisheng Xi


Research Article (11 pages), Article ID 5012427, Volume 2020 (2020)

### **Computation of Irregularity Indices of Certain Computer Networks**

Jiangnan Liu, Lulu Cai , Abaid ur Rehman Virk , Waheed Akhtar, Shahzad Ahmed Maitla, and Yang Wei


Research Article (17 pages), Article ID 2797286, Volume 2020 (2020)

### **Theoretical Calculation and Application Test of Lift Force for Ideal Electric Asymmetric Capacitor**

Jian-Guo Huang and Xiang-Yu Cheng 


Research Article (6 pages), Article ID 7230640, Volume 2020 (2020)

### **Irregularity Measures for Metal-Organic Networks**

Xuan Guo, Yu-Ming Chu , Muhammad Khalid Hashmi, Abaid Ur Rehman Virk , and Jingjing Li

Research Article (11 pages), Article ID 3978130, Volume 2020 (2020)

### **Algebraic Connectivity and Disjoint Vertex Subsets of Graphs**


Yan Sun and Faxu Li 

Research Article (6 pages), Article ID 5763218, Volume 2020 (2020)

## Contents

---

### **Risk Factors Discovery for Cancer Survivability Analysis Using Graph-Rule Mining**

Chaoyu Yang , Jie Yang, and Zhenyu Yang




Research Article (12 pages), Article ID 2384130, Volume 2020 (2020)

### **An Improved Mask R-CNN Model for Multiorgan Segmentation**

Jian-Hua Shu , Fu-Dong Nian , Ming-Hui Yu, and Xu Li 



Research Article (11 pages), Article ID 8351725, Volume 2020 (2020)

### **Construction for the Sequences of $Q$ -Borderenergetic Graphs**

Bo Deng , Caibing Chang, Haixing Zhao , and Kinkar Chandra Das 


Research Article (5 pages), Article ID 6176849, Volume 2020 (2020)

### **Research on Mathematical Model of Cost Budget in the Early Stage of Assembly Construction Project Based on Improved Neural Network Algorithm**

Xin Lin  and Yinan Lu 


Research Article (7 pages), Article ID 3674929, Volume 2020 (2020)

### **Multiple Positive Solutions for Fractional Three-Point Boundary Value Problem with $p$ -Laplacian Operator**

Dong Li, Yang Liu, and Chunli Wang 

Research Article (6 pages), Article ID 2327580, Volume 2020 (2020)

### **On Local Generalized Ulam–Hyers Stability for Nonlinear Fractional Functional Differential Equation**

Dongming Nie, Azmat Ullah Khan Niazi , and Bilal Ahmed

Research Article (12 pages), Article ID 3276873, Volume 2020 (2020)

## Retraction

# Retracted: Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining

### Mathematical Problems in Engineering

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their

agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y.-H. Xu, L.-F. Huang, R.-R. Guo, X.-Y. Zhang, and J.-M. Zhu, "Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining," *Mathematical Problems in Engineering*, vol. 2020, Article ID 9345901, 10 pages, 2020.

## Retraction

# Retracted: Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries

### Mathematical Problems in Engineering

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z.-J. Zhou, Y. Wang, M.-M. Lu, and J.-M. Zhu, "Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8723062, 10 pages, 2020.

## Corrigendum

# Corrigendum to “Research on Economic Mathematical Analysis and Construction Model of Prefabricated Building Structure Based on Improved Neural Network Algorithm”

**Xin Lin**

*School of Urban Construction Engineering, Chongqing Radio & TV University, Chongqing, China*

Correspondence should be addressed to Xin Lin; [linxin49840566@163.com](mailto:linxin49840566@163.com)

Received 17 August 2021; Accepted 17 August 2021; Published 6 September 2021

Copyright © 2021 Xin Lin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the article titled “Research on Economic Mathematical Analysis and Construction Model of Prefabricated Building Structure Based on Improved Neural Network Algorithm” [1], the Acknowledgments section should be corrected as follows.

This work was supported by the Chongqing Natural Science Foundation of China (cstc2020jcyj-msxmX0612) (name of the project: “The Seismic Performance of the Prefabricated ECC-RC Composite Pier”) and the Scientific and Technological Research Project of Chongqing Education Commission (KJQN201904003) (name of the project: “The Prefabricated ECC Support for Seismic Performance of Tube-Concrete Composite Bridge Piers”).

## References

- [1] X. Lin, “Research on economic mathematical analysis and construction model of prefabricated building structure based on improved neural network algorithm,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 5362357, 6 pages, 2021.



## Research Article

# Research on Economic Mathematical Analysis and Construction Model of Prefabricated Building Structure Based on Improved Neural Network Algorithm

Xin Lin 

*School of Urban Construction Engineering, Chongqing Radio & TV University, Chongqing, China*

Correspondence should be addressed to Xin Lin; [linxin49840566@163.com](mailto:linxin49840566@163.com)

Received 13 August 2020; Revised 26 November 2020; Accepted 21 January 2021; Published 17 April 2021

Academic Editor: Jia-Bao Liu

Copyright © 2021 Xin Lin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this article, a mathematical analysis model of economics of prefabricated building structure based on improved neural network algorithm is proposed in order to solve the low analysis accuracy in traditional methods. Firstly, by means of analyzing the costs of materials, labor, and equipment, the economic characteristics of the cost of fabricated building structures are determined. Secondly, the single neuron is analyzed and the weight coefficient is adjusted in accordance with the multilayer neural network structure, so as to minimize the construction error of the economic analysis model of the assembled building structure. Meanwhile, the weight vector is obtained, error-weighted square sum is calculated through choosing an adaptive filter and obtained, and the weight vector is updated by the least squares algorithm. Thirdly, the neural network algorithm training and learning process is designed and improved, the dependent variable is selected, the number of input points is determined, and then, the training and learning process of the improved neural network algorithm is completed. Finally, a fitness function is set to measure the authenticity of dataset, which is further defined as a combination of different weights to construct an economic mathematical analysis model. The experimental results indicate that the analysis results of this method can reach an accuracy up to 96%, so it has a broader application prospect in low-rise buildings.

## 1. Introduction

Nowadays, as social progress, the prefabricated building structures with low capital, simple structure, and good performance are in urgent need [1]. Compared with traditional concrete building structures, the buildings adopting prefabricated structural systems are better in terms of antiearthquake strength, firefighting level, corrosion resistance, and construction waste production [2]. Meanwhile, by virtue of a shorter construction period, the prefabricated structure buildings not only can greatly reduce the construction cost but also is more practical than the concrete building structure [3]. Considering the large number of floors, large area, and high manpower as well as material resources required for superhigh-rise buildings [4], the structure cost accounts for 1/3 of the total price. Therefore, it is of great significance to conduct economic analysis on the

fabricated building structure [5]. Although prefabricated building structures are advantageous in low carbon emissions, green environmental protection, and sustainable development, there are only a few researches focusing on the economic mathematical analysis of prefabricated building structures [6]. In response to this problem, the research from the perspective of structural economy is carried out based on the example of actual project [7]. As an artificial intelligence technology, neural network has received many experts' studies in the field of economic analysis, and a good analysis result is achieved [8]. However, there are still some problems such as insensitive to the initial spatial conditions of high dimensions and multiple local minima, so it is necessary to search for global data points by using as few control parameters as possible [9]. To this end, a mathematical analysis model for the economics of fabricated building structures based on improved neural network algorithms is

constructed, which can provide engineering companies with a scientific basis for grasping future economic development trends [10].

## 2. Economic Characteristics of Fabricated Building Structure

The cost of prefabricated building structure is mainly composed of direct and indirect engineering costs, in which the direct engineering costs include labor costs, material costs, equipment costs, site management fees, and construction costs [11]; indirect engineering costs contain material expenses, utilities, maintenance fees, transportation fees, installation fees, and measures fees [12].

*2.1. Material Expenses.* By fully taking into account the lifting requirements, prefabricated building structure generally maximizes the amount of components so that the amount of building materials has also increased [13]. In order to save money to the greatest extent during the production process, prefabricated building structure uses a higher turnover template to replace the wooden template. This means the consumption of prefabricated building structure materials and material expenses will be greatly reduced if the standardization level of the prefabricated building structure and the turnover frequency of the formwork are improved [14].

*2.2. Labor Costs.* The labor costs of prefabricated building structures mainly include labor costs in the production process, transportation costs, and labor costs in installation. Due to the continuous development of industrial production measures and processes, concrete building structures that require longer working hours for workers are gradually reduced [15].

*2.3. Equipment Costs.* Equipment costs are the major factor impeding the economic progress of fabricated building structures. The diversification of China's current buildings has generated a higher demand for fabricated building structures, which results in an increase in the cost of traditional concrete building structures, the one-time investment in equipment, and the pressure of equipment cost recovery [16].

## 3. Improved Neural Network Algorithm

The structure of multilayer neural network is shown in Figure 1.

The neuron structure of the neural network is shown in Figure 2.

The neural network aims to adjust the weight coefficients to ensure a minimized construction error of the economic mathematical analysis model of fabricated building structure [17]. Each neuron is connected by a linear combination and a nonlinear function. Since the linear combination has the same structure as the adaptive filter, the least squares

algorithm is applicable to update the weight vector [18]. The specific process is as follows.

First, an adaptive filter is selected, and the weight vector is obtained as follows:

$$Q(n) = [Q_0(n), Q_1(n) \Delta Q_{n-1}(n)]^T. \quad (1)$$

The data vector is input:

$$f(n) = [f(n), f(n-1) \Delta f(n+1)]^T. \quad (2)$$

The objective function is the error weighted sum of squares, and the specific calculation content is as follows:

$$Z(n) = \sum_{i=1}^n \alpha^{n-1} (d(i) - Q(n)f(i))^2. \quad (3)$$

In formula (3),  $\alpha$  represents the forgetting factor and  $d(i)$  represents the expected output value.

In this algorithm, the expected value output by using the neuron is affected by the multilayer structure, and it is unable to measure the expected response of the middle layer neuron, so the error backpropagation shall be adopted to calculate the expected response value [19].

The specific steps of the improved neural network algorithm are as follows:

Step 1: the weights are initialized, and the learning step length is determined to ensure that the forgetting factor is within the error range.

Step 2: the output value of the linear combination of each layer of the network is calculated according to the forward propagation of input mode.

Step 3: the weights are updated through the least square algorithm, and the output error is checked to complete the improvement of neural network algorithm.

## 4. Application of Improved Neural Network Algorithm in Mathematical Analysis Model

*4.1. Training and Learning.* The training and learning process of the improved neural network algorithm is shown in Figure 3.

- (1) The dependent variable is selected according to the actual situation, and the variables that have an impact on the dependent variable are screened out through self-organizing method so as to determine the number of input points.
- (2) The number of output layer nodes is set, and the learning accuracy, the number of iteration steps, the upper limit of hidden nodes, learning parameters, and momentum coefficients are initialized.
- (3) The learning sample is input to ensure that the sample value is within the specified threshold range.
- (4) The values are randomly generated within the specified threshold range, and initial weight matrix is assigned.
- (5) The entire network is trained according to the improved neural network algorithm.

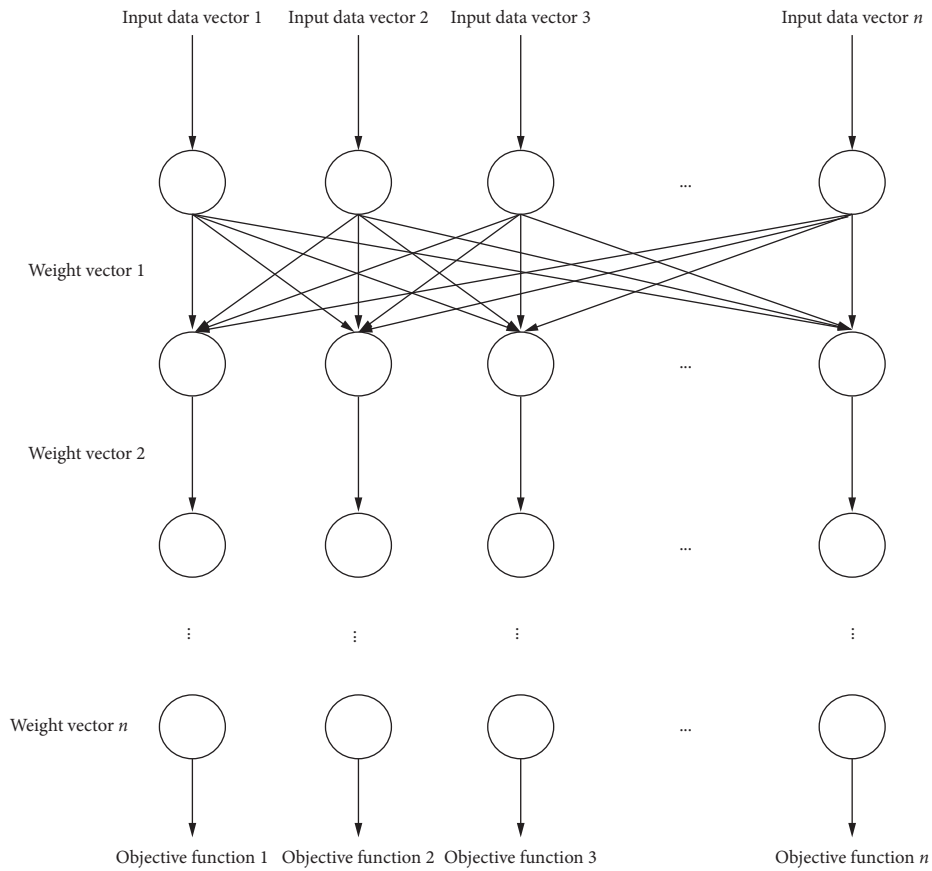


FIGURE 1: Structure of multilayer neural network.

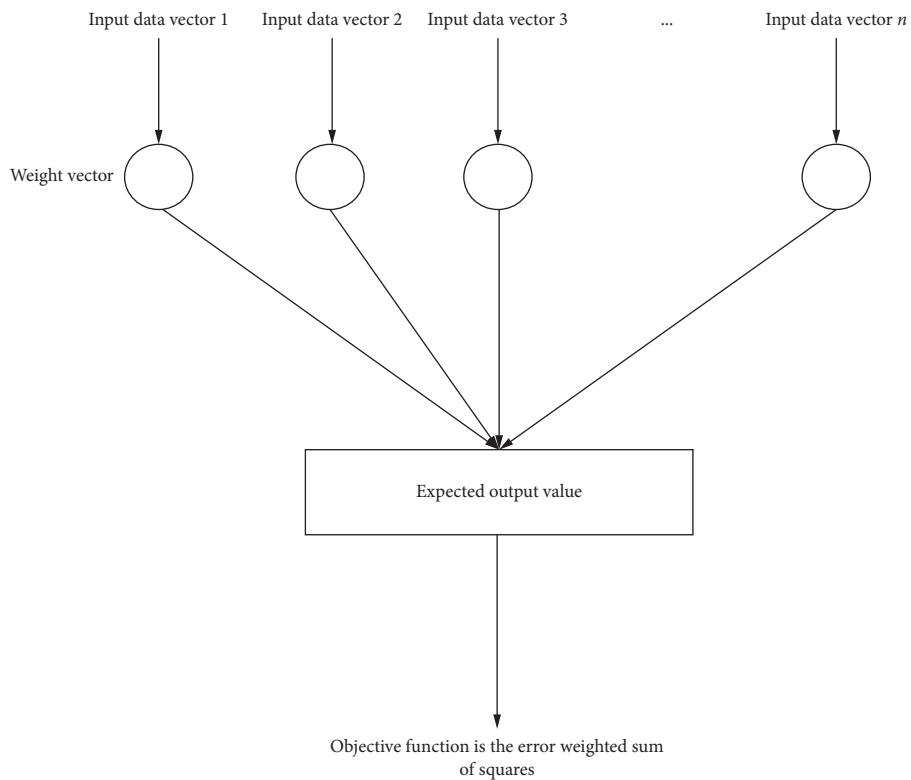


FIGURE 2: Neuron structure of neural network.

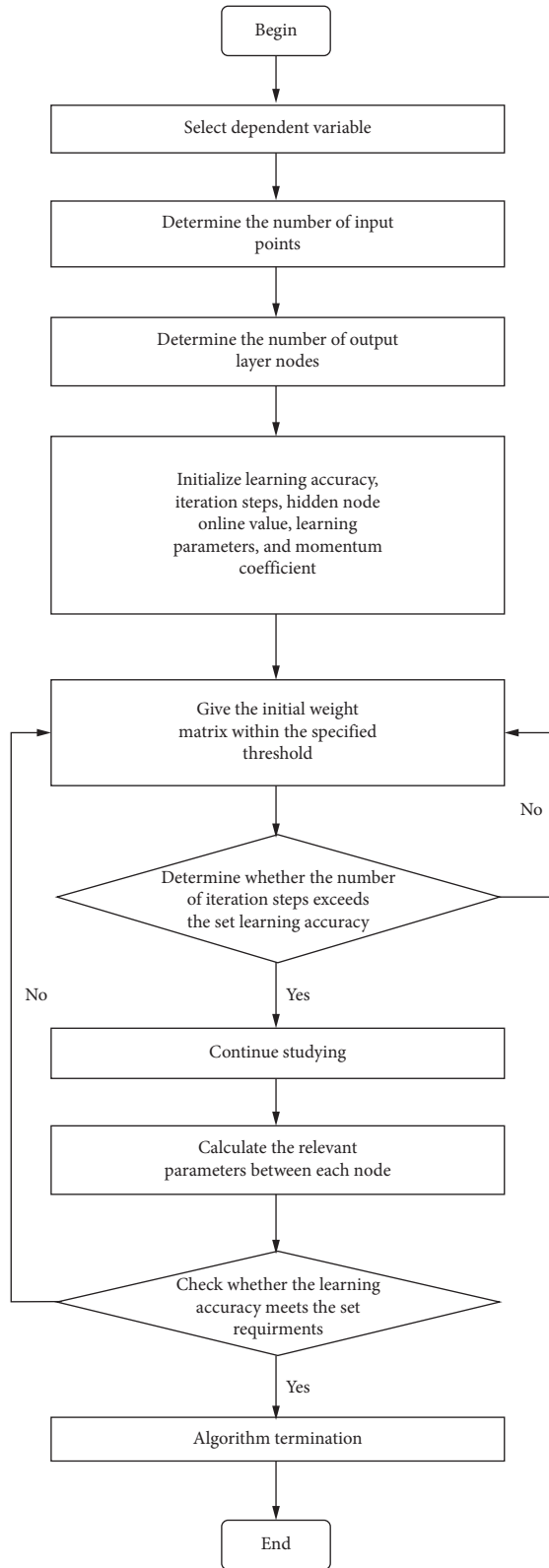


FIGURE 3: Training and learning process of improved neural network algorithm.

- (6) Judge whether the number of iteration steps exceeds the set learning accuracy; if yes, continue learning; otherwise, return to the previous step.

- (7) Calculate the relevant parameters between nodes and check whether the learning accuracy meets the set requirements. If yes, terminate the algorithm; otherwise, return to step (4) to complete the training and learning of the improved neural network algorithm.

4.2. *Model Building.* In the economic research of the fabricated building structure, the specific application steps of using the improved neural network algorithm in the mathematical analysis construction model are as follows:

- (1) First, an initial dataset containing around 10–20 data is constructed. The length of the individual is the sum of the weight and the bias value of the improved neural network algorithm. The sum is multiplied by the number of digits of the improved neural network algorithm code so that the gene of a single data can be randomly selected within a predefined range.
- (2) A fitness function is set to measure the authenticity of dataset. Accordingly, the fitness function is defined as a combination of different weights to generate a mean square error between the generated network output and the actual output of the corresponding sample; the two errors are superimposed to obtain the sum, and then, the economic mathematical analysis model is constructed as follows:

$$F = \frac{\max G - \sum_i (Y_i - X_i)^2}{\max O}. \quad (4)$$

In formula (4),  $i$  represents the number of samples,  $\max G$  represents the maximum fitness value of the neural network,  $\max O$  represents the maximum error that the neural network may output,  $Y_i$  represents the sample output value after neural network learning, and  $X_i$  represents the actual output value. The closer the neural network fitness value is to the maximum value, the higher the value of  $\max G$ , indicating that the dataset is real.

- (3) The original dataset is selected to generate a new dataset, and the new dataset is added to another dataset.
- (4) A single dataset is brought into the improved neural network algorithm after gene encoding so as to calculate the fitness of a single individual. If certain individual meets the set conditions, the application of the improved neural network algorithm ends.
- (5) Arrange according to the utility degree of a single dataset to remove redundant datasets while retaining the needed datasets.
- (6) Return to (3) and repeat continuously until all needed data are obtained.

The economic dataset of the fabricated building structure is determined based on all the acquired data so that the mathematical analysis is completed.

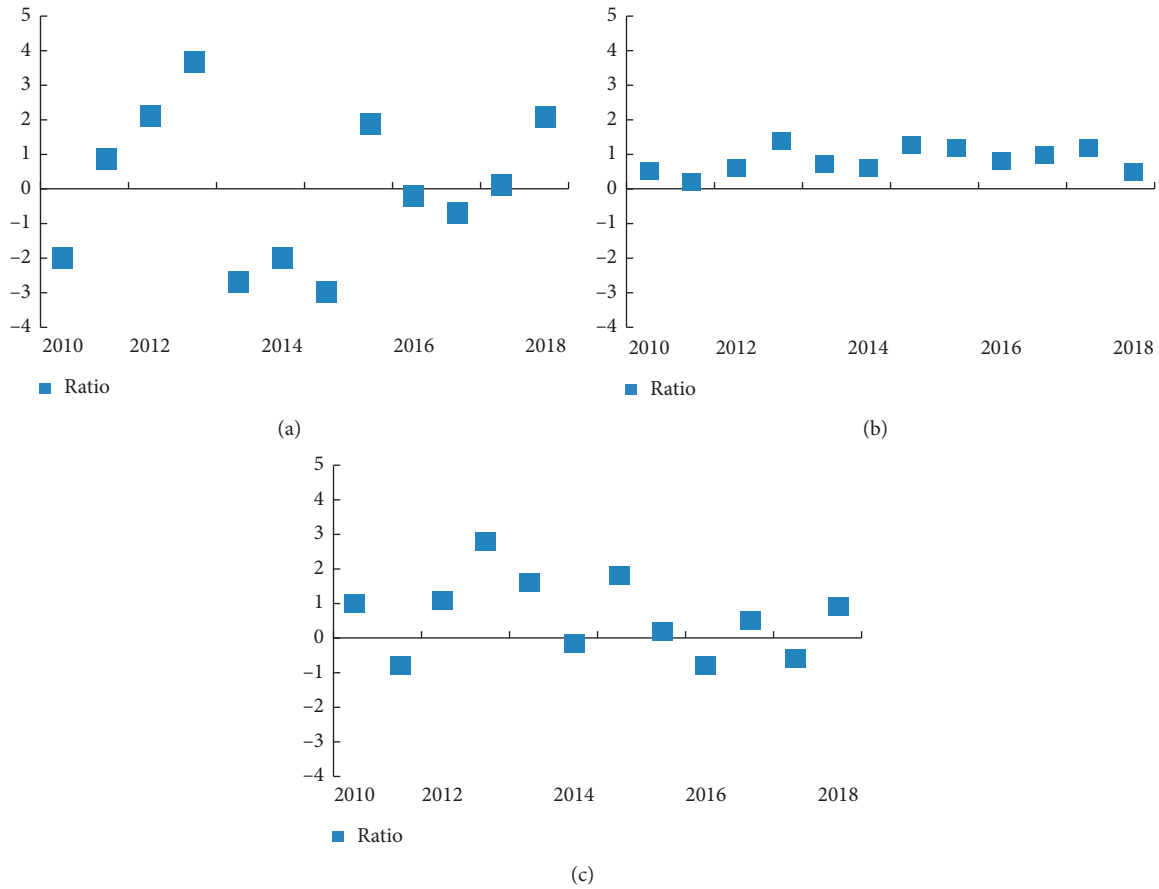


FIGURE 4: Relationship between added value of construction industry and other variables: (a) ratio to relative growth rate of construction area; (b) ratio to relative growth rate of total output of building industry; (c) ratio to relative growth rate of practitioners.

TABLE 1: Comparison results of analysis accuracy between two methods.

Time/year	Gross output/million yuan	Algorithm based on neural network (%)	Algorithm based on improved neural network (%)
2010	105744.1	38	96
2012	151395.1	43	95
2014	552896.3	30	96
2016	1238761.5	41	96
2018	2675431.6	39	95

### 5. Experimental Analysis

The actual experiments are performed so as to verify the rationality of studying the economic mathematical analysis and construction model of prefabricated building structure based on the improved neural network algorithm.

5.1. Data Analysis. Based on the prefabricated building structure in a city, the relationship between the added value of the construction industry and other variables from 2010 to 2018 is analyzed.

Figure 4(a) shows that, during 2010–2018, the added value of the construction industry has a greater impact on the construction area, and the construction area is also increasing steadily at this time; Figure 4(b) indicates that, during 2010–2018, the relative growth rate between the

added value of construction industry and the total output value is stable; as shown in Figure 4(c), during 2010–2018, the relative growth rate between construction industry’s added value and the number of employees is not stable.

5.2. Experimental Results and Analysis. As can be learned from the above content, the added value of the city’s construction industry mainly depends on the growth of total output value but is less related to the construction area and the number of employees. On account of this phenomenon, the traditional neural network algorithm and the analysis method based on the improved neural network algorithm are used, respectively, for comparative analysis. The results are as follows.

According to Table 1, the analysis accuracy of traditional neural network algorithm is always lower than 45%, while

the analysis accuracy of improved neural network algorithm is always higher than 95%. It can be seen that it is reasonable to research economic mathematical analysis and construction model of prefabricated building structure based on improved neural network algorithm.

## 6. Conclusions

With the rapid economic development, the stability and economics of traditional concrete building structures cannot meet the needs of modernization. Therefore, the research is based on an improved neural network algorithm for mathematical analysis and construction of fabricated building structures. The improved neural network algorithm can obtain a satisfactory economic analysis structure and provides a new effective way for future economic forecasting because it has strong generalization ability.

Despite the relatively accurate economic analysis results of this model, there are still some problems, such as less compressive analysis result due to the failure in considering the seismic performance of building structure. Therefore, the future research in this field shall focus on this aspect so as to widen the application prospects.

## Data Availability

Simulation data and our model and related parameters used are provided within the article.

## Conflicts of Interest

The author declares that there are no conflicts of interest.

## Acknowledgments

This work was supported by the Chongqing Municipal Education Commission Science and Technology Research Project (KJQN201904003) and Research Project of Chongqing Technology and Business Institute (ZD2016-04).

## References

- [1] G.-H. Kim, J.-E. Yoon, S.-H. An, H.-H. Cho, and K.-I. Kang, "Neural network model incorporating a genetic algorithm in estimating construction costs," *Building and Environment*, vol. 39, no. 11, pp. 1333–1340, 2004.
- [2] H. M. Günaydın and S. Z. Doğan, "A neural network approach for early cost estimation of structural systems of buildings," *International Journal of Project Management*, vol. 22, no. 7, pp. 595–602, 2004.
- [3] O. Tatari and M. Kucukvar, "Cost premium prediction of certified green buildings: a neural network approach," *Building and Environment*, vol. 46, no. 5, pp. 1081–1086, 2011.
- [4] M.-Y. Cheng, H.-C. Tsai, and E. Sudjono, "Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry," *Expert Systems with Applications*, vol. 37, no. 6, pp. 4224–4231, 2010.
- [5] M. W. Emsley, D. J. Lowe, A. R. Duff, A. Harding, and A. Hickson, "Data modelling and the application of a neural network approach to the prediction of total construction costs," *Construction Management and Economics*, vol. 20, no. 6, pp. 465–472, 2002.
- [6] G.-H. Kim, J.-M. Shin, S. Kim, and Y. Shin, "Comparison of school building construction costs estimation methods using regression analysis, neural network, and support vector machine," *Journal of Building Construction and Planning Research*, vol. 1, no. 1, pp. 1–7, 2013.
- [7] G. H. Kim, D. S. Seo, and K. I. Kang, "Hybrid models of neural networks and genetic algorithms for predicting preliminary cost estimates," *Journal of Computing in Civil Engineering*, vol. 19, no. 2, pp. 208–211, 2005.
- [8] M. Juszczak, "The use of artificial neural networks for residential buildings conceptual cost estimation," *AIP Conference Proceedings*, vol. 1558, no. 1, pp. 1302–1306, 2013.
- [9] B. B. Ekici and U. T. Aksoy, "Prediction of building energy consumption by using artificial neural networks," *Advances in Engineering Software*, vol. 40, no. 5, pp. 356–362, 2009.
- [10] M. Juszczak, "Application of committees of neural networks for conceptual cost estimation of residential buildings," *AIP Conference Proceedings*, vol. 1648, no. 1, Article ID 600008, 2015.
- [11] C. M. Tam and C. F. Fang, "Comparative cost analysis of using high-performance concrete in tall building construction by artificial neural networks," *Structural Journal*, vol. 96, no. 6, pp. 927–936, 1999.
- [12] G.-H. Kim, S.-H. An, and K.-I. Kang, "Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning," *Building and Environment*, vol. 39, no. 10, pp. 1235–1242, 2004.
- [13] W. J. Zhu, W. F. Feng, and Y. G. Zhou, "The application of genetic fuzzy neural network in project cost estimate," in *Proceedings of the 2010 International Conference on E-Product E-Service and E-Entertainment*, pp. 1–4, IEEE, Henan, China, November 2010.
- [14] M. Juszczak and A. Leśniak, "Modelling construction site cost index based on neural network ensembles," *Symmetry*, vol. 11, no. 3, p. 411, 2019.
- [15] H.-G. Cho, K.-G. Kim, J.-Y. Kim, and G.-H. Kim, "A comparison of construction cost estimation using multiple regression analysis and neural network in elementary school project," *Journal of the Korea Institute of Building Construction*, vol. 13, no. 1, pp. 66–74, 2013.
- [16] V. Chandanshive and A. R. Kambekar, "Estimation of building construction cost using artificial neural networks," *Journal of Soft Computing in Civil Engineering*, vol. 3, no. 1, pp. 91–107, 2019.
- [17] N. I. El-Sawalhi and O. Shehatto, "A neural network model for building construction projects cost estimating," *A Neural Network Model for Building Construction Projects Cost Estimating*, vol. 4, no. 4, 2014.
- [18] U. Gulcicek, O. Ozkan, M. Gunduz, and I. H. Demir, "Cost assessment of construction projects through neural networks," *Canadian Journal of Civil Engineering*, vol. 40, no. 6, pp. 574–579, 2013.
- [19] V. R. Ambrule and A. N. Bhirud, "Use of artificial neural network for pre design cost estimation of building projects," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 5, no. 2, pp. 173–176, 2017.



## Research Article

# On Partition Dimension of Some Cycle-Related Graphs

Changcheng Wei,<sup>1</sup> Muhammad Faisal Nadeem ,<sup>2</sup> Hafiz Muhammad Afzal Siddiqui ,<sup>2</sup>  
Muhammad Azeem ,<sup>2,3</sup> Jia-Bao Liu ,<sup>4</sup> and Adnan Khalil<sup>2</sup>

<sup>1</sup>Department of Mathematics and Computer Science, Tongling University, Tongling, Anhui, China

<sup>2</sup>Department of Mathematics, COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan

<sup>3</sup>Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, Seri Kembangan, Malaysia

<sup>4</sup>School of Mathematics and Physics, Anhui Jianzhu University, Hefei 230601, China

Correspondence should be addressed to Muhammad Faisal Nadeem; [mfaisalnadeem@ymail.com](mailto:mfaisalnadeem@ymail.com)

Received 21 August 2020; Revised 28 December 2020; Accepted 22 February 2021; Published 8 March 2021

Academic Editor: Isabella Torricollo

Copyright © 2021 Changcheng Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Let  $G$  be a simple connected graph. Suppose  $\Delta = \{\Delta_1, \Delta_2, \dots, \Delta_l\}$  an  $l$ -partition of  $V(G)$ . A partition representation of a vertex  $\alpha$  w.r.t  $\Delta$  is the  $l$ -vector  $(d(\alpha, \Delta_1), d(\alpha, \Delta_2), \dots, d(\alpha, \Delta_l))$ , denoted by  $r(\alpha|\Delta)$ . Any partition  $\Delta$  is referred as resolving partition if  $\forall \alpha_i \neq \alpha_j \in V(G)$  such that  $r(\alpha_i|\Delta) \neq r(\alpha_j|\Delta)$ . The smallest integer  $l$  is referred as the partition dimension  $pd(G)$  of  $G$  if the  $l$ -partition  $\Delta$  is a resolving partition. In this article, we discuss the partition dimension of kayak paddle graph, cycle graph with chord, and a graph generated by chain of cycles. It has been shown that the partition dimension of the said families of graphs is constant.

## 1. Introduction and Preliminaries

The concept of partition dimension is a natural generalization of metric dimension. It was proposed in [1]. This concept came from the study of metric dimension which was defined independently in [2, 3]. This parameter of a graph was proved NP-complete problem given [4]. The metric dimension of a connected graph is based on the distance between vertices, while partition dimension is based on the distance between a vertex and a set containing some vertices. Many researchers worked on this topic. A significant number of papers were published and some of them are given in [5–9]. The problem of finding partition dimension of a connected graph is still unsolved. There are only some lower and upper bounds for the partition dimension of general connected graph available in the literature and the exact values are still open. An upper bound for the partition dimension of a tree is given in [10]. The partition dimension of certain classes of graphs is given in [11–13].

Graph theory is very vast field of applied and computational mathematics. That is why most of the applied sciences field extensively uses the graph theory. The partition dimension is also considered an applied topic of graph theory, and some of them are Djokovic–Winkler relation [14], strategies for the mastermind game [15], network discovery and verification [16], and in chemistry to represent the chemical compounds [17, 18]. In digital world, it is used to recognize the pattern, in robotics, it used for image processing, and it also plays a key role for the management of hierarchical data structures [19]. However, the applications of partition are still limited because the computational cost to compute the partition dimension is very complex. Some other applications of this concept is the navigation of robots in networks, and for interest, few areas appear in these literature [2, 20].

The partition dimension of a graph can be constant, but some graphs have bounded. Here, are some literature works. Ahmed et al. [21] computed the metric dimension of kayak paddle graph and cycle graph with chords and

proved that adding an edge in a cycle graph will not affect the metric dimension of a cycle graph. Mehreen et al. [5] computed the partition dimension of fullerene graph is 3. Rajan et al. [7] described the constant partition dimension of hexagonal and honeycomb networks. Partition dimension of certain honeycomb-derived networks is computed in [6]. The partition dimension of some wheel-related graphs is computed in [22]. Rodriguez-Velázquez et al. [8, 9] worked on the trees and unicyclic graphs and computed the bounds on the partition dimension. For more details, we refer to [23–26].

The distance between two vertices  $v_1, v_2 \in V(G)$  is a shortest path between them and is denoted by  $d(v_1, v_2)$ . Let  $\Delta = \{\Delta_1, \dots, \Delta_l\}$  be an  $l$ -partition of the vertex set  $V(G)$  and  $r(v|\Delta) = \{d(v, \Delta_1), d(v, \Delta_2), \dots, d(v, \Delta_a)\}$  be an  $l$ -tuple representing a unique code of  $v$  w.r.t to  $\Delta$ . If all representation codes of the vertex set of the graph  $G$  are unique w.r.t to  $\Delta$ , then  $\Delta$  is a resolving partition. This partition with the minimum value of  $l$  is referred as partition dimension of  $G$  and is denoted by  $pd(G)$ .

This paper deals with the partition dimension of the kayak paddle graph, a cycle graph with a chord, and a graph generated by a chain of cycles. More precisely, we computed the partition dimension of said families of graphs and it turns out to be constant, i.e., 3.

Following theorems are very helpful in finding the partition dimension of a connected graph.

**Theorem 1** (see [1]). *Let  $\Pi$  be a resolving partition of  $V(G)$  of a graph  $G$  and  $u, v \in V(G)$ . If  $d(u, w) = d(v, w)$  for all vertices  $w \in V(G) \setminus \{u, v\}$ , then  $u$  and  $v$  belong to different classes of  $\Pi$ .*

**Theorem 2** (see [1]). *Let  $G$  be a simple and connected graph of order  $n$ ; then,*

- (i)  $pd(G)$  is 2 iff  $G$  is a path graph
- (ii)  $pd(G)$  is  $n$  iff  $G$  is a complete graph

The remaining article is managed as follows. In Section 2, the partition dimension of kayak paddle graph is computed, which is constant for all three parameters given in the definition. In Section 3, the partition dimension of cycle graph with chord is discussed, and in Section 4, we study partition dimension of a family of graph generated by chain of cycles. In the end, conclusion and references have been given.

## 2. Kayak Paddle Graph

This section deals with the graph generated by cycle, known as kayak paddle graph, denoted by  $KP(x, y, z)$ . This family of graphs can be constructed by considering two cycles of length  $x$  and  $y$  and a path of length  $z$  with  $x, y \geq 3$  and  $z \geq 2$ . The vertex set and edge set of kayak paddle graph are as follows, respectively:

$$\begin{aligned} V(KP(x, y, z)) &= \{\alpha_1, \alpha_2, \dots, \alpha_x\} \cup \{\beta_1, \beta_2, \dots, \beta_y\} \cup \{\gamma_1, \dots, \gamma_{z-1}\}, \\ E(KP(x, y, z)) &= \{\alpha_i \alpha_{i+1} : 1 \leq i \leq x\} \cup \{\beta_i \beta_{i+1} : 1 \leq i \leq y\} \cup \{\gamma_i \gamma_{i+1} : 1 \leq i \leq z-2\} \cup \{\alpha_1 \gamma_1, \gamma_{z-1} \beta_1\}, \end{aligned} \quad (1)$$

where  $\alpha_{x+1} = \alpha_1$  and  $\beta_1 = \beta_{y+1}$ .

Figure 1 shows a kayak paddle graph  $KP(12, 8, 5)$  with two cycles of length  $x = 12$  and  $y = 8$  and a path of length  $z = 5$  joining them.

**2.1. Partition Dimension of Kayak Paddle Graph.** In the following result, partition dimension of the kayak paddle graph generated by two cycles and a path graph has been discussed.

**Theorem 3.** *Let  $G = KP(x, y, z)$  be a kayak paddle graph with  $x \geq 3$ ,  $y \geq 4$ , and  $z \geq 2$ . Then,  $pd(KP(x, y, z)) = 3$ .*

*Proof.* To show that  $pd(KP(x, y, z)) = 3$ , we use double inequality. First of all, we prove that  $pd(KP(x, y, z)) \leq 3$  by constructing a resolving partition with three elements. For this, consider the following cases on  $x$  and  $y$ :

- (i) Case 1: when  $x = 2s$  and  $y = 2t$  and  $s, t \geq 2$ . We claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving set, where  $\Delta_1 = \{\alpha_s\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ , where  $q$  is any of the vertex of the type  $\alpha_i, \beta_i$ , or  $\gamma_i$  other than the vertices of  $\Delta_1$  and  $\Delta_2$ . Then, Table 1 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  vertices with respect to resolving partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = s$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.
- (iii) Case 2: when  $x = 2s + 1$ ,  $y = 2t$  and  $t \geq 2$ . Then, we split  $x$  into the following three subcases:

Subcase 2.1: for  $x = 3, 5$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition with  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 2 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = 2$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

Subcase 2.2: for  $x = 7$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition with  $\Delta_1 = \{\alpha_3\}$ ,  $\Delta_2 = \{\beta_t\}$ , and



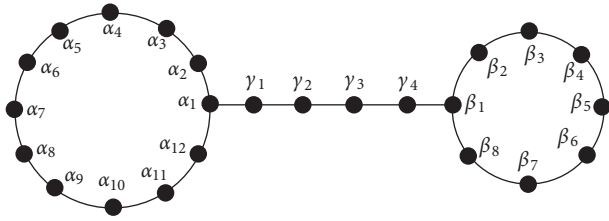


FIGURE 1: A kayak paddle graph:  $KP(12, 8, 5)$ .

TABLE 1: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: 1 \leq \varepsilon \leq s$	$s - \varepsilon$	$t + z - 2 + \varepsilon$	$a$
$\alpha_\varepsilon: s + 1 \leq \varepsilon \leq x$	$\varepsilon - s$	$2s + t + z - \varepsilon$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$s + z - 2 + \varepsilon$	$t - \varepsilon$	$b$
$\beta_\varepsilon: t + 1 \leq \varepsilon \leq x$	$2t + s + z - \varepsilon$	$\varepsilon - t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$s + \varepsilon - 1$	$t - \varepsilon + z - 1$	$0$

TABLE 2: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$t + z - 2 + \varepsilon$	$a$
$\alpha_\varepsilon: \varepsilon = x = 3$	$1$	$t + z$	$0$
$\alpha_\varepsilon: \varepsilon = 3, 4, x = 5$	$\varepsilon - 2$	$t + z + 1$	$0$
$\alpha_\varepsilon: \varepsilon = x = 5$	$2$	$t + z$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$z + \varepsilon$	$t - \varepsilon$	$b$
$\beta_\varepsilon: t + 1 \leq \varepsilon \leq y$	$2t + 2 + z - \varepsilon$	$\varepsilon - t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$1 + \varepsilon$	$t - \varepsilon + z - 1$	$0$

$\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 3 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = 3$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

Subcase 2.3: for  $x = 2s + 1$  and  $s \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition  $\Delta_1 = \{\alpha_s\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 4 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = s$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

- (iv) Case 3: when  $x = 2s + 1$  and  $y = 2t + 1$ , then we split  $x$  and  $y$  in the following 9 subcases:

Subcase 3.1: for  $x = 3$  and  $y = 3, 5$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition with  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_2\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 5 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = b = 1$  when  $\varepsilon = 2$  and  $a, b = 0$ , otherwise.

Subcase 3.2: for  $x = 3$  and  $y = 7$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition with  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_3\}$ ,  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 6 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$

TABLE 3: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2, 3$	$3 - \varepsilon$	$t + z - 2 + \varepsilon$	$a$
$\alpha_\varepsilon: \varepsilon = 4, 5$	$\varepsilon - 3$	$t + z + 2$	$0$
$\alpha_\varepsilon: \varepsilon = 6, 7$	$3$	$t + z - \varepsilon + 7$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$z + \varepsilon + 1$	$t - \varepsilon$	$b$
$\beta_\varepsilon: t + 1 \leq \varepsilon \leq y$	$2t + z - \varepsilon + 3$	$\varepsilon - t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$\varepsilon + 2$	$t - \varepsilon + z - 1$	$0$

TABLE 4: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: 1 \leq \varepsilon \leq s$	$s - \varepsilon$	$t + z - 2 + \varepsilon$	$a$
$\alpha_\varepsilon: \varepsilon = s + 1, s + 2$	$\varepsilon - s$	$s + t + z - 1$	$0$
$\alpha_\varepsilon: s + 3 \leq \varepsilon \leq 2s - 1$	$\varepsilon - s$	$2s + t + z - \varepsilon + 1$	$0$
$\alpha_\varepsilon: 2s \leq \varepsilon \leq x$	$s$	$2s + t + z - \varepsilon + 1$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$s + z + \varepsilon - 2$	$t - \varepsilon$	$b$
$\beta_\varepsilon: t + 1 \leq \varepsilon \leq y$	$2t + s + z - \varepsilon$	$\varepsilon - t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$\varepsilon + s - 1$	$t - \varepsilon + z - 1$	$0$

TABLE 5: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$z + \varepsilon$	$a$
$\alpha_\varepsilon: \varepsilon = 3$	$1$	$z + 2$	$0$
$\beta_\varepsilon: \varepsilon = 1, 2$	$z + \varepsilon$	$2 - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = y = 3$	$z + 2$	$1$	$0$
$\beta_\varepsilon: \varepsilon = 3, 4, y = 5$	$z + 3$	$\varepsilon - 2$	$0$
$\beta_\varepsilon: \varepsilon = y = 5$	$z + 2$	$2$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$\varepsilon + 1$	$1 - \varepsilon + z$	$0$

TABLE 6: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$z + \varepsilon + 1$	$a$
$\alpha_\varepsilon: \varepsilon = 3$	$1$	$z + 3$	$0$
$\beta_\varepsilon: \varepsilon = 1, 2, 3$	$z + \varepsilon$	$3 - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = 4, 5$	$z + 4$	$\varepsilon - 3$	$0$
$\beta_\varepsilon: \varepsilon = 6, 7$	$z + 9 - \varepsilon$	$3$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$\varepsilon + 1$	$2 - \varepsilon + z$	$0$

when  $\varepsilon = 2$ ,  $b = 1$  when  $\varepsilon = 3$ , and  $a, b = 0$ , otherwise.

Subcase 3.3: for  $x = 3$  and  $y = 2t + 1$  and  $t \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition, where  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 7 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = 2$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

Subcase 3.4: for  $x = 5$  and  $y = 5$  and  $t \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition, where  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_2\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 8 shows different representation of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ ,

TABLE 7: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$t + z + \varepsilon - 2$	$a$
$\alpha_\varepsilon: \varepsilon = 3$	$1$	$t + z$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$z + \varepsilon$	$t - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = t + 1, t + 2$	$t + z + 1$	$\varepsilon - t$	$0$
$\beta_\varepsilon: t + 3 \leq \varepsilon \leq 2t - 1$	$2t + z - \varepsilon + 3$	$\varepsilon - t$	$0$
$\beta_\varepsilon: 2t \leq \varepsilon \leq y$	$2t + z - \varepsilon + 3$	$t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$1 + \varepsilon$	$t - \varepsilon + z - 1$	$0$

TABLE 8: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$z + \varepsilon$	$a$
$\alpha_\varepsilon: \varepsilon = 3, 4$	$\varepsilon - 1$	$z + 3$	$0$
$\alpha_\varepsilon: \varepsilon = 5$	$2$	$z + 2$	$0$
$\beta_\varepsilon: \varepsilon = 1, 2$	$z + \varepsilon$	$2 - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = 3, 4$	$z + 3$	$\varepsilon - 2$	$0$
$\beta_\varepsilon: \varepsilon = 5$	$2 + z$	$2$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$1 + \varepsilon$	$z - \varepsilon + 1$	$0$

where  $a = b = 1$  when  $\varepsilon = 2$  and otherwise  $a, b$  are 0.

Subcase 3.5: for  $x = 5$  and  $y = 7$  and  $t \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition, where  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_3\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 9 shows different representation of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = 2$ ,  $b = 1$  when  $\varepsilon = 3$ , and  $a, b = 0$ , otherwise.

Subcase 3.6: for  $x = 5$  and  $y = 2t + 1$  and  $t \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition, where  $\Delta_1 = \{\alpha_2\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 10 shows different representation of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = 2$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

Subcase 3.7: for  $x = 7$  and  $y = 7$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition, where  $\Delta_1 = \{\alpha_3\}$ ,  $\Delta_2 = \{\beta_3\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 11 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = b = 1$  when  $\varepsilon = 3$  and  $a, b = 0$ , otherwise.

Subcase 3.8: for  $x = 7$  and  $y = 2t + 1$  and  $t \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving with  $\Delta_1 = \{\alpha_3\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 12 shows different representations of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = 3$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

Subcase 3.9: for  $x = 2s + 1$  and  $y = 2t + 1$  and  $s, t \geq 4$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving 3-partition with  $\Delta_1 = \{\alpha_s\}$ ,  $\Delta_2 = \{\beta_t\}$ , and  $\Delta_3 = \{q \in V(G) | q \notin \Delta_1, \Delta_2\}$ . Then, Table 13 shows different representation of  $\alpha_\varepsilon, \beta_\varepsilon$ , and  $\gamma_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = s$ ,  $b = 1$  when  $\varepsilon = t$ , and  $a, b = 0$ , otherwise.

TABLE 9: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$z + \varepsilon + 1$	$a$
$\alpha_\varepsilon: \varepsilon = 3, 4$	$\varepsilon - 1$	$z + 4$	$0$
$\alpha_\varepsilon: \varepsilon = 5$	$2$	$z + 3$	$0$
$\beta_\varepsilon: \varepsilon = 1, 2, 3$	$z + \varepsilon$	$3 - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = 4, 5$	$z + 4$	$\varepsilon - 3$	$0$
$\beta_\varepsilon: \varepsilon = 5$	$z + 9 - \varepsilon$	$3$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$1 + \varepsilon$	$z - \varepsilon + 2$	$0$

TABLE 10: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2$	$2 - \varepsilon$	$t + z + \varepsilon - 2$	$a$
$\alpha_\varepsilon: \varepsilon = 3, 4$	$\varepsilon - 1$	$t + z + 1$	$0$
$\alpha_\varepsilon: \varepsilon = 5$	$2$	$z + t$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$z + \varepsilon$	$t - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = t + 1, t + 2$	$t + z + 1$	$\varepsilon - t$	$0$
$\beta_\varepsilon: t + 3 \leq \varepsilon \leq 2t - 1$	$2t + z - \varepsilon + 3$	$t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$1 + \varepsilon$	$z - \varepsilon + 1$	$0$

TABLE 11: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2, 3$	$3 - \varepsilon$	$z + \varepsilon + 1$	$a$
$\alpha_\varepsilon: \varepsilon = 4, 5$	$\varepsilon - 3$	$z + 5$	$0$
$\alpha_\varepsilon: \varepsilon = 6, 7$	$3$	$z + 10 - \varepsilon$	$0$
$\beta_\varepsilon: \varepsilon = 1, 2, 3$	$z + 1 + \varepsilon$	$3 - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = 4, 5$	$z + 5$	$\varepsilon - 3$	$0$
$\beta_\varepsilon: \varepsilon = 6, 7$	$z + 10 - \varepsilon$	$3$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$2 + \varepsilon$	$z - \varepsilon + 2$	$0$

TABLE 12: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: \varepsilon = 1, 2, 3$	$3 - \varepsilon$	$t + z + \varepsilon - 2$	$a$
$\alpha_\varepsilon: \varepsilon = 4, 5$	$\varepsilon - 3$	$t + z + 2$	$0$
$\alpha_\varepsilon: \varepsilon = 6, 7$	$3$	$t + z + \varepsilon + 7$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$z + 1 + \varepsilon$	$t - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = t + 1, t + 2$	$t + z + 2$	$\varepsilon - t$	$0$
$\beta_\varepsilon: t + 3 \leq \varepsilon \leq 2t - 1$	$2t + z - \varepsilon + 4$	$\varepsilon - t$	$0$
$\beta_\varepsilon: 2t \leq \varepsilon \leq y$	$2t + z - \varepsilon + 4$	$t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$2 + \varepsilon$	$t + z - \varepsilon - 1$	$0$

TABLE 13: Distance codes for the vertices of  $KP(x, y, z)$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$\alpha_\varepsilon: 1 \leq \varepsilon \leq s$	$s - \varepsilon$	$t + z + \varepsilon - 2$	$a$
$\alpha_\varepsilon: \varepsilon = s + 1, s + 2$	$\varepsilon - s$	$t + s + z - 1$	$0$
$\alpha_\varepsilon: s + 3 \leq \varepsilon \leq 2s - 1$	$\varepsilon - s$	$2s + t + z - \varepsilon + 1$	$0$
$\alpha_\varepsilon: 2s \leq \varepsilon \leq x$	$s$	$2s + t + z - \varepsilon + 1$	$0$
$\beta_\varepsilon: 1 \leq \varepsilon \leq t$	$s + z - 1 + \varepsilon$	$t - \varepsilon$	$b$
$\beta_\varepsilon: \varepsilon = t + 1, t + 2$	$s + 2t + z - \varepsilon + 1$	$\varepsilon - t$	$0$
$\beta_\varepsilon: t + 3 \leq \varepsilon \leq 2t - 1$	$s + 2t + z - \varepsilon + 1$	$\varepsilon - t$	$0$
$\beta_\varepsilon: 2t \leq \varepsilon \leq y$	$s + 2t + z - \varepsilon + 1$	$t$	$0$
$\gamma_\varepsilon: 1 \leq \varepsilon \leq z - 1$	$s - 1 + \varepsilon$	$t + z - \varepsilon - 1$	$0$

Since all the vertices  $\alpha_\varepsilon$ ,  $\beta_\varepsilon$ , and  $\gamma_\varepsilon$  have different representations with respect to the partition  $\Delta$ , we obtain

$$\text{pd}(\text{KP}(x, y, z)) \leq 3. \quad (2)$$

*Conversely.* To prove that  $\text{pd}(\text{KP}(x, y, z)) \geq 3$ , suppose, on contrary,  $\text{pd}(\text{KP}(x, y, z)) = 2$ . It is not possible because  $\text{pd}(G) = 2$  if and only if  $G$  is a path graph. Our assumption is wrong. Therefore,

$$\text{pd}(\text{KP}(x, y, z)) \geq 3. \quad (3)$$

Hence, from inequalities (2) and (3), we conclude that

$$\text{pd}(\text{KP}(x, y, z)) = 3. \quad (4)$$

### 3. Partition Dimension of Cycle with Chord Graph

In this section, we consider a cycle with chord graph, denoted by  $C_n^m$ . This graph is constructed by joining any two nonadjacent vertices of the cycle. A graph of cycle with chord with vertex set  $V(C_n^m) = \{v_1, v_2, \dots, v_n\}$  and  $E(C_n^m) = \{v_i v_{i+1} : 1 \leq i \leq n\} \cup \{v_1 v_{m+1}\}$  is shown in Figure 2. To find the partition dimension of cycle with chord graph  $C_n^m$ , it suffices to consider  $t \leq \lfloor (n/2) \rfloor$  for given value of  $n$ .

The partition dimension of the cycle with chord graph is given in the following result.

**Theorem 4.** *Let  $C_n^m$  be a cycle with chord graph,  $n \geq 4$  and  $2 \leq m \leq n - 2$ . Then,  $\text{pd}(C_n^m) = 3$ .*

*Proof.* To show that  $\text{pd}(C_n^m) = 3$ , first, we will prove  $\text{pd}(C_n^m) \leq 3$  by constructing a resolving partition having 3 elements. For this, we split  $n$  and  $m$  into the following two cases:

- (i) Case 1: for  $n = 2t$  and  $m = 2s$ , we claim  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition with  $\Delta_1 = \{v_{t+s}\}$ ,  $\Delta_2 = \{v_{t+s+1}\}$ , and  $\Delta_3 = \{v_\varepsilon \in V(G) | v_\varepsilon \notin \Delta_1, \Delta_2\}$ . Then, Table 14 shows different representations of  $v_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = t + s$  and  $t + s + 1$  and 0, otherwise.
- (iii) Case 2: when  $n = 2t + 1$  and  $m = 2s + 1$ , we claim that  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$  is a resolving partition with  $\Delta_1 = \{v_{t+s+1}\}$ ,  $\Delta_2 = \{v_{t+s+2}\}$ , and  $\Delta_3 = \{v_\varepsilon \in V(G) | v_\varepsilon \notin \Delta_1, \Delta_2\}$ . Then, Table 15 shows different representations of  $v_\varepsilon$  with respect to the partition  $\Delta$ , where  $a = 1$  when  $\varepsilon = t + s + 1$  and  $t + s + 2$  and 0, otherwise.

Case 2.1: when  $n = 2t$  and  $m = 2s + 1$ , where  $a = 1$  when  $\varepsilon = t + s + 1$  and  $t + s + 2$  and 0, otherwise. Table 16 shows different representations of  $v_\varepsilon$  with respect to the partition  $\Delta$ .

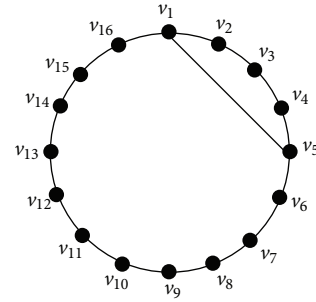


FIGURE 2: A cycle with chord graph  $C_{16}^4$ .

TABLE 14: Distance codes for the vertices of  $C_n^m$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$v_\varepsilon : 1 \leq \varepsilon \leq s$	$t - s - 1 + \varepsilon$	$t - s + \varepsilon - 1$	0
$v_\varepsilon : s + 1 \leq \varepsilon \leq t + s$	$t + s - \varepsilon$	$t + s - \varepsilon + 1$	$a$
$v_\varepsilon : t + s + 1 \leq \varepsilon \leq n$	$\varepsilon - t - s$	$\varepsilon - t - s - 1$	$a$

TABLE 15: Distance codes for the vertices of  $C_n^m$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$v_\varepsilon : 1 \leq \varepsilon \leq s + 1$	$t - s - 1 + \varepsilon$	$t - s + \varepsilon - 1$	0
$v_\varepsilon : s + 2 \leq \varepsilon \leq t + s + 1$	$t + s - \varepsilon + 1$	$t + s - \varepsilon + 2$	$a$
$v_\varepsilon : t + s + 2 \leq \varepsilon \leq n$	$\varepsilon - t - s - 1$	$\varepsilon - t - s - 2$	$a$

TABLE 16: Distance codes for the vertices of  $C_n^m$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$v_\varepsilon : 1 \leq \varepsilon \leq s + 1$	$t - s - 1 + \varepsilon$	$t - s + \varepsilon - 2$	0
$v_\varepsilon : s + 2 \leq \varepsilon \leq t + s + 1$	$t + s - \varepsilon + 1$	$t + s - \varepsilon + 2$	$a$
$v_\varepsilon : t + s + 2 \leq \varepsilon \leq n$	$\varepsilon - t - s - 1$	$\varepsilon - t - s - 2$	$a$

Case 2.2: when  $n = 2t + 1$  and  $m = 2s$ , where  $a = 1$  when  $\varepsilon = t + s + 1$  and  $t + s + 2$  and 0, otherwise. Table 17 shows different representations of  $v_\varepsilon$  with respect to the partition  $\Delta$ .

Since all the vertices of  $C_n^m$  have different representations with respect to the partition  $\Delta$ , therefore,

$$\text{pd}(C_n^m) \leq 3. \quad (5)$$

*Conversely.* To prove that  $\text{pd}(C_n^m) \geq 3$ , suppose, on the contrary,  $\text{pd}(C_n^m) = 2$ . It is not possible because  $\text{pd}(G) = 2$  if and only if  $G$  is a path graph. Therefore,

$$\text{pd}(C_n^m) \geq 3. \quad (6)$$

Hence, from inequalities (5) and (6), we can conclude that

TABLE 17: Distance codes for the vertices of  $C_n^m$  w.r.t  $\Delta$ .

$d(\cdot, \cdot)$	$\Delta_1$	$\Delta_2$	$\Delta_3$
$v_\varepsilon: 1 \leq \varepsilon \leq s$	$t - s + \varepsilon$	$t - s + \varepsilon - 1$	0
$v_\varepsilon: \varepsilon = s + 1$	$t$	$t$	0
$v_\varepsilon: s + 2 \leq \varepsilon \leq t + s + 1$	$t + s - \varepsilon + 1$	$t + s - \varepsilon + 2$	$a$
$v_\varepsilon: t + s + 2 \leq \varepsilon \leq n$	$\varepsilon - t - s - 1$	$\varepsilon - t - s - 2$	$a$

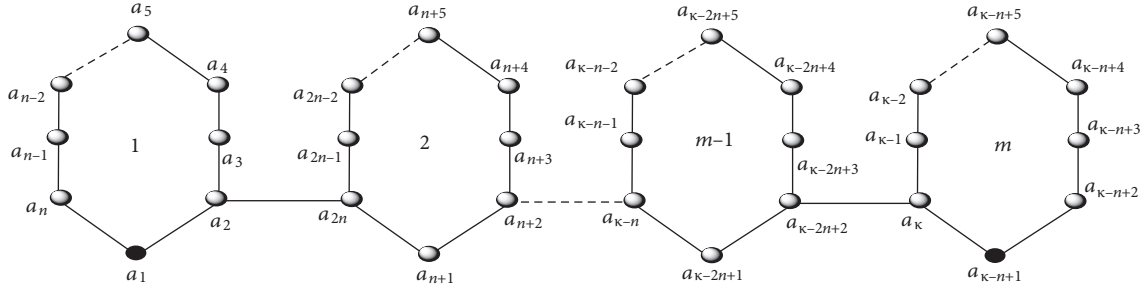


FIGURE 3: Chain of cycles' graph  $CC_n^m$ .

TABLE 18: Distance codes for  $a_k$  w.r.t  $\Delta_1$ .

$d(\cdot, \cdot)$	$\Delta_1$
$a_k: k = 1, 2, 3, \dots, \lceil (n + 1/2) \rceil$	$k - 1$
$a_k: k = \lceil (n + 1/2) \rceil + 1, \lceil (n + 1/2) \rceil + 2, \dots, n$	$n - k + 1$
$a_k: k \equiv 1, 2, 3, \dots, \lfloor (n - 1/2) \rfloor \pmod{n}$ and $k \geq n + 1$	$3 \lfloor (k/n) \rfloor + k - 1 - n \lfloor (k/n) \rfloor$
$a_k: k \equiv 0, \lfloor (n - 1/2) \rfloor + 1, \lfloor (n - 1/2) \rfloor + 2, \dots, n - 1 \pmod{n}$ and $n + 1 \leq k \leq 2n$	$n \lfloor (k/n) \rfloor + 2 - k$
$a_k: k \equiv 0, \lfloor (n - 1/2) \rfloor + 1, \lfloor (n - 1/2) \rfloor + 2, \dots, n - 1 \pmod{n}$ and $k > 2n$	$n \lfloor (k + \lfloor (n - 1/2) \rfloor + 1/n) \rfloor - k + 2 + 3(\lfloor (k + \lfloor (n - 1/2) \rfloor + 1/n) \rfloor - 2)$

TABLE 19: Distance codes for  $a_k$  w.r.t  $\Delta_2$ ,  $m = 2$ .

$d(\cdot, \cdot)$	$\Delta_1$
$a_k: k = 1$	3
$a_k: k = 2, 3, \dots, \lfloor (n + 2/2) \rfloor$	$k$
$a_k: k = \lfloor (n + 2/2) \rfloor + 1, \dots, n$	$n - k + 4$

TABLE 20: Distance codes for  $a_k$  w.r.t  $\Delta_2$ ,  $m \geq 3$ .

$d(\cdot, \cdot)$	$\Delta_2$
$a_k: k = 1$	$3(\lfloor (n(m - 1) + 1/n) \rfloor - 1) + k$
$a_k: k = 2, 3, \dots, \lfloor (n + 2/2) \rfloor$	$3(\lfloor (k/n) \rfloor - 1) + k$
$a_k: k = \lfloor (n + 2/2) \rfloor + 1, \dots, n$ and $k \geq n + 1$	$n - k + 4 + 3(\lfloor (n(m - 1) + 1/n) \rfloor - 1)$

$$\text{pd}(C_n^m) = 3. \tag{7}$$

□

### 4. Cycles' Chain Graph

This section deals with the graph generated by taking and joining  $m$  copies of cycles and  $m - 1$  copies of path of length 1 alternately. This graph is denoted by  $CC_n^m$  with vertex set  $V(G) = \{a_1, a_2, \dots, a_{nm}\}$  and edge set  $E(G) = \{a_i a_{i+1} : 1 \leq i \leq n\} \cup \{a_{(m-1)n+i} a_{(m-1)n+i+1} : m \geq 2\} \cup \{a_{(j-1)n+2} a_{(j+1)n} : 1 \leq j \leq m - 1\}$ , where

$m$  represents the number of cycles and  $n$  represents the number of vertices in a cycle, and for each cycle,  $a_{mn+1} = a_{(m-1)n+1}$ , i.e., for  $m = 1$  and  $a_{n+1} = a_1$  and for  $m = 2$   $a_{2n+1} = a_{n+1}$ . The order and size of the chain graph is  $\kappa = |V(CC_n^m)| = nm$  and  $|E(CC_n^m)| = m(n + 1) - 1$ , respectively. Figure 3 shows chain of cycles' graph  $CC_n^m$ .

4.1. Partition Dimension of Chain of Cycles' Graph  $CC_n^m$ . In this section, we study the partition dimension of graph generated by chain of cycles  $CC_n^m$ . The following theorem presents the partition dimension of the graph  $CC_n^m$ .

TABLE 21: Distance codes for  $a_k$  w.r.t  $\Delta_2$ ,  $m \geq 2$ .

$d(\cdot, \cdot)$	$\Delta_2$
$a_k: k = n(m-1) + 1, n(m-1) + 2, \dots, n(m-1) + \lfloor (n/2) \rfloor$	$k - n(m-1) - 1$
$a_k: k = n(m-1) + \lfloor (n/2) \rfloor + 1, n(m-1) + \lfloor (n/2) \rfloor + 2, \dots, nm$	$nm - k + 1$
$a_k: k \equiv 2, 3, \dots, \lfloor (n+2/2) \rfloor \pmod{n}$ and $n+1 \leq k \leq n(m-1)$	$3(m - \lfloor (k-1/n) \rfloor - 2) + k - n \lfloor (k/n) \rfloor$
$a_k: k \equiv 0, \lfloor (n+2/2) \rfloor + 1, \dots, n-1 \pmod{n}$ and $n+1 \leq k \leq n(m-1)$	$3(m - \lfloor (k-1/n) \rfloor - 2) + 4 - k + n \lfloor (k + \lfloor (n+2/2) \rfloor / n) \rfloor$
$a_k: k \equiv 1 \pmod{n}$ and $n+1 \leq k \leq n(m-1)$	$3(m - \lfloor (k/n) \rfloor - 1)$

**Theorem 5.** *If  $n \geq 5$  and  $m \geq 2$ , then  $pd(CC_n^m) = 3$ .*

*Proof.* First, we prove  $pd(CC_n^m) \leq 3$  by constructing a resolving partition set  $\Delta = \{\Delta_1, \Delta_2, \Delta_3\}$   $\Delta_1 = \{a_1\}$ ,  $\Delta_2 = \{a_{\kappa-n+1}\}$ , and  $\Delta_3 = \{a_k \in V(G) | a_k \notin \Delta_1, \Delta_2\}$  from the vertex set of  $CC_n^m$ . We assume the following cases on the vertex set of  $G = CC_n^m$  and on the copies of cycle graph, i.e.,  $m$ .

Table 18 shows different representations of  $a_k$  with respect to the partition  $\Delta_1$ .

Tables 19–21 show different representations of  $a_k$  with respect to the partition  $\Delta_2$ .

Third vector representations are

$$r(a_k | \Delta_3) = \begin{cases} 1, & k = 1, \kappa - n + 1, \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

Hence, it follows from the above discussion that  $pd(CC_n^m) \leq 3$  because all the vertices of  $CC_n^m$  have unique representations with respect to resolving partition set  $\Delta$ .

The reverse inequality  $pd(CC_n^m) \geq 3$  can be easily followed from the fact that partition dimension of the graph  $G$  is 2 if and only if  $G$  is a path graph. Thus, we conclude that, for  $n \geq 5$  and  $m \geq 2$ ,

$$pd(CC_n^m) = 3. \quad (9) \quad \square$$

### 5. Conclusion and Discussion

In this paper, we computed partition dimension of cycle-related graph such as paddle graph and cycle graph with chord and chain of cycles' graph. It has been shown that partition dimension of the aforementioned graph is 3. It was proved that the partition dimension of the cycle graph [1] is 3. We conclude that, by making small or significant changes in the cycle graph, do not affect its partition dimension.

#### Data Availability

No data were used to support the study.

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### Authors' Contributions

All the authors contributed equally for the preparation of this article.

### Acknowledgments

This paper was supported by Anhui Natural Science Research Project (2020) under Grant no. KJ2020A0696.

### References

- [1] G. Chartrand, E. Salehi, and P. Zhang, "The partition dimension of a graph," *Aequationes Mathematicae*, vol. 59, no. 1, pp. 45–54, 2000.
- [2] F. Harary and R. A. Melter, "On the metric dimension of a graph," *Ars Combinatoria*, vol. 2, pp. 191–195, 1976.
- [3] P. J. Slater, "Leaves of trees," *Congress Numerantium*, vol. 14, pp. 549–559, 1975.
- [4] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, New York, NY, USA, 1979.
- [5] N. Mehreen, R. Farooq, R. Farooq, and S. Akhter, "On partition dimension of fullerene graphs," *AIMS Mathematics*, vol. 3, no. 3, pp. 343–352, 2018.
- [6] M. C. Monica and S. Santhakumar, "Partition dimension of certain honeycomb derived networks," *International Journal of Pure and Applied Mathematics*, vol. 108, no. 4, pp. 809–818, 2016.
- [7] B. Rajan, A. William, I. Rajasingh, C. Grigorious, and S. Stephen, "On certain networks with partition dimension three," in *Proceedings of International Conference on Mathematics Engineering and Business Management*, Chengdu, China, March 2012.
- [8] J. A. Rodríguez-Velázquez, I. Lemańska, and M. Lemanska, "On the partition dimension of trees," *Discrete Applied Mathematics*, vol. 166, pp. 204–209, 2014.
- [9] J. A. Rodríguez-Velázquez, I. G. Yero, and H. Fernau, "On the partition dimension of unicyclic graphs," *Bulletin Mathematics Society of Science and Mathematics*, vol. 57, pp. 381–391, 2014.
- [10] R. Juan, I. G. Yero, and M. Lemanska, "On the partition dimension of trees," *Discrete Applied Mathematics*, vol. 166, pp. 204–209, 2014.
- [11] H. A. Amrullah and E. T. Baskoro, "The partition dimension for a subdivision of homogeneous caterpillars," *AKCE International Journal of Graphs and Combinatorics*, vol. 130, pp. 157–168, 1998.
- [12] D. Amrullah and E. T. Baskoro, "The partition dimension for a homogeneous firecrackers," *Far East Journal of Applied Mathematics*, vol. 90, no. 1, pp. 77–98, 2015.
- [13] H. M. A. Siddiqui and M. Imran, "Computation of metric dimension and partition dimension of Nanotubes," *Journal of Computational and Theoretical Nanoscience*, vol. 12, no. 2, pp. 199–203, 2015.
- [14] J. Cáceres, C. Hernando, M. Mora et al., "On the metric dimension of Cartesian products of graphs," *SIAM Journal on Discrete Mathematics*, vol. 21, no. 2, pp. 423–441, 2007.

- [15] V. Chvatal, *Mastermind, Combinatorica*, vol. 3, no. 3–4, pp. 325–329, 1983.
- [16] Z. Beerliova, F. Eberhard, T. Erlebach et al., “Network discovery and verification,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 12, pp. 2168–2181, 2006.
- [17] M. A. Johnson, “Structure-activity maps for visualizing the graph variables arising in drug design,” *Journal of Biopharmacy and Statistics*, vol. 3, pp. 203–236, 1993.
- [18] M. A. Johnson, “Browsable structure-activity datasets,” *Advances in Molecular Similarity*, pp. 153–170, JAI Press Connecticut, Stamford, CT, USA, 1998.
- [19] R. A. Melter and I. Tomescu, “Metric bases in digital geometry,” *Computer Vision, Graphics, and Image Processing*, vol. 25, no. 1, pp. 113–121, 1984.
- [20] G. Chartrand, L. Eroh, M. A. Johnson, and O. R. Oellermann, “Resolvability in graphs and the metric dimension of a graph,” *Discrete Applied Mathematics*, vol. 105, no. 1–3, pp. 99–113, 2000.
- [21] A. Ahmad, M. Baća, and S. Sultan, “Computing the metric dimension of kayak paddle graph and cycles with chord,” *Proyecciones Journal of Mathematics*, vol. 39, no. 2, pp. 287–300, 2020.
- [22] I. Javaid and S. Shokat, “On the partition dimension of some wheel related graphs,” *Journal of Prime Research in Mathematics*, vol. 4, pp. 154–164, 2008.
- [23] Y.-M. Chu, M. F. Nadeem, M. Azeem, and M. K. Siddiqui, “On sharp bounds on partition dimension of convex polytopes,” *IEEE Access*, vol. 8, Article ID 224781, 2020.
- [24] S. Pirzada, M. Aijaz, and S. P. Redmond, “Upper dimension and basis of zero divisor graphs of commutative rings,” *AKCE International Journal of Graphs and Combinatorics*, vol. 17, pp. 168–173, 2019.
- [25] S. Pirzada, M. Aijaz, and S. P. Redmond, “On upper dimension of graphs and their bases sets,” *Discrete Mathematics Letter*, vol. 3, pp. 37–43, 2020.
- [26] S. Pirzada and M. Aijaz, “On graphs with same metric and upper dimension,” *Discrete Mathematics, Algorithms and Applications*, vol. 13, no. 02, Article ID 2150015, 2021.



## Research Article

# Computing Analysis for First Zagreb Connection Index and Coindex of Resultant Graphs

Muhammad Javaid <sup>1</sup>, Usman Ali,<sup>1</sup> and Jia-Bao Liu <sup>2</sup>

<sup>1</sup>Department of Mathematics, School of Science, University of Management and Technology, Lahore 54770, Pakistan

<sup>2</sup>School of Mathematics and Physics, Anhui Jianzhu University, Hefei 230601, China

Correspondence should be addressed to Muhammad Javaid; [javidmath@gmail.com](mailto:javidmath@gmail.com)

Received 20 August 2020; Revised 2 January 2021; Accepted 13 January 2021; Published 1 March 2021

Academic Editor: Yong Zhang

Copyright © 2021 Muhammad Javaid et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A numeric parameter which studies the behaviour, structural, toxicological, experimental, and physicochemical properties of chemical compounds under several graphs' isomorphism is known as topological index. In 2018, Ali and Trinajstić studied the first Zagreb connection index ( $ZC_1$ ) to evaluate the value of a molecule. This concept was first studied by Gutman and Trinajstić in 1972 to find the solution of  $\pi$ -electron energy of alternant hydrocarbons. In this paper, the first Zagreb connection index and coindex are obtained in the form of exact formulae and upper bounds for the resultant graphs in terms of different indices of their factor graphs, where the resultant graphs are obtained by the product-related operations on graphs such as tensor product, strong product, symmetric difference, and disjunction. At the end, an analysis of the obtained results for the first Zagreb connection index and coindex on the aforesaid resultant graphs is interpreted with the help of numerical values and graphical depictions.

## 1. Introduction

Topological indices (TIs) are used in the study of cheminformatics which has key role in the formational structure of quantitative structures' activity and property relationships to examine the chemical reactivity and experimental activity of a chemical compound in a molecular graph [1]. So, TIs predict both physical and chemical features that are defined in the molecular graphs such as surface tension, solubility, connectivity, freezing point, boiling point, melting point, critical temperature, polarizability, heat of evaporation, and formation [2]. In addition, the medical behaviours and a number of drug particles of different compounds have formed with the help of various TIs in the pharmaceutical networks (see [3]). In particular, the TIs called by connection number-based Zagreb indices are used to compute the correlation values among various octane isomers such as heat of evaporation, acentric factor, molecular weight, connectivity, critical temperature, stability, and density (see [4, 5]).

Product graphs play an essential part to develop new molecular graphs from simple graphs. For this purpose,

Ashrafi et al. [6] defined the concept of coindices for several products on graphs. Das et al. [7] computed upper bounds for multiplicative Zagreb indices of operations such as join, corona, Cartesian, disjunction, and composition. The reformulated, multiplicative, hyper, first, second, and third Zagreb coindices with certain properties are defined in [8–13]. Relations between Zagreb coindices and some distance-based indices are computed in [14]. For more details, see [15–18]. For this purpose, we can define some operations such as tensor product, strong product, symmetric difference, and disjunction in the following definitions.

*Definition 1.* Tensor (or Kronecker or conjunctive or direct product of two graphs  $G_1$  and  $G_2$  is a graph  $G_1 \otimes G_2$  with vertex set  $V(G_1 \otimes G_2) = V(G_1) \otimes V(G_2)$  and edge set  $E(G_1 \otimes G_2) = \{(a_1, b_1)(a_2, b_2), \text{ where } (a_1, b_1), (a_2, b_2) \in V(G_1) \otimes V(G_2)\}$  and

$$[a_1 a_2 \in E(G_1)] \wedge [b_1 b_2 \in E(G_2)]. \quad (1)$$

The connection number of a vertex  $(a, b)$  of  $G_1 \otimes G_2$  is defined by  $\tau_{Q_1 \otimes Q_2}(a, b) = \tau_{G_1}(a) + \tau_{G_2}(b) + \tau_{G_1}(a)\tau_{G_2}(b)$ . For more details, see Figure 1.

**Definition 2.** Strong product or normal product  $(G_1 \otimes G_2)$  of two graphs  $G_1$  and  $G_2$  is obtained by taking the vertex set and edge set as  $V(Q_1 \otimes Q_2) = V(Q_1) \otimes V(Q_2)$  and  $E(Q_1 \otimes Q_2) = \{(a_1, b_1)(a_2, b_2), \text{ where } (a_1, b_1), (a_2, b_2) \in V(Q_1) \otimes V(Q_2)\}$  with condition

$$\begin{aligned} & \text{either } [a_1 = a_2 \in V(G_1) \wedge b_1 b_2 \in E(Q_2)] \\ & \text{or } [b_1 = b_2 \in V(Q_2) \wedge a_1 a_2 \in E(Q_1)] \\ & \text{or } [a_1 a_2 \in E(Q_1)] \wedge [b_1 b_2 \in E(Q_2)]. \end{aligned} \quad (2)$$

The connection number of a vertex  $(a, b)$  of  $Q_1 \otimes Q_2$  is defined by

$$\begin{aligned} \tau_{G_1 \otimes G_2}(a, b) &= \tau_{G_1}(a) + \tau_{G_2}(b) + d_{G_1}(a)\tau_{G_1}(a) \\ & \quad + d_{G_2}(b)\tau_{G_2}(b) + \tau_{G_1}(a)\tau_{G_2}(b) \\ &= \tau_{G_1}(a)[1 + d_{G_1}(a)] \\ & \quad + \tau_{G_2}(b)[1 + d_{G_2}(b) + \tau_{G_1}(a)], \end{aligned} \quad (3)$$

if  $n_1, n_2 \geq 4$ . For more details, see Figure 2.

**Definition 3.** Symmetric difference of two graphs  $Q_1$  and  $Q_2$  is a graph  $G_1 \ominus G_2$  with vertex set  $V(Q_1 \ominus Q_2) = V(Q_1) \ominus V(Q_2)$  and edge set  $E(Q_1 \ominus Q_2) = \{(a_1, b_1)(a_2, b_2), \text{ where } (a_1, b_1), (a_2, b_2) \in V(Q_1) \ominus V(Q_2)\}$  and

$$\begin{aligned} & [a_1 a_2 \in E(Q_1)] \\ & \text{or } [b_1 b_2 \in E(Q_2)], \end{aligned} \quad (4)$$

but not both hold at the same time, respectively.

The connection number of a vertex  $(a, b)$  of  $G_1 \ominus G_2$  is defined by

$$\begin{aligned} \tau_{G_1 \ominus G_2}(a, b) &= n_1 n_2 - 1 - n_2 d_{G_1}(a) - n_1 d_{G_2}(b) \\ & \quad + 2d_{G_1}(a)d_{G_2}(b) \\ &= n_2 [n_1 - d_{G_1}(a)] + d_{G_2}(b)[2d_{G_1}(a) - n_1] - 1. \end{aligned} \quad (5)$$

For more details, see Figure 3.

**Definition 4.** Disjunction of two graphs  $G_1$  and  $G_2$  is a graph  $G_1 \oplus G_2$  with vertex set  $V(Q_1 \oplus Q_2) = V(Q_1) \oplus V(Q_2)$  and edge set  $E(Q_1 \oplus Q_2) = \{(a_1, b_1)(a_2, b_2), \text{ where } (a_1, b_1), (a_2, b_2) \in V(Q_1) \oplus V(Q_2)\}$  and

$$\begin{aligned} & [a_1 a_2 \in E(Q_1)] \\ & \text{or } [b_1 b_2 \in E(Q_2)]. \end{aligned} \quad (6)$$

The connection number of a vertex  $(a, b)$  of  $G_1 \oplus G_2$  is defined by

$$\begin{aligned} \tau_{G_1 \oplus G_2}(a, b) &= n_1 n_2 - 1 - n_2 d_{G_1}(a) - n_1 d_{G_2}(b) \\ & \quad + d_{G_1}(a)d_{G_2}(b) \\ &= n_2 [n_1 - d_{G_1}(a)] + d_{G_2}(b)[d_{G_1}(a) - n_1] - 1. \end{aligned} \quad (7)$$

For more details, see Figure 4.

The graph theory provides the significant tools in the field of modern chemistry that is exploited to develop several types of molecular graphs and also predicts their chemical properties. In 1972, Gutman and Trinajstić [19] defined the first degree-based (number of vertices at distance one) TI called the first Zagreb index to compute the total  $\pi$ -electron energy of the molecules in molecular graphs. There are several TIs in literature, but degree-based TIs are studied more than others (see [20–29]).

Ali and Trinajstić [30] restudied the concept of connection number-based (number of vertices at distance two) TIs that were also defined by Gutman and Trinajstić in the same paper in 1972 to compute the total electron energy of the alternant hydrocarbons. They recalled them as Zagreb connection indices and reported that the chemical capability of the Zagreb connection indices is better than the classical Zagreb indices for the thirteen physicochemical properties of octane isomers. After two years, a few works have delivered on these connection number-based descriptors. Ali et al. [31] computed the analysis of Zagreb connection indices and coindices for some chemical structures of operations on graphs. For further studies and properties of the Zagreb connection indices, we refer to [32–40]. Recently, Gong et al. [41, 42] developed blocking lot-streaming flow shop scheduling problems and dynamic interval multi-objective optimization problems. These problems have been considered in various studies which have a close relation with the topic considered in this paper. For more details, see [43, 44].

In this paper, we compute the analysis for the first Zagreb connection index and coindex of the resultant graphs in the shape of exact formulae and upper bounds in terms of their factor graphs, where resultant graphs are obtained by operations such as tensor product, strong product, symmetric difference, and disjunction. Moreover, at the end, an analysis of the first Zagreb connection index and coindex on the aforesaid operations is included with the help of their numerical values and graphical depictions.

The rest of the paper is as follows: Section 2 represents the preliminary definitions and key results which are used in the main results, Section 3 contains the main results of product on graphs, and Section 4 includes the analysis and conclusions.

## 2. Preliminaries

Let  $G = (V(G), E(G))$  be a simple graph such that the order and size of  $G$  are  $|V(G)| = n$  and  $|E(G)| = e$ . The distance  $d(a, b)$  between any two vertices  $a$  and  $b$  of a graph  $G$  is equal to the length of a shortest path connecting them. For  $b \in V(G)$  and a positive integer  $p$ ,  $N_p(b/G) =$



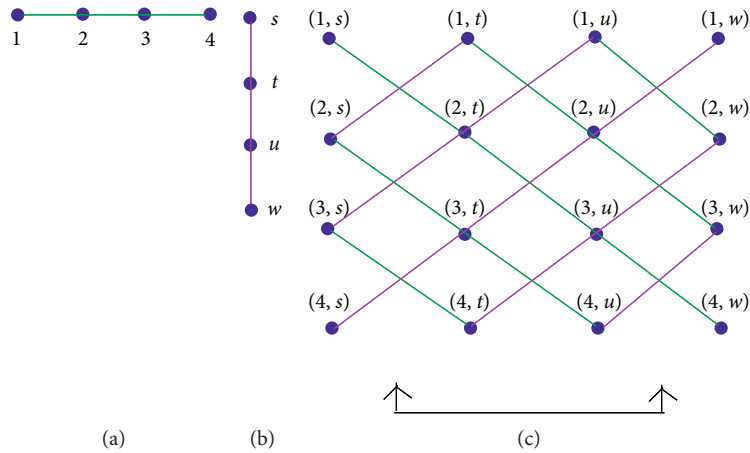


FIGURE 1: (a)  $G_1 \cong P_4$ , (b)  $G_2 \cong P_4$ , and (c) tensor product  $(P_4 \otimes P_4)$ .

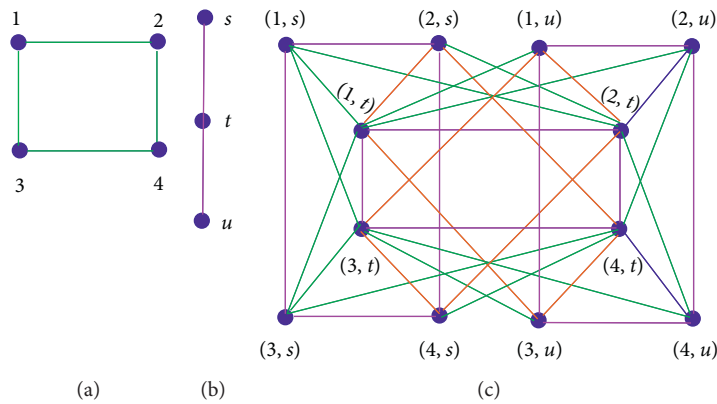


FIGURE 2: (a)  $G_1 \cong C_4$  (b)  $G_2 \cong P_3$ , and (c) strong product  $(C_4 \otimes P_3)$ .

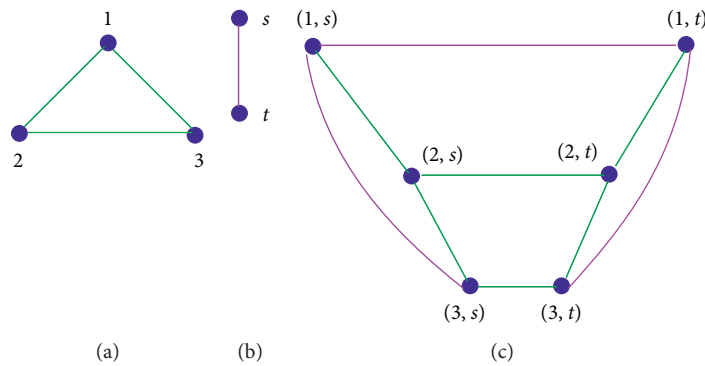


FIGURE 3: (a)  $G_1 \cong C_3$ , (b)  $G_2 \cong P_2$ , and (c) symmetric difference  $(C_3 \oplus P_2)$ .

$\{u \in V(G) : d(a, u) = p\}$  denotes the open  $p$ -neighborhood of  $a$  in a graph  $G$ , where  $d_p(a/G) = |N_p(a/G)|$  is called  $p$ -distance degree of a vertex  $a$  in any graph  $G$ . The degree of a vertex  $a$  in a graph  $G$  is the number of edges incident on it, and it is denoted by  $d_G(a)$ . In particular:

If  $p = 1$ ,  $d_1(a/G) = |N_1(a/G)| = d_G(a) = \text{degree of } a$  (number of vertices at distance one from  $a$ )

If  $p = 2$ ,  $d_2(a/G) = |N_2(a/G)| = \tau_G(a) = \text{connection number of } a$  (number of vertices at distance two from  $a$ )

The complement of a graph  $G$  is denoted by  $\bar{G}$ . It is also simple with the same vertex set as of  $G$ , but edge set is defined as  $E(\bar{G}) = \{ab : a, b \in V(G) \wedge ab \notin E(G)\}$ . Thus,  $E(G) \cup E(\bar{G}) = E(K_n)$ , where  $K_n$  is a complete graph of

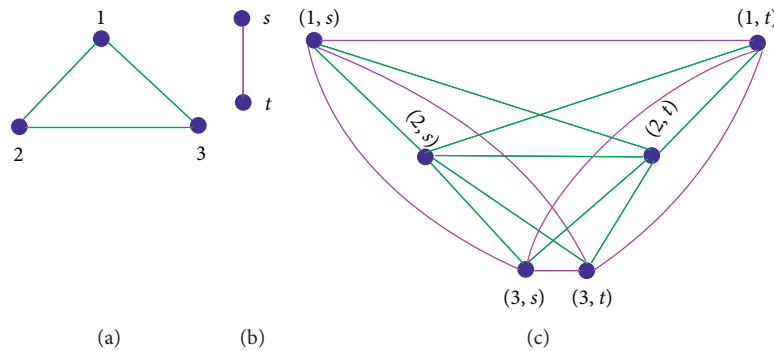


FIGURE 4: (a)  $G_1 \cong C_3$ , (b)  $G_2 \cong P_2$ , and (c) disjunction  $(C_3 \cdot P_2)$ .

order  $n$  and size  $|E(K_n)| = \binom{n}{2}$ . Moreover, if  $|E(G)| = e$ , then  $|E(\overline{G})| = \binom{n}{2} - e = \mu$  and  $d_{\overline{G}}(b) = n - 1 - d_G(b)$ , where  $d_G(b)$  and  $d_{\overline{G}}(b)$  are the degrees of the vertex  $b$  in  $G$  and  $\overline{G}$ , respectively. Now, throughout the paper, for two graphs  $G_1$  and  $G_2$ , we assume that  $|V(G_1)| = n_1$ ,  $|V(G_2)| = n_2$ ,  $|E(G_1)| = e_1$ , and  $|E(G_2)| = e_2$ . Finally, it is important to note that Zagreb connection coindices of  $G$  are not Zagreb connection indices of  $\overline{G}$  because the connection number works according to  $G$ . For any terminology or notion which are not mentioned here, we refer to [45, 46].

**Definition 5.** For a graph  $G$ , the first Zagreb index ( $M_1(G)$ ) and second Zagreb index ( $M_2(G)$ ) are defined as

$$\begin{aligned} M_1(G) &= \sum_{ab \in E(G)} [d_G(a) + d_G(b)], \\ M_2(G) &= \sum_{ab \in E(G)} [d_G(a) \times d_G(b)]. \end{aligned} \quad (8)$$

These degree-based indices are defined by Gutman, Trinajstić, and Ruscic (see [19, 47]) which are frequently used to predict better outcomes of the various parameters related to the molecular networks such as chirality, complexity, entropy, heat energy, ZE-isomerism, heat capacity, absolute value of correlation coefficient, chromatographic, retention times in chromatographic, pH, and molar ratio [19, 48]. Symmetrical to these degree-based TIs, the connection-based TIs are discussed in Definitions 6 and 7.

**Definition 6.** For a graph  $G$ , the first Zagreb connection index ( $ZC_1(G)$ ) is defined as

$$ZC_1(G) = \sum_{b \in V(G)} [\tau_G(b)]^2. \quad (9)$$

**Definition 7.** For a graph  $G$ , the modified first Zagreb connection index ( $ZC_1^*(G)$ ) and second Zagreb connection index ( $ZC_2(G)$ ) are defined as

$$\begin{aligned} ZC_1^*(G) &= \sum_{ab \in E(G)} [\tau_G(a) + \tau_G(b)], \\ ZC_2(G) &= \sum_{ab \in E(G)} [\tau_G(a) \times \tau_G(b)]. \end{aligned} \quad (10)$$

**Definition 8.** For a graph  $G$ , the first Zagreb coindex ( $\overline{M}_1(G)$ ) and second Zagreb coindex ( $\overline{M}_2(G)$ ) are defined as

$$\begin{aligned} \overline{M}_1(G) &= \sum_{ab \notin E(G)} [d_G(a) + d_G(b)], \\ \overline{M}_2(G) &= \sum_{ab \notin E(G)} [d_G(a) \times d_G(b)]. \end{aligned} \quad (11)$$

These coindices associated with the degree-based classical Zagreb indices are defined by Ashrafi et al. (see [6]). The coindices associated with the connection-based Zagreb indices are defined in Definition 9.

**Definition 9.** For a graph  $G$ , the first Zagreb connection coindex ( $\overline{ZC}_1(G)$ ) and second Zagreb connection coindex ( $\overline{ZC}_2(G)$ ) are defined as

$$\begin{aligned} \overline{ZC}_1(G) &= \sum_{ab \notin E(G)} [\tau_G(a) + \tau_G(b)], \\ \overline{ZC}_2(G) &= \sum_{ab \notin E(G)} [\tau_G(a) \times \tau_G(b)]. \end{aligned} \quad (12)$$

The modified coindices associated with the connection number-based Zagreb indices are defined in Definition 10. These modified coindices associated with the connection number-based Zagreb indices are defined by Ali et al. (see [49]).

**Definition 10.** For a graph  $G$ , the modified first Zagreb connection coindex ( $M\overline{ZC}_1(G)$ ) and the modified second Zagreb connection coindex ( $M\overline{ZC}_2(G)$ ) are defined as

$$\begin{aligned}
 M\bar{Z}C_1(K) &= \sum_{uv \notin E(K)} [d_K(u)\tau_K(v) + d_K(v)\tau_K(u)], \\
 M\bar{Z}C_2(K) &= \sum_{uv \notin E(K)} [d_K(u)\tau_K(u) + d_K(v)\tau_K(v)]. \tag{13}
 \end{aligned}$$

The degree/connection-based coindices defined in Definitions 8–10 study the various physicochemical and isomer properties of molecules on the basis of the adjacency and nonadjacency pairs of vertices in the molecular networks. For more details, see [6, 30, 31, 36].

Now, we present some important results which are used in the main results.

**Lemma 1** (see [50]). *Let  $G$  be a connected graph and  $\bar{G}$  be its complement with  $n$  vertices and  $e$  edges. Then,*

- (a)  $\sum_{b \in V(G)} d_G(b) = 2e$
- (b)  $M_1(\bar{G}) \geq M_1(G) - 4$

**Lemma 2** (see [36]). *Let  $G$  be a connected and  $\{C_3, C_4\}$ -free graph with  $n$  vertices and  $e$  edges. Then,*

- (a)  $\sum_{b \in V(G_1)} \tau_{G_1}(b) = M_1(G_1) - 2e_1 = A$
- (b)  $\sum_{b \in V(G_2)} \tau_{G_2}(b) = M_1(G_2) - 2e_2 = B$

### 3. Main Results

This section contains the main results for the first Zagreb connection index ( $ZC_1$ ) and first Zagreb connection coindex ( $\bar{Z}C_1$ ) of the product on graphs obtained under the operations of tensor product, strong product, symmetric difference, and disjunction.

**Theorem 1.** *Let  $G_1$  and  $G_2$  be two connected and  $\{C_3, C_4\}$ -free graphs. Then,  $ZC_1$  and  $\bar{Z}C_1$  of the tensor product  $(G_1 \otimes G_2)$  are*

$$ZC_1(G_1 \otimes G_2) = n_2 ZC_1(G_1) + n_1 ZC_1(G_2) + ZC_1(G_1)ZC_1(G_2) + 2AB + 2AZC_1(G_2) + 2BZC_1(G_1), \tag{14}$$

$$\bar{Z}C_1(G_1 \otimes G_2) = 2\mu_2 \bar{Z}C_1(G_1) + 2\mu_1 \bar{Z}C_1(G_2) + 2 \sum_{a_1, a_2 \notin E(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1}(a_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)]. \tag{15}$$

*Proof.* Since  $\tau_{G_1 \otimes G_2}(a, b) = \tau_{G_1}(a) + \tau_{G_2}(b) + \tau_{G_1}(a)\tau_{G_2}(b)$ , where  $a \in V(G_1)$ ,  $b \in V(G_2)$ , and  $(a, b) \in G_1 \otimes G_2$ ,

$$\begin{aligned}
 ZC_1(G_1 \otimes G_2) &= \sum_{(a,b) \in V(G_1 \otimes G_2)} [\tau_{G_1 \otimes G_2}(a, b)]^2 \\
 &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} [\tau_{G_1}(a) + \tau_{G_2}(b) + \tau_{G_1}(a)\tau_{G_2}(b)]^2 \\
 &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} [\tau_{G_1}^2(a) + \tau_{G_2}^2(b) + \tau_{G_1}^2(a)\tau_{G_2}^2(b) + 2\tau_{G_1}(a) + \tau_{G_2}(b) + 2\tau_{G_1}(a)\tau_{G_2}^2(b) + 2\tau_{G_1}^2(a)\tau_{G_2}(b)] \\
 &= \sum_{a \in V(G_1)} [n_2 \tau_{G_1}^2(a) + ZC_1(G_2) + \tau_{G_1}^2(a)ZC_1(G_2) + 2\tau_{G_1}(a)\{M_1(G_2) - 2e_2\} \\
 &\quad + 2\tau_{G_1}(a)ZC_1(G_2) + 2\tau_{G_1}^2(a)\{M_1(G_2) - 2e_2\}]. \tag{16}
 \end{aligned}$$

Using Lemma 2 (a) and (b),

$$ZC_1(G_1 \otimes G_2) = n_2 ZC_1(G_1) + n_1 ZC_1(G_2) + ZC_1(G_1)ZC_1(G_2) + 2AB + 2AZC_1(G_2) + 2BZC_1(G_1), \tag{17}$$

$$\begin{aligned}
 \bar{Z}C_1(G_1 \otimes G_2) &= \sum_{(a_1, b_1), (a_2, b_2) \notin E(G_1 \otimes G_2)} [\tau_{G_1 \otimes G_2}(a_1, b_1) + \tau_{G_1 \otimes G_2}(a_2, b_2)] \\
 &= 2 \sum_{a_1, a_2 \notin E(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\{\tau_{G_1}(a_1) + \tau_{G_2}(b_1) + \tau_{G_1}(a_1)\tau_{G_2}(b_1)\} + \{\tau_{G_1}(a_2) + \tau_{G_2}(b_2) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)\}] \\
 &= 2 \sum_{a_1, a_2 \notin E(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\{\tau_{G_1}(a_1) + \tau_{G_1}(a_2)\} + \{\tau_{G_2}(b_1) + \tau_{G_2}(b_2)\} + \{\tau_{G_1}(a_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)\}] \\
 \bar{Z}C_1(G_1 \otimes G_2) &= 2\mu_2 \bar{Z}C_1(G_1) + 2\mu_1 \bar{Z}C_1(G_2) + 2 \sum_{a_1, a_2 \notin E(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1}(a_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)]. \tag{18}
 \end{aligned}$$

□

**Theorem 2.** Let  $G_1$  and  $G_2$  be two connected and  $\{C_3, C_4\}$ -free graphs. Then,  $ZC_1$  and  $\overline{ZC}_1$  of the strong product  $(G_1 \otimes G_2)$  are

$$\begin{aligned} ZC_1(G_1 \otimes G_2) &= 2BZC_1^*(G_1) + 2AZC_1^*(G_2) + 2ZC_1^*(G_1)ZC_1^*(G_2) + (n_2 + 2B)ZC_1(G_1) \\ &\quad + (n_1 + 2A)ZC_1(G_2) + ZC_1(G_1)ZC_1(G_2) + 2AB + 2(n_2 + B) \sum_{a \in V(G_1)} d_{G_1}(a)\tau_{G_1}^2(a) + 2(n_1 + A) \\ &\quad \cdot \sum_{b \in V(G_2)} d_{G_2}(b)\tau_{G_2}^2(b) + n_2 \sum_{a \in V(G_1)} d_{G_1}^2(a)\tau_{G_1}^2(a) + n_1 \sum_{b \in V(G_2)} d_{G_2}^2(b)\tau_{G_2}^2(b), \end{aligned} \quad (19)$$

$$\begin{aligned} \overline{ZC}_1(G_1 \otimes G_2) &\leq (n_2 + 2\mu_2)M\overline{ZC}_2(G_1) + (n_1 + 2\mu_1)M\overline{ZC}_2(G_2) + (n_2 + 2\mu_2 + B)\overline{ZC}_1(G_1) \\ &\quad + (n_1 + 2\mu_1 + A)\overline{ZC}_1(G_2) + 2[\mu_2 ZC_1^*(G_1) + \mu_1 ZC_1^*(G_2)] + 2[\mu_2 A + \mu_1 B] \\ &\quad + 2 \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1}(a_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)]. \end{aligned} \quad (20)$$

*Proof.* Since  $\tau_{G_1 \otimes G_2}(a, b) = \tau_{G_1}(a)[1 + d_{G_1}(a)] + \tau_{G_2}(b)[1 + d_{G_2}(b) + \tau_{G_1}(a)]$ , where  $a \in V(G_1)$ ,  $b \in V(G_2)$ , and  $(a, b) \in G_1 \otimes G_2$ ,

$$\begin{aligned} ZC_1(G_1 \otimes G_2) &= \sum_{(a,b) \in V(G_1 \otimes G_2)} [\tau_{G_1 \otimes G_2}(a, b)]^2 \\ &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} [\tau_{G_1}(a)\{1 + d_{G_1}(a)\} + \tau_{G_2}(b)\{1 + d_{G_2}(b) + \tau_{G_1}(a)\}]^2 \\ &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} \left[ \tau_{G_1}^2(a)\{1 + d_{G_1}(a)\}^2 + \tau_{G_2}^2(b)\{1 + d_{G_2}(b) + \tau_{G_1}(a)\}^2 + \right. \\ &\quad \left. 2\tau_{G_1}(a)\tau_{G_2}(b)\{1 + d_{G_1}(a)\}\{1 + d_{G_2}(b) + \tau_{G_1}(a)\} \right] \\ &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} \left[ \tau_{G_1}^2(a)\{1 + d_{G_1}^2(a) + 2d_{G_1}(a)\} + \tau_{G_2}^2(b)\{1 + d_{G_2}(b)\}^2 + \tau_{G_1}^2(a) + 2\{1 + d_{G_2}(b)\}\tau_{G_1}(a) \right] \\ &\quad + 2\tau_{G_1}(a)\tau_{G_2}(b)\{1 + d_{G_2}(b) + \tau_{G_1}(a) + d_{G_1}(a) + d_{G_1}(a)d_{G_2}(b) + d_{G_1}(a)\tau_{G_1}(a)\} \\ &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} \left[ \tau_{G_1}^2(a) + d_{G_1}^2(a)\tau_{G_1}^2(a) + 2d_{G_1}(a)\tau_{G_1}^2(a) + \tau_{G_2}^2(b) + d_{G_2}^2(b)\tau_{G_2}^2(b) + d_{G_2}(b)\tau_{G_2}^2(b) \right. \\ &\quad + \tau_{G_1}^2(a)\tau_{G_2}^2(b) + 2\tau_{G_1}(a)\tau_{G_2}^2(b) + 2d_{G_2}(b)\tau_{G_1}(a)\tau_{G_2}^2(b) + 2\tau_{G_1}(a)\tau_{G_2}(b) + 2d_{G_2}(b)\tau_{G_1}(a)\tau_{G_2}(b) \\ &\quad \left. + 2\tau_{G_1}^2(a)\tau_{G_2}(b) + 2d_{G_1}(a)\tau_{G_1}(a)\tau_{G_2}(b) + 2d_{G_1}(a)\tau_{G_1}(a)d_{G_2}(b)\tau_{G_2}(b) + 2d_{G_1}(a)\tau_{G_1}^2(a)\tau_{G_2}(b) \right]. \end{aligned} \quad (21)$$

Using Lemma 2 (a) and (b),

$$\begin{aligned}
 ZC_1(G_1 \otimes G_2) &= 2BZC_1^*(G_1) + 2AZC_1^*(G_2) \\
 &\quad + 2ZC_1^*(G_1)ZC_1^*(G_2) + (n_2 + 2B)ZC_1(G_1) \\
 &\quad + (n_1 + 2A)ZC_1(G_2) + ZC_1(G_1)ZC_1(G_2) \\
 &\quad + 2AB + 2(n_2 + B) \sum_{a \in V(G_1)} d_{G_1}(a)\tau_{G_1}^2(a) \\
 &\quad + 2(n_1 + A) \sum_{b \in V(G_2)} d_{G_2}(b)\tau_{G_2}^2(b) \\
 &\quad + n_2 \sum_{a \in V(G_1)} d_{G_1}^2(a)\tau_{G_1}^2(a) + n_1 \sum_{b \in V(G_2)} d_{G_2}^2(b)\tau_{G_2}^2(b),
 \end{aligned} \tag{22}$$

$$\begin{aligned}
 \bar{Z}C_1(G_1 \otimes G_2) &= \sum_{(a_1, b_1)(a_2, b_2) \notin E(G_1 \otimes G_2)} [\tau_{G_1 \otimes G_2}(a_1, b_1) + \tau_{G_1 \otimes G_2}(a_2, b_2)] \\
 &= \sum_{a \in V(G_1)} \sum_{b_1 b_2 \notin E(G_2)} [\tau_{G_1 \otimes G_2}(a, b_1) + \tau_{G_1 \otimes G_2}(a, b_2)] \\
 &\quad + \sum_{b \in V(G_2)} \sum_{a_1 a_2 \notin E(G_1)} [\tau_{G_1 \otimes G_2}(a_1, b) + \tau_{G_1 \otimes G_2}(a_2, b)] \\
 &\quad + \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1 \otimes G_2}(a_1, b_1) + \tau_{G_1 \otimes G_2}(a_2, b_2)].
 \end{aligned} \tag{23}$$

We take

$$\begin{aligned}
 &\sum_{a \in V(G_1)} \sum_{b_1 b_2 \notin E(G_2)} [\tau_{G_1 \otimes G_2}(a, b_1) + \tau_{G_1 \otimes G_2}(a, b_2)] \\
 &= \sum_{a \in V(G_1)} \sum_{b_1 b_2 \notin E(G_2)} \left[ \{\tau_{G_1}(a) + \tau_{G_2}(b_1) + d_{G_1}(a)\tau_{G_1}(a) + d_{G_2}(b_1)\tau_{G_2}(b_1) + \tau_{G_1}(a)\tau_{G_2}(b_1)\} \right. \\
 &\quad \left. + \{\tau_{G_1}(a) + d_{\tau_{G_2}(b_2)G_1}(a)\tau_{G_1}(a) + d_{G_2}(b_2)\tau_{G_2}(b_2) + \tau_{G_1}(a)\tau_{G_2}(b_2)\} \right] \\
 &= \sum_{a \in V(G_1)} \sum_{b_1 b_2 \notin E(G_2)} [2\tau_{G_1}(a) + \{\tau_{G_2}(b_1) + \tau_{G_2}(b_2)\} + 2d_{G_1}(a)\tau_{G_1}(a) + \{d_{G_2}(b_1)\tau_{G_2}(b_1) \\
 &\quad + d_{G_2}(b_2)\tau_{G_2}(b_2)\} + \tau_{G_1}(a)\{\tau_{G_2}(b_1) + \tau_{G_2}(b_2)\}].
 \end{aligned} \tag{24}$$

We know that

$$\begin{aligned}
 \sum_{b_1 b_2 \notin E(G_2)} &= \binom{n_2}{2} - e_2 = \mu_2 \\
 &= 2\mu_2 A + n_1 \bar{Z}C_1(G_2) + 2\mu_2 ZC_1^*(G_1) + n_1 M \bar{Z}C_2(G_2) + A \bar{Z}C_1(G_2).
 \end{aligned} \tag{25}$$

Also, we take

$$\begin{aligned}
 &\sum_{b \in V(G_2)} \sum_{a_1 a_2 \notin E(G_1)} [\tau_{G_1 \otimes G_2}(a_1, b) + \tau_{G_1 \otimes G_2}(a_2, b)] \\
 &\sum_{b \in V(G_2)} \sum_{a_1 a_2 \notin E(G_1)} \left[ \{\tau_{G_1}(a_1) + \tau_{G_2}(b) + d_{G_1}(a_1)\tau_{G_1}(a_1) + d_{G_2}(b)\tau_{G_2}(b) + \tau_{G_1}(a_1)\tau_{G_2}(b)\} \right. \\
 &\quad \left. + \{\tau_{G_1}(a_2) + \tau_{G_2}(b) + d_{G_1}(a_2)\tau_{G_1}(a_2) + d_{G_2}(b)\tau_{G_2}(b) + \tau_{G_1}(a_2)\tau_{G_2}(b)\} \right] \\
 &= n_2 \bar{Z}C_1(G_1) + 2\mu_1 B + n_2 M \bar{Z}C_2(G_1) + 2\mu_1 ZC_1^*(G_2) + B \bar{Z}C_1(G_1).
 \end{aligned} \tag{26}$$

Again, we take

$$\begin{aligned}
 & \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1 \otimes G_2}(a_1, b_1) + \tau_{G_1 \otimes G_2}(a_2, b_2)] \\
 & \leq 2 \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\{\tau_{G_1}(a_1) + \tau_{G_2}(b_1) + d_{G_1}(a_1)\tau_{G_1}(a_1) + d_{G_2}(b_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_1)\tau_{G_2}(b_1) \\
 & + \{\tau_{G_1}(a_2) + \tau_{G_2}(b_2) + d_{G_1}(a_2)\tau_{G_1}(a_2) + d_{G_2}(b_2)\tau_{G_2}(b_2) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)\}]. \tag{27}
 \end{aligned}$$

Similarly =  $2\mu_2 \overline{ZC}_1(G_1) + 2\mu_1 \overline{ZC}_1(G_2) + 2\mu_2 M \overline{ZC}_2(G_1) + 2\mu_1 M \overline{ZC}_2(G_2)$

$$+ 2 \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1}(a_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)].$$

Consequently,

$$\begin{aligned}
 \overline{ZC}_1(G_1 \otimes G_2) & \leq (n_2 + 2\mu_2)M \overline{ZC}_2(G_1) + (n_1 + 2\mu_1)M \overline{ZC}_2(G_2) + (n_2 + 2\mu_2 + B)\overline{ZC}_1(G_1) \\
 & + (n_1 + 2\mu_1 + A)\overline{ZC}_1(G_2) + 2[\mu_2 ZC_1^*(G_1) + \mu_1 ZC_1^*(G_2)] + 2[\mu_2 A + \mu_1 B] \\
 & + 2 \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1}(a_1)\tau_{G_2}(b_1) + \tau_{G_1}(a_2)\tau_{G_2}(b_2)]. \tag{28}
 \end{aligned}$$

□

**Theorem 3.** Let  $G_1$  and  $G_2$  be two connected and  $\{C_3, C_4\}$ -free graphs. Then,  $ZC_1$  and  $\overline{ZC}_1$  of the symmetric difference  $(G_1 \ominus G_2)$  are

$$\begin{aligned}
 ZC_1(G_1 \ominus G_2) & = n_2(n_2^2 - 8e_2)M_1(G_1) + n_1(n_1^2 - 8e_1)M_1(G_2) + 4M_1(G_1)M_1(G_2) \\
 & + n_1 n_2 (n_1^2 n_2^2 - 4n_2^2 e_1 - 4n_1^2 e_2 - 2n_1 n_2 + 24e_1 e_2 + 1) + 4(n_1^2 e_2 + n_2^2 e_1 - 4e_1 e_2), \tag{29}
 \end{aligned}$$

$$\begin{aligned}
 \overline{ZC}_1(G_1 \ominus G_2) & = (4e_2 - n_2^2)\overline{M}_1(G_1) + (4e_1 - n_1^2)\overline{M}_1(G_2) + 2\mu_1(n_1 n_2^2 - n_2 - 2n_1 e_2) \\
 & + 2\mu_2(n_1^2 n_2 - n_1 - 2n_2 e_1). \tag{30}
 \end{aligned}$$

*Proof.* Since  $\tau_{G_1 \ominus G_2}(a, b) = n_2[n_1 - d_{G_1}(a)] + d_{G_2}(b) [2d_{G_1}(a) - n_1] - 1$ , where  $a \in V(G_1)$ ,  $b \in V(G_2)$ , and  $(a, b) \in G_1 \ominus G_2$ ,

$$\begin{aligned}
 ZC_1(G_1 \ominus G_2) &= \sum_{(a,b) \in V(G_1 \ominus G_2)} [\tau_{G_1 \ominus G_2}(a, b)]^2 \\
 &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} [n_2 \{n_1 - d_{G_1}(a)\} + d_{G_2}(b) \{2d_{G_1}(a) - n_1\} - 1]^2 \\
 &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} [n_2^2 \{n_1 - d_{G_1}(a)\}^2 + d_{G_2}^2(b) \{2d_{G_1}(a) - n_1\}^2 + 1 + 2n_2 d_{G_2}(b) \{n_1 - d_{G_1}(a)\} \\
 &\quad \{2d_{G_1}(a) - n_1\} - 2d_{G_2}(b) \{2d_{G_1}(a) - n_1\} - 2n_2 \{n_1 - d_{G_1}(a)\}] \\
 &= n_1^3 n_2^3 + n_2^3 M_1(G_1) - 4n_1 n_2^3 e_1 + 4M_1(G_1) M_1(G_2) + n_1^3 M_1(G_2) - 8n_1 e_1 M_1(G_2) + n_1 n_2 \\
 &\quad + 16n_1 n_2 e_1 e_2 - 4n_1^3 n_2 e_2 - 8n_2 e_2 M_1(G_1) + 8n_1 n_2 e_1 e_2 - 16e_1 e_2 + 4n_1^2 e_2 - 2n_1^2 n_2^2 + 4n_2^2 e_1, \\
 ZC_1(G_1 \ominus G_2) &= n_2(n_2^2 - 8e_2)M_1(G_1) + n_1(n_1^2 - 8e_1)M_1(G_2) + 4M_1(G_1)M_1(G_2) + n_1 n_2 \\
 &\quad (n_1^2 n_2^2 - 4n_2^2 e_1 - 4n_1^2 e_2 - 2n_1 n_2 + 24e_1 e_2 + 1) + 4(n_1^2 e_2 + n_2^2 e_1 - 4e_1 e_2).
 \end{aligned}
 \tag{31}$$

$$\begin{aligned}
 \bar{Z}C_1(G_1 \ominus G_2) &= \sum_{(a_1, b_1), (a_2, b_2) \notin E(G_1 \ominus G_2)} [\tau_{G_1 \ominus G_2}(a_1, b_1) + \tau_{G_1 \ominus G_2}(a_2, b_2)] \\
 &= \sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1 \ominus G_2}(a, b_1) + \tau_{G_1 \ominus G_2}(a, b_2)] + \sum_{b \in V(G_2)} \sum_{a_1, a_2 \notin E(G_1)} [\tau_{G_1 \ominus G_2}(a_1, b) + \tau_{G_1 \ominus G_2}(a_2, b)] \\
 &\quad \cdot \sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1 \ominus G_2}(a, b_1) + \tau_{G_1 \ominus G_2}(a, b_2)] \sum_{b \in V(G_2)} \sum_{a_1, a_2 \notin E(G_1)} [\tau_{G_1 \ominus G_2}(a_1, b) + \tau_{G_1 \ominus G_2}(a_2, b)] \\
 &\quad \cdot [\tau_{G_1 \ominus G_2}(a_1, b_1) + \tau_{G_1 \ominus G_2}(a_2, b_2)].
 \end{aligned}
 \tag{32}$$

We take

$$\begin{aligned}
 &\sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1 \ominus G_2}(a, b_1) + \tau_{G_1 \ominus G_2}(a, b_2)] \\
 &= \sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\{n_1 n_2 - 1 - n_2 d_{G_1}(a) - n_1 d_{G_2}(b_1) + 2d_{G_1}(a) d_{G_2}(b_1)\} \\
 &\quad + \{n_1 n_2 - 1 - n_2 d_{G_1}(a) - n_1 d_{G_2}(b_2) + 2d_{G_1}(a) d_{G_2}(b_2)\}].
 \end{aligned}
 \tag{33}$$

$$\text{Similarly} = 2n_1^2 n_2 \mu_2 - 2n_1 \mu_2 - 4n_2 e_1 \mu_2 - n_1^2 \bar{M}_1(G_2) + 4e_1 \bar{M}_1(G_2).$$

Also, we take

$$\sum_{b \in V(G_2)} \sum_{a_1, a_2 \notin E(G_1)} [\tau_{G_1 \ominus G_2}(a_1, b) + \tau_{G_1 \ominus G_2}(a_2, b)].
 \tag{34}$$

$$\text{Similarly} = 2n_1 n_2^2 \mu_1 - 2n_2 \mu_1 - n_2^2 \bar{M}_1(G_1) - 4n_1 e_2 \mu_1 + 4e_2 \bar{M}_1(G_1).$$

Again, we take (null case)

$$N = \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \uparrow b_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] = 0. \quad (35)$$

We further take (also null case)

$$N = \sum_{\substack{b_1 b_2 \notin E(G_1) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} \sum_{\substack{a_1 a_2 \notin E(G_2) \\ a_1 \neq a_2 \wedge a_1 \uparrow a_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] = 0. \quad (36)$$

Consequently,

$$\begin{aligned} \bar{Z}C_1(G_1 \oplus G_2) &= (4e_2 - n_2^2)\bar{M}_1(G_1) + (4e_1 - n_1^2)\bar{M}_1(G_2) \\ &\quad + 2\mu_1(n_1 n_2^2 - n_2 - 2n_1 e_2) \\ &\quad + 2\mu_2(n_1^2 n_2 - n_1 - 2n_2 e_1). \end{aligned} \quad (37)$$

□

**Theorem 4.** Let  $G_1$  and  $G_2$  be two connected and  $\{C_3, C_4\}$ -free graphs. Then,  $ZC_1$  and  $\bar{Z}C_1$  of the disjunction  $(G_1 \oplus G_2)$  are

$$\begin{aligned} ZC_1(G_1 \oplus G_2) &= n_2(n_2^2 - 4e_2)M_1(G_1) + n_1(n_1^2 - 4e_1)M_1(G_2) + M_1(G_1)M_1(G_2) + n_1 n_2 \\ &\quad (n_1^2 n_2^2 - 4n_2^2 e_1 - 4n_1^2 e_2 - 2n_1 n_2 + 16e_1 e_2 + 1) + 4(n_1^2 e_2 + n_2^2 e_1 - 2e_1 e_2), \end{aligned} \quad (38)$$

$$\begin{aligned} \bar{Z}C_1(G_1 \oplus G_2) &= (2e_2 - n_2^2 - 2n_2 \mu_2)\bar{M}_1(G_1) + (2e_1 - n_1^2 - 2n_1 \mu_1)\bar{M}_1(G_2) \\ &\quad + 2\mu_1(n_1 n_2^2 - n_2 - 2n_1 e_2) + 2\mu_2(n_1^2 n_2 - n_1 - 2n_2 e_1) + 4\mu_1 \mu_2 (n_1 n_2 - 1) \\ &\quad + 2 \sum_{\substack{a_1 a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1 b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [d_{G_1}(a_1)d_{G_2}(b_1) + d_{G_1}(a_2)d_{G_2}(b_2)]. \end{aligned} \quad (39)$$

*Proof.* Since  $\tau_{G_1 \oplus G_2}(a, b) = n_2[n_1 - d_{G_1}(a)] + d_{G_2}(b)[d_{G_1}(a) - n_1] - 1$ , where  $a \in V(G_1)$ ,  $b \in V(G_2)$ , and  $(a, b) \in G_1 \oplus G_2$ ,

$$\begin{aligned} ZC_1(G_1 \oplus G_2) &= \sum_{(a,b) \in V(G_1 \oplus G_2)} [\tau_{G_1 \oplus G_2}(a, b)]^2 \\ &= \sum_{a \in V(G_1)} \sum_{b \in V(G_2)} [n_2\{n_1 - d_{G_1}(a)\} + d_{G_2}(b)\{d_{G_1}(a) - n_1\} - 1]^2. \end{aligned} \quad (40)$$



Similarly,

$$ZC_1(G_1 \oplus G_2) = n_2(n_2^2 - 4e_2)M_1(G_1) + n_1(n_1^2 - 4e_1)M_1(G_2) + M_1(G_1)M_1(G_2) + n_1n_2 \cdot (n_1^2n_2^2 - 4n_2^2e_1 - 4n_1^2e_2 - 2n_1n_2 + 16e_1e_2 + 1) + 4(n_1^2e_2 + n_2^2e_1 - 2e_1e_2). \quad (41)$$

$$\begin{aligned} \bar{Z}C_1(G_1 \oplus G_2) &= \sum_{(a_1, b_1)(a_2, b_2) \notin E(G_1 \oplus G_2)} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] \\ &= \sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1 \oplus G_2}(a, b_1) + \tau_{G_1 \oplus G_2}(a, b_2)] + \sum_{b \in V(G_2)} \sum_{a_1, a_2 \notin E(G_1)} [\tau_{G_1 \oplus G_2}(a_1, b) + \tau_{G_1 \oplus G_2}(a_2, b)] \\ &\quad \sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] \\ &\quad \sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \uparrow b_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] \\ &+ \sum_{\substack{b_1, b_2 \notin E(G_1) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} \sum_{\substack{a_1, a_2 \notin E(G_2) \\ a_1 \neq a_2 \wedge a_1 \uparrow a_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)]. \end{aligned} \quad (42)$$

We take

$$\begin{aligned} &\sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\tau_{G_1 \oplus G_2}(a, b_1) + \tau_{G_1 \oplus G_2}(a, b_2)] \\ &= \sum_{a \in V(G_1)} \sum_{b_1, b_2 \notin E(G_2)} [\{n_1n_2 - 1 - n_2d_{G_1}(a) - n_1d_{G_2}(b_1) + d_{G_1}(a)d_{G_2}(b_1)\} \\ &\quad + \{n_1n_2 - 1 - n_2d_{G_1}(a) - n_1d_{G_2}(b_2) + d_{G_1}(a)d_{G_2}(b_2)\}]. \end{aligned} \quad (43)$$

$$\text{Similarly} = 2n_1^2n_2\mu_2 - 2n_1\mu_2 - 4n_2e_1\mu_2 - n_1^2\bar{M}_1(G_2) + 2e_1\bar{M}_1(G_2).$$

Also, we take

$$\begin{aligned} &\sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] \\ &= 2 \sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [\{n_1n_2 - 1 - n_2d_{G_1}(a_1) - n_1d_{G_2}(b_1) + d_{G_1}(a_1)d_{G_2}(b_1)\} \\ &\quad + \{n_1n_2 - 1 - n_2d_{G_1}(a_2) - n_1d_{G_2}(b_2) + d_{G_1}(a_2)d_{G_2}(b_2)\}]. \end{aligned} \quad (44)$$

$$\text{Similarly} = 4n_1n_2\mu_1\mu_2 - 4\mu_1\mu_2 - 2n_2\mu_2\bar{M}_1(G_1) - 2n_1\mu_1\bar{M}_1(G_2)$$

$$+ 2 \sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [d_{G_1}(a_1)d_{G_2}(b_1) + d_{G_1}(a_2)d_{G_2}(b_2)].$$

Again, we take

$$\sum_{b \in V(G_2)} \sum_{a_1, a_2 \notin E(G_1)} [\tau_{G_1 \oplus G_2}(a_1, b) + \tau_{G_1 \oplus G_2}(a_2, b)], \quad (45)$$

$$\text{Similarly} = 2n_1 n_2^2 \mu_1 - 2n_2 \mu_1 - n_2^2 \overline{M}_1(G_1) - 4n_1 e_2 \mu_1 + 2e_2 \overline{M}_1(G_1).$$

We further take (null case)

$$N = \sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \uparrow b_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] = 0. \quad (46)$$

Furthermore, we take (also null case)

$$N = \sum_{\substack{b_1, b_2 \notin E(G_1) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} \sum_{\substack{a_1, a_2 \notin E(G_2) \\ a_1 \neq a_2 \wedge a_1 \uparrow a_2}} [\tau_{G_1 \oplus G_2}(a_1, b_1) + \tau_{G_1 \oplus G_2}(a_2, b_2)] = 0. \quad (47)$$

Consequently,

$$\begin{aligned} \overline{ZC}_1(G_1 \oplus G_2) &= (2e_2 - n_2^2 - 2n_2 \mu_2) \overline{M}_1(G_1) + (2e_1 - n_1^2 - 2n_1 \mu_1) \overline{M}_1(G_2) \\ &+ 2\mu_1(n_1 n_2^2 - n_2 - 2n_1 e_2) + 2\mu_2(n_1^2 n_2 - n_1 - 2n_2 e_1) + 4\mu_1 \mu_2 (n_1 n_2 - 1) \\ &+ 2 \sum_{\substack{a_1, a_2 \notin E(G_1) \\ a_1 \neq a_2 \wedge a_1 \sim a_2}} \sum_{\substack{b_1, b_2 \notin E(G_2) \\ b_1 \neq b_2 \wedge b_1 \sim b_2}} [d_{G_1}(a_1) d_{G_2}(b_1) + d_{G_1}(a_2) d_{G_2}(b_2)]. \end{aligned} \quad (48)$$

□

## 4. Analysis and Conclusions

In this section, we compute the analysis for the first Zagreb connection index ( $ZC_1$ ) and coindex ( $\overline{ZC}_1$ ) of product on graphs such as tensor product, strong product, symmetric difference, and disjunction with the help of Tables 1–5 which are constructed by using numerical values of the aforesaid Zagreb index and coindex, respectively. The graphical depictions of the exact formulae and upper bounds for  $ZC_1$  and  $\overline{ZC}_1$  are also depicted in Figures 5–9. Moreover, in this section, for particular cases of main results, we compute all the results in the shape of upper bounds as  $\{C_3, C_4\}$  are not free graphs and  $G_1$  and  $G_2$  are two undirected graphs.

**4.1. Tensor Product.** Let  $P_m$  and  $P_n$  be two particular alkanes called by paths, then the tensor product ( $P_m \otimes P_n$ ) is obtained by the product of  $P_m$  and  $P_n$ . For  $m = 3$  and  $n = 3$ , see Figure 10.

Using Theorem 1, the exact formulae for the first Zagreb connection index ( $ZC_1(\theta_1)$ ) and first Zagreb connection coindex ( $\overline{ZC}_1(\psi_1)$ ) of tensor product are obtained as follows:

- (a)  $ZC_1(P_m \otimes P_n) \leq 64mn - 138m - 138n + 292$
- (b)  $\overline{ZC}_1(P_m \otimes P_n) \leq 4mn - 4m - 6n + 6$

The upper bounds for the first Zagreb connection index ( $ZC_1(\lambda_1)$ ) of tensor product are obtained as follows [39]:

$$ZC_1(P_m \otimes P_n) \leq 64mn - 56m - 124n + 100. \quad (49)$$

Table 1 and Figure 5 depict the numerical and graphical behaviours of the analysis between exact formulae and upper bounds for the first Zagreb connection index and coindex of tensor product by using values  $m = n$ .

**4.2. Strong Product.** Let  $P_m$  and  $P_n$  be two particular alkanes called by paths, then the strong product ( $P_m \otimes P_n$ ) is

TABLE 1: Analysis for index and coindex of exact formulae and upper bounds of  $\theta_1 = P_m \circ P_n$ ,  $\lambda_1 = P_m \otimes P_n$ , and  $\psi_1 = P_m \oplus P_n$ , respectively.

$(m = n)$	$ZC_1(\theta_1)$	$ZC_1(\lambda_1)$	$\overline{ZC}_1(\psi_1)$
1	80	-16	0
2	-4	-4	2
3	40	136	12
4	212	404	30
5	512	800	56
6	940	1324	90
7	1496	1976	132
8	2180	2756	182
9	2992	3664	240
10	3932	4700	306
11	5000	5864	380
12	6196	7156	462
13	7520	8576	552
14	8972	10124	650
15	10552	11800	756

TABLE 2: Analysis for index and coindex of exact formulae and upper bounds of  $\theta_2 = P_m \circ P_n$ ,  $\lambda_2 = P_m \otimes P_n$ , and  $\psi_2 = P_m \oplus P_n$ , respectively.

$(m = n)$	$ZC_1(\theta_2)$	$ZC_1(\lambda_2)$	$\overline{ZC}_1(\psi_2)$
1	98	-13	22
2	56	-40	88
3	218	91	234
4	584	644	460
5	1154	1979	766
6	1928	4552	1152
7	2906	8915	1618
8	4088	15716	2164
9	5474	25699	2790
10	7064	39704	3496
11	8858	58667	4282
12	10856	83620	5148
13	13058	115691	6094
14	15464	156104	7120
15	18074	206179	8226

obtained by the product of  $P_m$  and  $P_n$ . For  $m = 4$  and  $n = 4$ , see Figure 11.

Using Theorem 2, the exact formulae for the first Zagreb connection index ( $ZC_1(\theta_2)$ ) and first Zagreb connection coindex ( $\overline{ZC}_1(\psi_2)$ ) of strong product are obtained as follows:

- (a)  $ZC_1(P_m \otimes P_n) \leq 102mn - 174m - 174n + 344$
- (b)  $\overline{ZC}_1(P_m \otimes P_n) \leq 40mn - 26m - 28n + 36$

The upper bounds for first Zagreb connection index ( $ZC_1(\lambda_2)$ ) of strong product are obtained as follows [39]:

$$ZC_1(P_m \otimes P_n) \leq 4mn^3 + 16mn^2 - 5mn - 4m - 12n^3 - 40n^2 + 24n + 4. \tag{50}$$

Table 2 and Figure 6 depict the numerical and graphical behaviours of the analysis between exact formulae and upper

TABLE 3: Analysis for index and coindex of exact formulae and upper bounds of  $\theta_3 = P_m \circ P_n$ ,  $\lambda_3 = P_m \otimes P_n$ , and  $\psi_3 = P_m \oplus P_n$ , respectively.

$(m = n)$	$ZC_1(\theta_3)$	$ZC_1(\lambda_3)$	$\overline{ZC}_1(\psi_3)$
1	12	33	0
2	4	34	0
3	116	153	44
4	912	390	264
5	4348	745	888
6	15332	1218	2240
7	44004	1809	4740
8	108736	2518	8904
9	239852	3345	15344
10	484068	4290	24768
11	909652	5353	37980
12	1612304	6534	55880
13	2721756	7833	79464
14	4409092	9250	109824
15	6894788	10785	148148

TABLE 4: Analysis for index and coindex of exact formulae and upper bounds of  $\theta_4 = P_m \circ P_n$ ,  $\lambda_4 = P_m \otimes P_n$ , and  $\psi_4 = P_m \oplus P_n$ , respectively.

$(m = n)$	$ZC_1(\theta_4)$	$ZC_1(\lambda_4)$	$\overline{ZC}_1(\psi_4)$
1	0	141	0
2	0	40	-2
3	40	189	40
4	492	588	378
5	2928	1237	1456
6	11680	2136	3910
7	36120	3285	8568
8	93660	4684	16450
9	213472	6333	28768
10	440928	8232	46926
11	842760	10381	72520
12	1512940	12780	107338
13	2579280	15429	153360
14	4210752	18328	212758
15	6676153	21477	287896

TABLE 5: Particular numeric values of obtained results for index and coindex.

Product graphs	$ZC_1(\theta)$	$ZC_1(\lambda)$	$\overline{ZC}_1(\psi)$
$(P_3 \circ P_3)$	40	136	12
$(P_4 \otimes P_4)$	584	644	460
$(P_3 \oplus P_3)$	116	153	44
$(P_3 \oplus P_3)$	40	189	40

bounds for the first Zagreb connection index and coindex of strong product by using values  $m = n$ .

4.3. *Symmetric Difference.* Let  $P_m$  and  $P_n$  be two particular alkanes called by paths, then the symmetric difference ( $P_m \ominus P_n$ ) is obtained by the product of  $P_m$  and  $P_n$ . For  $m = 3$  and  $n = 3$ , see Figure 12.

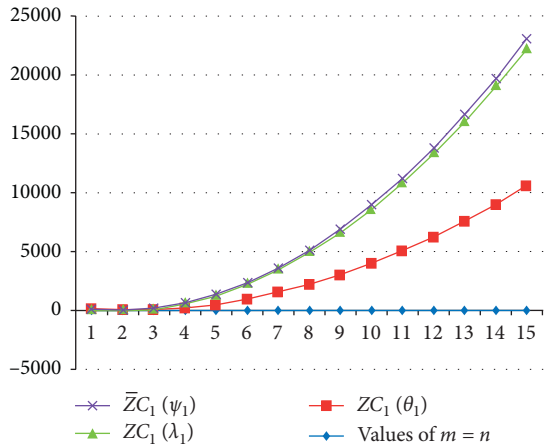


FIGURE 5: Analysis for index and coindex of tensor product with respect to Table 1.

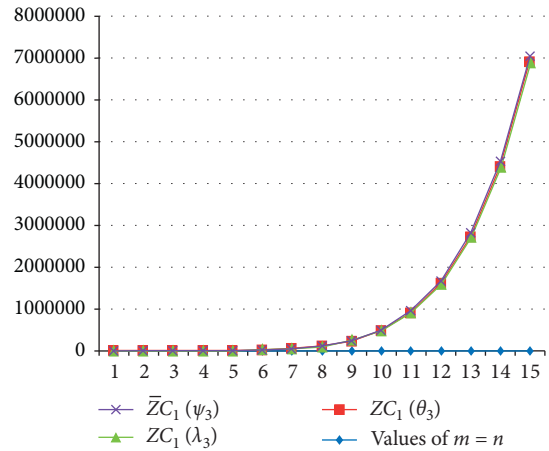


FIGURE 7: Analysis for index and coindex of symmetric difference with respect to Table 3.

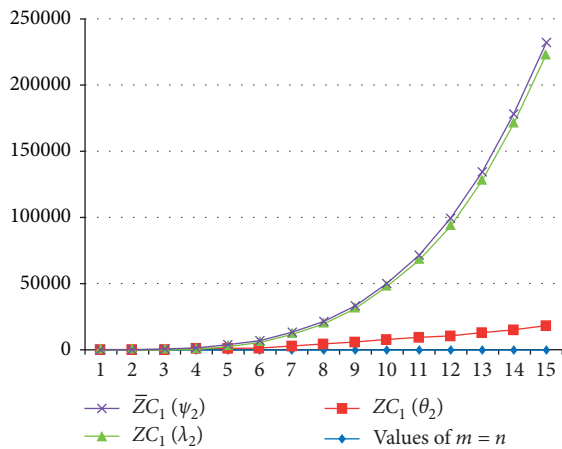


FIGURE 6: Analysis for index and coindex of strong product with respect to Table 2.

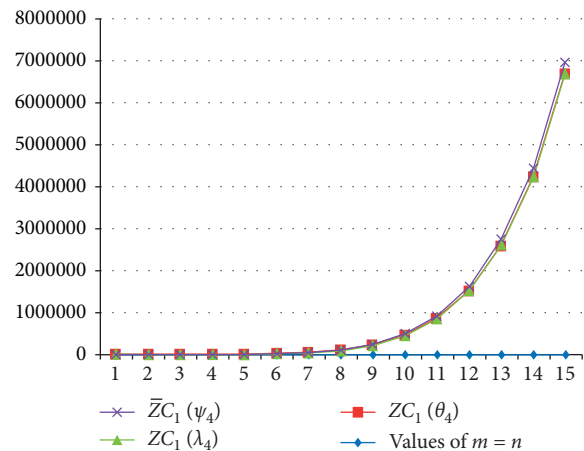


FIGURE 8: Analysis for index and coindex of disjunction with respect to Table 4.

Using Theorem 3, the exact formulae for the first Zagreb connection index ( $ZC_1(\theta_3)$ ) and first Zagreb connection

coindex ( $\bar{Z}C_1(\psi_3)$ ) of symmetric difference are obtained as follows:

$$ZC_1(P_m \ominus P_n) \leq m^3 n^3 - 4m^3 n^2 + 8m^3 n - 6m^3 - 4m^2 n^3 + 22m^2 n^2 - 52m^2 n + 44m^2 + 8mn^3 - 52mn^2 + 137mn - 128m - 6n^3 + 44n^2 - 128n + 128, \tag{51}$$

$$\bar{Z}C_1(P_m \ominus P_n) \leq 4m^2 n^2 - 9m^2 n + 5m^2 - 9mn^2 + 20mn - 12m + 5n^2 - 12n + 8. \tag{52}$$

The upper bounds for first Zagreb connection index ( $ZC_1(\lambda_3)$ ) of symmetric difference are obtained as follows [40]:

$$ZC_1(P_m \ominus P_n) \leq 59mn - 88m - 88n + 150. \tag{53}$$

Table 3 and Figure 7 depict the numerical and graphical behaviours of the analysis between exact formulae and upper bounds for first Zagreb connection index and coindex of symmetric difference by using values  $m = n$ .

4.4. Disjunction. Let  $P_m$  and  $P_n$  be two particular alkanes called by paths, then the disjunction ( $P_m \oplus P_n$ ) is obtained by the product of  $P_m$  and  $P_n$ . For  $m = 3$  and  $n = 3$ , see Figure 13.

Using Theorem 4, the exact formulae for the first Zagreb connection index ( $ZC_1(\theta_4)$ ) and first Zagreb connection coindex ( $\bar{Z}C_1(\psi_4)$ ) of disjunction are obtained as follows:

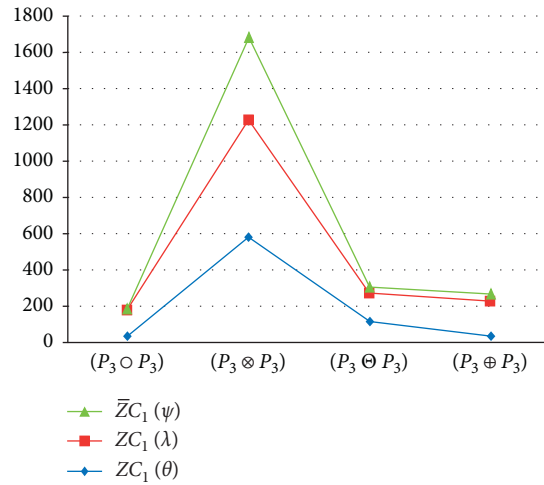


FIGURE 9: Analysis of obtained results for product on graphs with respect to Table 5.

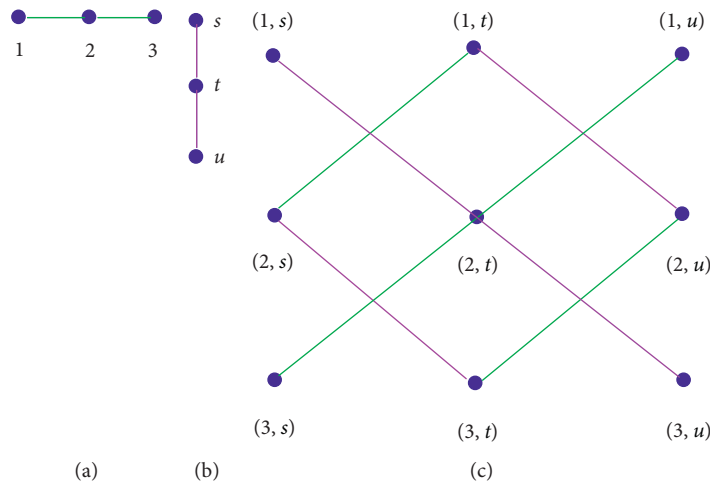


FIGURE 10: (a)  $G_1 \cong P_3$  (b)  $G_2 \cong P_3$ , and (c) tensor product  $(P_3 \otimes P_3)$ .

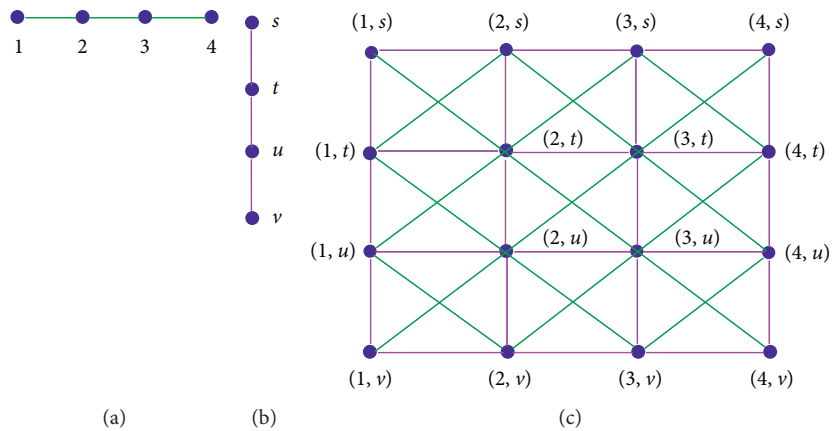


FIGURE 11: (a)  $G_1 \cong P_4$  (b)  $G_2 \cong P_4$ , and (c) strong product  $(P_4 \otimes P_4)$ .

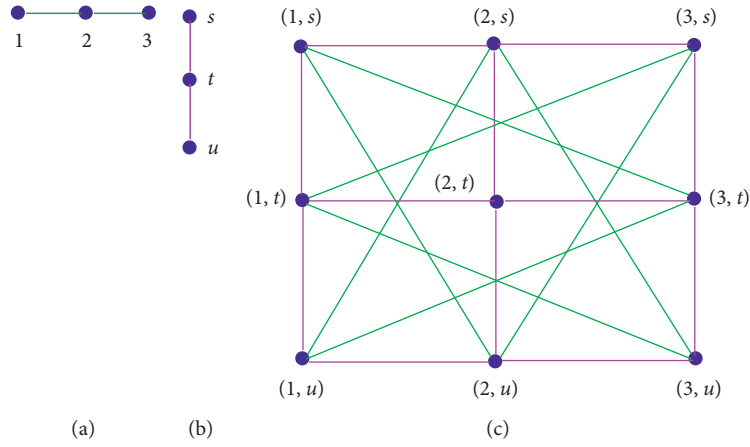


FIGURE 12: (a)  $G_1 \cong P_3$  (b)  $G_2 \cong P_3$ , and (c) symmetric difference  $(P_3 \oplus P_3)$ .

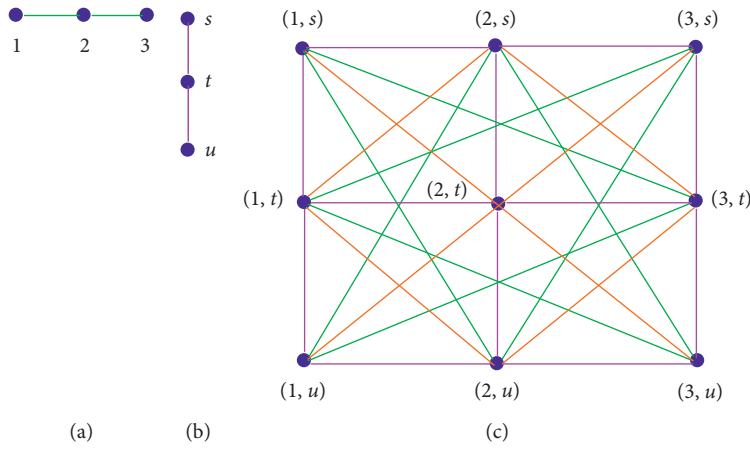


FIGURE 13: (a)  $G_1 \cong P_3$  (b)  $G_2 \cong P_3$ , and (c) disjunction  $(P_3 \oplus P_3)$ .

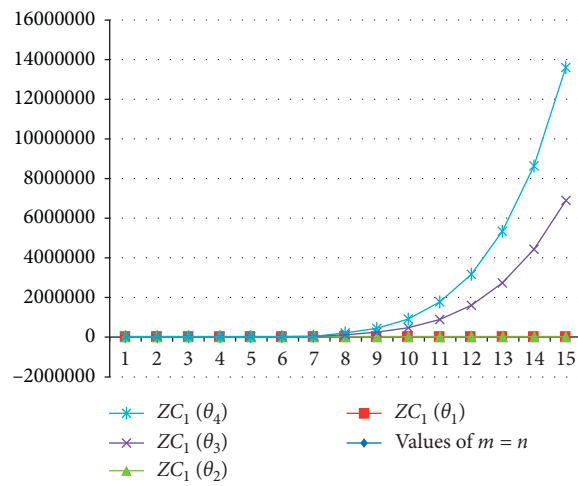


FIGURE 14: Analysis for  $ZC_1(\theta)$  of exact formula.

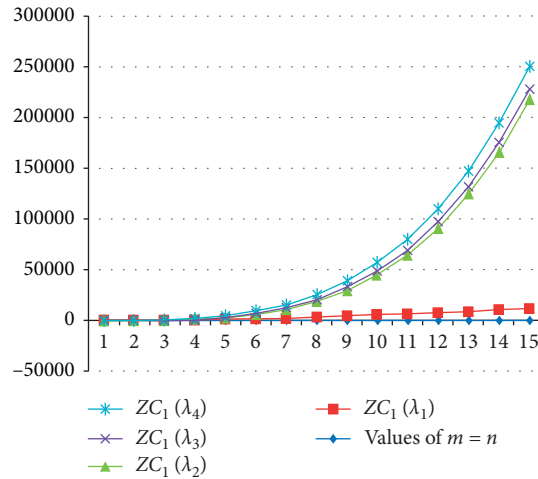


FIGURE 15: Analysis for  $ZC_1(\lambda)$  of upper bound.

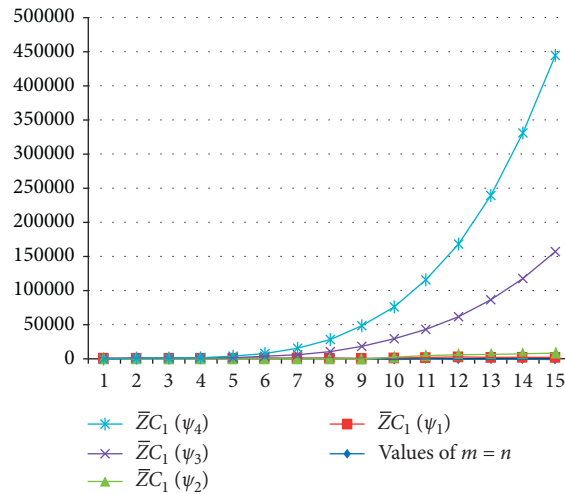


FIGURE 16: Analysis for  $\bar{Z}C_1(\psi)$  of exact formula.

$$ZC_1(P_m \oplus P_n) \leq m^3 n^3 - 4m^3 n^2 + 8m^3 n - 6m^3 - 4m^2 n^3 + 14m^2 n^2 - 28m^2 n + 20m^2 + 8mn^3 - 28mn^2 + 41mn - 40m - 6n^3 + 36n^2 - 40n + 28, \tag{54}$$

$$\bar{Z}C_1(P_m \otimes P_n) \leq 8m^2 n^2 - 19m^2 n + 7m^2 - 19mn^2 + 36mn - 2m + 7n^2 - 4n - 14. \tag{55}$$

The upper bounds for first Zagreb connection index ( $ZC_1(\lambda_4)$ ) of disjunction are obtained as follows [39]:

$$ZC_1(P_m \oplus P_n) \leq 125mn - 238m - 238n + 492. \tag{56}$$

Table 4 and Figure 8 depict the numerical and graphical behaviours of the analysis between exact formulae and upper bounds for the first Zagreb connection index and coindex of disjunction by using values  $m = n$ .

Now, from Tables 1–5 and Figures 5–9 and 14–16, we close our discussion with the following conclusions:

The behaviours of all the connection number-based Zagreb index and coindex for operations on graphs

such as tensor product, strong product, symmetric difference, and disjunction are increased in the following order, respectively, as  $\bar{Z}C_1(\psi) \geq ZC_1(\lambda) \geq ZC_1(\theta)$ .

For increasing values of  $m$  and  $n$ , the upper bound for the first Zagreb connection index of products on graphs are working rapidly than all the exact formulae for the first Zagreb connection index, respectively.

In certain intervals of the values of  $m$  and  $n$ , all the first Zagreb connection coindices attain the maximum values on increasing values of  $m$  and  $n$ . In Figures 5–8, we analyse that the first Zagreb connection coindex

attains more upper layer than other TIs in all the operations.

Table 5 and Figure 9 interpret the particular analysis of the obtained results for index and coindex on operations such as tensor product, strong product, symmetric difference, and disjunction. This particular analysis also concludes that the first Zagreb connection coindex attains more upper layer than other TIs in all the operations.

In particular, Figures 14–16 interpret the exact formula for the first Zagreb connection index, upper bound for the first Zagreb connection index, and exact formula for the first Zagreb connection coindex which are dominant on operations from tensor product to disjunction, respectively. In addition, we analyse that the first Zagreb connection coindex of operation disjunction has attained more upper layer than all the other operations for connection number-based index and coindex.

The investigation of these indices and coindices for the resultant graphs obtained from other operations of graphs (subtraction, switching, zig-zag product, addition, rooted product, modular product etc.) is still open.

## Data Availability

All data used to support the findings of this study are included within the article. However, additional data will be made available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Humanities and Social Science Project of Anhui Provincial Education Department, Research on social crisis management in Anhui cities under the background of big data (subject no: SK2016A0233).

## References

- [1] R. Todeschini and V. Consonni, *Molecular Descriptors for Chemoinformatics*, Wiley-VCH Verlag GmbH, Weinheim, Germany, 2009.
- [2] A. R. Matamala and E. Estrada, "Generalised topological indices: optimisation methodology and physico-chemical interpretation," *Chemical Physics Letters*, vol. 410, no. 4-6, pp. 343–347, 2005.
- [3] H. Gonzalez-Diaz, S. Vilar, L. Santana, and E. Uriarte, "Medicinal chemistry and bioinformatics - current trends in drugs discovery with networks topological indices," *Current Topics in Medicinal Chemistry*, vol. 7, no. 10, pp. 1015–1029, 2007.
- [4] G. Rücker and C. Rücker, "On topological indices, boiling points, and cycloalkanes," *Journal of Chemical Information and Computer Sciences*, vol. 39, no. 5, pp. 788–802, 1999.
- [5] B. Furtula and I. Gutman, "A forgotten topological index," *Journal of Mathematical Chemistry*, vol. 53, no. 4, pp. 1184–1190, 2015.
- [6] A. R. Ashrafi, T. Došlić, and A. Hamzeh, "The Zagreb coindices of graph operations," *Discrete Applied Mathematics*, vol. 158, no. 15, pp. 1571–1578, 2010.
- [7] K. C. Das, A. Yurttas, M. Togan, A. S. Cevik, and I. N. Cangul, "The multiplicative Zagreb indices of graph operations," *Journal of Inequalities and Applications*, vol. 90, p. 14, 2013.
- [8] K. Xu, K. C. Das, and K. Tang, "On the multiplicative Zagreb coindex of graphs," *Opuscula Mathematica*, vol. 33, no. 1, pp. 191–204, 2013.
- [9] S. M. Hosamani and N. Trinajstić, "On reformulated zagreb coindices, research gate, T 09:07:00 UTC," 2015.
- [10] K. Pattabiraman, S. Nagarajan, and M. Chendrasekharan, "Zagreb indices and coindices of product graphs," *Journal of Progressive Research in Mathematics*, vol. 10, pp. 80–91, 2015.
- [11] B. Basavanagoud, I. Gutman, and C. S. Gali, "On second Zagreb index and coindex of some derived graphs," *Kragujevac Journal of Science*, vol. 37, pp. 113–121, 2015.
- [12] N. De, S. M. A. Nayeem, and A. Pal, "The F-coindex of some graph operations," *SpringerPlus*, vol. 5, no. 221, p. 13, 2016.
- [13] M. Veylaki, M. J. Nikmehr, and H. A. Tavallaee, "The third and hyper-Zagreb coindices of some graph operations," *Journal of Applied Mathematics and Computing*, vol. 50, no. 1-2, pp. 315–325, 2016.
- [14] H. Hua and S. Zhang, "Relations between Zagreb coindices and some distance-based topological indices," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 68, pp. 199–208, 2012.
- [15] M. Azari and A. Iranmanesh, "Chemical graphs constructed from rooted product and their Zagreb indices," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 70, pp. 901–919, 2013.
- [16] N. De, A. Pal, and S. M. A. Nayeem, "On some bounds and exact formulae for connective eccentric indices of graphs under some graph operations," *International Journal of Mathematical Combinatorics*, vol. 2014, Article ID 579257, 5 pages, 2014.
- [17] Z. Luo and J. Wu, "Zagreb eccentricity indices of the generalized Hierarchical product graphs and their applications," *Journal of Applied Mathematics*, vol. 2014, Article ID 241712, 8 pages, 2014.
- [18] W. Gao, M. K. Jamil, and M. R. Farahani, "The hyper-Zagreb index and some graph operations," *Journal of Applied Mathematics and Computing*, vol. 54, no. 1-2, pp. 263–275, 2017.
- [19] I. Gutman and N. Trinajstić, "Graph theory and molecular orbitals. Total  $\phi$ -electron energy of alternant hydrocarbons," *Chemical Physics Letters*, vol. 17, no. 4, pp. 535–538, 1972.
- [20] I. Gutman, "Degree-based topological indices," *Croatica Chemica Acta*, vol. 86, no. 4, pp. 351–361, 2013.
- [21] J.-B. Liu, S. Javed, M. Javaid, and K. Shabbir, "Computing first general Zagreb index of operations on graphs," *IEEE Access*, vol. 7, pp. 47494–47502, 2019.
- [22] J.-B. Liu, M. Javaid, and H. M. Awais, "Computing Zagreb indices of the subdivision-related generalized operations of graphs," *IEEE Access*, vol. 7, pp. 105479–105488, 2019.
- [23] X. Zhang, H. M. Awais, M. Javaid, and M. K. Siddiqui, "Multiplicative zagreb indices of molecular graphs," *Journal of Chemistry*, vol. 2019, pp. 1–19, 2019.
- [24] H. M. Awais, M. Javaid, and M. Jamal, "Forgotten index of generalized F-sum graphs," *Journal of Prime Research in Mathematics*, vol. 15, pp. 115–128, 2019.



- [25] H. M. Awais, M. Javaid, and A. Raheem, "Hyper-Zagreb index of graphs based on generalized subdivision related operations," *Punjab University Journal of Mathematics*, vol. 52, no. 5, pp. 89–103, 2019.
- [26] G. Hong, Z. Gu, M. Javaid, H. M. Awais, and M. K. Siddiqui, "Degree-based topological invariants of metal-organic networks," *IEEE Access*, vol. 8, pp. 68288–68300, 2020.
- [27] H. M. Awais, M. Jamal, and M. Javaid, "Topological properties of metal-organic frameworks," *Main Group Metal Chemistry*, vol. 43, no. 1, pp. 67–76, 2020.
- [28] H. M. Awais, M. Javaid, and A. Akbar, "First general Zagreb index of generalized F-sum graphs," *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 2954975, 2020.
- [29] Y.-M. Chu, S. Javed, M. Javaid, and M. Kamran Siddiqui, "On bounds for topological descriptors of  $\phi$ -sum graphs," *Journal of Taibah University for Science*, vol. 14, no. 1, pp. 1288–1301, 2020.
- [30] A. Ali and N. Trinajstić, "A novel/old modification of the first Zagreb index," *Molecular Informatics*, vol. 37, pp. 1–7, 2018.
- [31] U. Ali, M. Javaid, and A. M. Alanazi, "Computing analysis of connection-based indices and coindices for product of molecular networks," *Symmetry*, vol. 12, no. 1320, 23 pages, 2020.
- [32] G. Ducoffe, R. Marinescu-Ghemeci, C. Obeja, A. Popa, and R. M. Tache, "Extremal graphs with respect to the modified first Zagreb connection Index," in *Proceedings of the 16th cologne-twente workshop on graphs and combinatorial optimization*, Paris, France, June 2018.
- [33] Z. Shao, I. Gutman, Z. Li, S. Wang, and P. Wu, "Leap Zagreb indices of trees and unicyclic Graphs," *Communications in Combinatorics and Optimization*, vol. 3, pp. 179–194, 2018.
- [34] A. M. Naji and N. D. Soner, "The first leap Zagreb index of some graph operations," *International Journal of Applied Graph Theory*, vol. 2, no. 1, pp. 7–18, 2018.
- [35] I. Gutman, E. Milovanović, and I. Milovanović, "Beyond the Zagreb indices," *AKCE International Journal of Graphs and Combinatorics*, vol. 22, 2018, (In Press).
- [36] J.-H. Tang, U. Ali, M. Javaid, and K. Shabbir, "Zagreb connection indices of subdivision and semi-total point operations on graphs," *Journal of Chemistry*, vol. 2019, Article ID 9846913, 14 pages, 2019.
- [37] U. Ali, M. Javaid, and A. Kashif, "Modified Zagreb connection indices of the T-sum graphs," *Main Group Metal Chemistry*, vol. 43, no. 1, pp. 43–55, 2020.
- [38] J. Cao, U. Ali, M. Javaid, and C. Huang, "Zagreb connection indices of molecular graphs based on operations," *Complexity*, vol. 2020, Article ID 7385682, 15 pages, 2020.
- [39] U. Ali and M. Javaid, "Upper bounds of Zagreb connection indices of tensor and strong product on graphs," *Punjab University Journal of Mathematics*, vol. 52, no. 4, pp. 89–100, 2020.
- [40] U. Ali and M. Javaid, "Zagreb connection indices of disjunction and symmetric difference operations on graphs," *Journal of Prime Research in Mathematics*, vol. 16, no. 2, pp. 1–15, 2020.
- [41] D. Gong, Y. Han, and J. Sun, "A novel hybrid multi-objective artificial bee colony algorithm for blocking lot-streaming flow shop scheduling problems," *Knowledge-Based Systems*, vol. 16, no. 2, pp. 1–37, 2018.
- [42] D. Gong, B. Xu, Y. Zhang, Y. Guo, and S. Yang, "A similarity-base cooperative co-evolutionary algorithm for dynamic interval multi-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 17, 2019.
- [43] J. Sun, D. Gong, and X. Sun, "Solving interval multi-objective optimization problems using evolutionary algorithms with preference polyhedron," in *Proceedings of the Conference Paper in Information Sciences*, Kuala Lumpur, Malaysia, November 2011.
- [44] B. Xu, Y. Zhang, D. Gong, Y. Guo, and M. Rong, "Environment sensitivity-based cooperative co-evolutionary algorithms for dynamic multi-objective optimization," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 14, no. 1, 2020.
- [45] F. Harary, *Graph Theory*, Addison-Wesley, Boston, MA, USA, 1994.
- [46] W. Imrich and S. Klavzar, *Product Graphs: Structure and Recognition*, Wiley, New York, NY, USA, 2000.
- [47] I. Gutman, B. Ruscic, N. Trinajstić, and C. F. Wilson, "Graph theory and molecular orbitals. XII. Acyclic polyenes," *The Journal of Chemical Physics*, vol. 62, no. 9, pp. 3399–3405, 1975.
- [48] Klavzar, S. ar and I. Gutman, "Selected properties of the Schultz molecular topological index," *Journal of Chemical Information and Modeling*, vol. 36, pp. 1001–1003, 1996.
- [49] U. Ali, M. Javaid, and Y.-M. Chu, "Computing relations between connection based modified indices and coindices for product of molecular-networks," *Journal of Molecular Liquids*, vol. 46, no. 1, 2021, (In Press).
- [50] D. B. West, *An Introduction to Graph Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.

## Research Article

# A Network Flow Algorithm for Solving Generalized Assignment Problem

Yongwen Hu <sup>1,2</sup> and Qunpo Liu <sup>3,4</sup>

<sup>1</sup>Key Laboratory of Power System Design and Test for Electrical Vehicle, Hubei University of Arts and Science, Xiangyang 441053, China

<sup>2</sup>School of Mechanical Engineering, Hubei University of Arts and Science, Xiangyang 441053, China

<sup>3</sup>Department of Robotics Engineering, Henan Polytechnic University, Jiaozuo 454003, China

<sup>4</sup>Henan International Joint Laboratory of Direct Drive and Control of Intelligent Equipment, Jiaozuo 454003, China

Correspondence should be addressed to Qunpo Liu; [lqpny@hpu.edu.cn](mailto:lqpny@hpu.edu.cn)

Received 14 July 2020; Revised 24 August 2020; Accepted 24 December 2020; Published 12 January 2021

Academic Editor: M Javaid

Copyright © 2021 Yongwen Hu and Qunpo Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The generalized assignment problem (GAP) is an open problem in which an integer  $k$  is given and one wants to assign  $k'$  agents to  $k$  ( $k' \leq k$ ) jobs such that the sum of the corresponding cost is minimal. Unlike the traditional  $\mathcal{K}$ -cardinality assignment problem, a job can be assigned to many, but different, agents and an agent may undertake several, but different, jobs in our problem. A network model with a special structure of GAP is given and an algorithm for GAP is proposed. Meanwhile, some important properties of the GAP are given. Numerical experiments are implemented, and the results indicate that the proposed algorithm can globally and efficiently optimize the GAP with a large range cost.

## 1. Introduction

Assignment problem (AP) is one of the fundamental combinatorial optimization problems with various applications in real life. The classical AP is proved to be an  $\mathcal{NP}$ -hard problem, and it deals with a situation of assigning  $n$  jobs to  $n$  agents such that each job must be processed by exactly one agent and vice versa. And the popular systematic method for solving classical AP is Kuhn's Hungarian method [1, 2] which can be extended for general network flow problems in a polynomial time. However, it is quite difficult to ensure that the number of jobs is exactly equal to the number of agents in real-life situation. Furthermore, jobs and agents are set to be unique in the classical AP, which leads to any row (column) in  $c_{ij}, \forall i \in \mathbf{M}, j \in \mathbf{N}$  ( $\mathbf{M}$  is an index set of agents and  $\mathbf{N}$  is an index set of jobs) being different from every other row (column). In fact, although all agents are unique, some jobs can be identical and vice versa in practical applications. Kennington and Wang [3] give some typical problems, such as manpower planning,

scheduling, and planning, in which more than one agent is assigned to the same job group. Also, a shortest augmenting path algorithm is presented for solving them.

Dell'Amico and Martello [4] considered a generalization of AP (called  $\mathcal{K}$ -cardinality linear assignment problem) where one wants to assign  $k$  (out of  $m$ ) agents to  $k$  (out of  $n$ ) jobs ( $k \leq \min(m, n)$ ) so that the sum of the corresponding cost is minimal. Some potential applications of  $\mathcal{K}$ -cardinality linear assignment problem can be found in [5, 6]. Further, the  $\mathcal{K}$ -cardinality assignment problem has various applications, for example, in assigning agents to machines when there are multiple alternatives and only a subset of agents and machines has to be assigned. In this paper, we consider a more generalized assignment problem (GAP) in which a cost matrix  $A_{m \times n}$  and positive integer  $k$  ( $0 < k \leq \min\{m, n\}$ ) are given and one wants to assign  $k'$  agents to  $k$  ( $k' \leq k$ ) jobs so that the sum of the corresponding cost is minimal.

Dell'Amico and Martello developed some algorithms for solving  $\mathcal{K}$ -cardinality linear assignment problem by min-

cost flow or shortest augmenting path techniques described in [4]. Also, heuristic processing techniques were given in [7]. Later, Volgenant [8] developed an algorithm for it by solving a sequence of classical AP. Recently, based on semi-Lagrangian relaxation, Belik and Jörnsten [9] proposed an algorithm for the  $\mathcal{K}$ -cardinality AP with large scale. Clearly, the GAP considered in this paper is the  $\mathcal{K}$ -cardinality assignment problem when  $k' = k$ . The GAP is an  $\mathcal{NP}$ -hard [10] problem. Later, Chu and Beasley [11] prove that GAP is  $\mathcal{NP}$ -complete.

Several algorithms have been proposed for GAP in the past few years. Most of the small sizes of the test instances have been solved by some exact methods such as branch and bound, branch and price, and cutting plane [12–16]. However, the exact methods are mostly unable to find an optimal solution within a reasonable computational time. Metaheuristic and heuristics have been proposed for solving GAP in order to reduce the computational time of the exact method, such as the combination of a heuristics and exact method [17], tabu search [18], simulated annealing [19], genetic algorithm [11, 20], differential evolution algorithm [21, 22], and bees algorithm [23].

Clearly, the GAP considered in this paper can be transformed into a classical AP when some dummy jobs or agents are introduced. And the Hungarian method and its variants [24–26] are presented for optimizing GAP recently. However, the transformation will result in a larger-scale problem and worse space complexity. Currently, several exact algorithms are presented for solving GAP with minimum cost, such as variable-fixing algorithm [27], branch and bound algorithm [12], and KM algorithm [28], but these algorithms can only solve GAP with some specific assumption.

This study focuses on GAP where the number of agents is not equal to the number of jobs. Specifically, one job can be assigned to many, but different, agents, and one agent may undertake many, but different, tasks. It is well known that AP is thought of as a special min-cost flow problem, which suggests that classical AP or GAP can be solved by some methods with network flow theory. Recently, an efficient algorithm was proposed for solving min-cost flow problem [29]. This paper attempts to give a variation of the algorithm presented in [29] for GAP based on the advantage of the special structure of the transformed network model.

This paper is organized as follows. Section 2 gives a mathematical formulation and network flow model for the GAP; in Section 3, we describe an algorithm for solving GAP. Also, the convergence of the proposed algorithm and some important features of GAP are given in this section; Section 4 presents numerical experiments to evaluate the performance of the proposed algorithm. Concluding remarks and some further research are provided in Section 5.

## 2. Mathematical Formulation and Network Flow Model

Given an  $m \times n$  matrix  $C = [c_{ij}]$  and an integer  $k$  ( $0 < k \leq \min\{m, n\}$ ), the GAP can be formulated as follows:

$$\text{GAP} \left\{ \begin{array}{l} \min \quad z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \sum_{j=1}^n x_{ij} = y_i, \quad i = 1, 2, \dots, m, \\ \sum_{i=1}^m x_{ij} = z_j, \quad j = 1, 2, \dots, n, \\ \text{s.t.} \quad \sum_{i=1}^m y_i = k, \\ \sum_{j=1}^n z_j = k, \\ x_{ij} = 0 \text{ or } 1, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n, \end{array} \right. \quad (1)$$

where  $k, y_i, z_j \in I; c_{ij} \geq 0; i \in \mathbf{M} = \{1, 2, \dots, m\}; j \in \mathbf{N} = \{1, 2, \dots, n\}; x_{ij} = 1$  if agent  $i$  is assigned to job  $j$ . First, we consider the case that an agent can be assigned to more than one job, but a job is assigned to exactly one agent, that is,  $z_j = 1, \forall j \in \mathbf{N}$  in (1). We leave the situation that a group job can be processed by more than one agent later.

It is well known that classical AP is a bipartite matching problem which can be transformed into a maximum flow problem in a simple network. To transform the GAP defined on an undirected graph  $G = (\mathbf{M} \cup \mathbf{N}, \mathbf{A})$ , we create a directed version of  $G$  by designating each arc in  $\mathbf{A}$  as pointing nodes in  $\mathbf{M}$  to nodes in  $\mathbf{N}$ . Then, we introduce a source node  $s$  and a sink node  $t$ , with  $(s, i), \forall i \in \mathbf{M}$ , and  $(j, t), \forall j \in \mathbf{N}$ . We refer to the transformed network as  $G' = (\mathbf{N}', \mathbf{A}', \mathbf{C}, \mathbf{U})$  with a cost set  $\mathbf{C}$  and a capacity set  $\mathbf{U}$  associated with each arc  $(i, j) \in \mathbf{A}'$ . Let  $c(s, i) = 0, c(j, t) = 0, 0 \leq u(s, i) \leq y_i, 0 \leq u(j, t) \leq 1, 0 \leq u(i, j) \leq 1, \forall i \in \mathbf{M}, j \in \mathbf{N}$ .

Figure 1 illustrates the transformation. Mathematically, the transformed model is given by

$$\text{MCF} \left\{ \begin{array}{l} \min \quad z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \sum_{i=1}^m x_{si} = k, \quad i \in \mathbf{M}, \\ \sum_{j=1}^n x_{jt} = -k, \quad j \in \mathbf{N}, \\ \sum_j x_{ij} - \sum_i x_{ij} = 0, \quad \forall i, j \in \mathbf{M} \cup \mathbf{N}, \\ 0 \leq x_{ij} \leq 1, \quad \forall i, j \in \mathbf{M} \cup \mathbf{N} \cup \{s\} \cup \{t\}. \end{array} \right. \quad (2)$$

It is easy to deal with the case that some identical jobs can be grouped into a set. More specifically, if  $1'$  is a group set which consists of  $z_{1'}$  jobs, then the capacity of  $(1', t)$  is to be set to  $z_{1'}$  instead of 1.

According to the transformation mentioned above and the *Integrality Theorem* [30], it is easy to get the following theorem.

**Theorem 1.** *The GAP is equivalent to the MCF.*

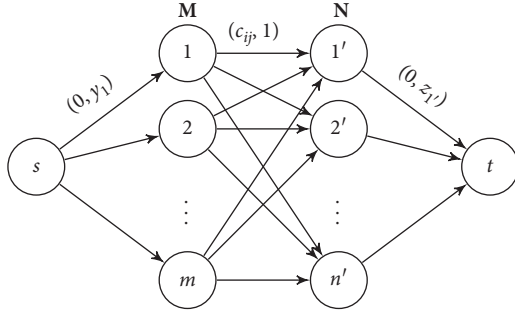


FIGURE 1: Unit capacity network flow model.

Theorem 1 indicates that we can find an integral flow of value  $k$  in the transformed network in order to get an optimal solution of GAP. Although several algorithms are presented for optimizing minimum-cost flow problem, both cost and capacity are to be considered for iteratively defining residual network. Hu et al. [29] introduced an approach that can solve min-cost flow. Unlike the traditional algorithms, the proposed algorithm in [29] can find an augmenting path in the original network by updating node potentials. We will describe an algorithm for optimizing GAP based on the network model defined above in the next section.

### 3. Proposed Algorithm for Solving GAP

Hu et al. [29] gave a detail algorithm for solving minimum-cost flow problem by introducing some optimality conditions. Since the transformed network flow model, shown in Figure 1, is a simple network, a modified version of the algorithm in [29] is introduced for GAP for self-contained in this section.

Let us consider a dual problem of problem (2)

$$\text{DP} \left\{ \begin{array}{l} \max \quad w = kp_s - kp_t + \sum_{(i,j) \in A'} u_{ij} p_{ij}, \\ s.t. \quad p_i - p_j + p_{ij} \leq c_{ij}, \quad \forall (i,j) \in A', \\ \quad p_i \text{ is free}, \quad \forall i \in \mathbf{M} \cup \mathbf{N} \cup \{s\} \cup \{t\}, \\ \quad p_{ij} \leq 0, \quad \forall (i,j) \in A', \end{array} \right. \quad (3)$$

where  $p_i$  and  $p_{ij}$  are the dual variables of MCF. And we refer to  $p_i$  and  $p_{ij}$  as the potential of node  $i$  and arc  $(i,j) \in A'$ , respectively.

Let  $x_{ij}$  and  $p = \{p_i, p_{ij}\}$  be the feasible solutions of MCF and DP, respectively. If  $x_{ij}$  and  $p = \{p_i, p_{ij}\}$  are optimal solutions, according to strong duality, we have

$$\begin{aligned} x_{ij} &= 0, & \text{if } p_i - p_j + p_{ij} < c_{ij}, \quad \forall i \in \mathbf{N}', \\ x_{ij} &= u_{ij}, & \text{if } p_i - p_j > c_{ij}, \quad \forall i \in \mathbf{N}', (i,j) \in A'. \end{aligned} \quad (4)$$

Since  $p_i$  is free  $\forall i \in \mathbf{N}'$ , let  $p_{ij} = \min\{0, c_{ij} + p_j - p_i\}$ , then  $p_{ij} \leq 0$  will be held. Therefore, a feasible solution  $p = \{p_i, p_{ij}\}$  of DP can be found. Thus, conditions (4) are equal to the conditions as follows:

$$\begin{aligned} x_{ij} &= 0, & \text{if } p_i - p_j < c_{ij}, \quad \forall i \in \mathbf{N}', \\ x_{ij} &= u_{ij}, & \text{if } p_i - p_j > c_{ij}, \quad \forall i \in \mathbf{N}', (i,j) \in A'. \end{aligned} \quad (5)$$

Clearly, an optimal solution  $x = x_{ij}$  will satisfy conditions (5). And we have the following theorem.

**Theorem 2.** Let  $x_{ij}$  be a feasible solution of MCF, then  $x_{ij}$  is an optimal solution if and only if the following conditions are satisfied:

$$\begin{cases} p_i - p_j < c_{ij} \implies x_{ij} = 0, \\ p_i - p_j > c_{ij} \implies x_{ij} = u_{ij}. \end{cases} \quad (6)$$

We refer the arc satisfying  $p_i - p_j = c_{ij}, \forall (i,j) \in A'$  to an admissible arc. Similarly, a network is an admissible network in which each arc is an admissible arc. Clearly,  $x_{ij} = 0, \forall (i,j) \in A'$  is an optimal solution with an integral flow of value 0. The proposed algorithm starts sending 0 unit of flow from  $s$  to  $t$ . Furthermore, by holding condition (6) at each iteration, the proposed algorithm can find at least an augmenting path from the source node  $s$  to the sink node  $t$  after updating node potential in finite iteration. Therefore, flow can be augmented when the current amount flow is not maximal.

Let  $\mu^+$  and  $\mu^-$  be a forward arc set and a backward arc set on an augmenting set  $\mu$ . Now we describe an algorithm for solving GAP corresponding the network  $G' = (\mathbf{N}', A', C, U)$ , where  $\mathbf{N}' = \mathbf{M} \cup \mathbf{N}$ , and we refer an admissible network to  $R$ .

Clearly, for an optimal solution  $x_{ij}^*, \forall i \in \mathbf{M}, j \in \mathbf{N}$ , agent  $i (i \in \mathbf{M})$  will be assigned to job  $j (j \in \mathbf{N})$  if  $x_{ij}^* = 1$ . We give the following definition in advance in order to get some important properties on GAP.

**Definition 1.** Let  $\mathbf{V}_i$  and  $\mathbf{V}_j$  be the node sets of any two augmenting paths (referred to  $\mu_i, \mu_j$ ) from source node  $s$  to sink node  $t$ , then  $\mu_i, \mu_j$  are independent augmenting paths if  $(\mathbf{V}_i \setminus \{s\} \setminus \{t\}) \cap (\mathbf{V}_j \setminus \{s\} \setminus \{t\}) = \emptyset$ .

**Theorem 3.** Suppose the current amount of flow obtained by Algorithm 1 is  $k' (k' < k)$ , GAP has  $C_q^{k-k'}$  optimal solutions if we can find  $q (q \geq k - k')$  independent augmenting paths in the following iteration.

Theorem 3 can be proved easily according to the detail algorithm and definition of independent augmenting path mentioned above, and here, we do not discuss the proof of Theorem 3.

**Theorem 4.** Let  $c_{i,j} = \min c_{ij}, \forall i \in \mathbf{M}, j \in \mathbf{N}$  be in a row set  $\mathbf{R}_o$  and in a column set  $\mathbf{C}_o$ , then at least an agent  $i' \in \mathbf{R}_o$  is to be assigned to a job in  $\mathbf{N}$ . Similarly, at least a job  $j' \in \mathbf{C}_o$  is to be processed by an agent in  $\mathbf{M}$ .



```

Initial settings:  $p_i = 0, \forall i \in \mathbf{N}', x_{ij} = 0, \forall (i, j) \in \mathbf{A}', \mathbf{S} = \{s\} \cup \mathbf{M}, \bar{\mathbf{s}} = \mathbf{N}' \setminus \mathbf{S}$ . Label source node  $(0, 1)$ . Let  $k$  be a given integer number
such that  $0 < k \leq \min\{m, n\}$ .
while  $\sum_{i=1}^m x_{si} < k, i \in \mathbf{M}$  do
  Remove all labels of node  $i, \forall i \in \mathbf{M} \cup \mathbf{N} \cup \{t\}$ .
  while  $t \notin \mathbf{S}$  do
     $p_i = p_i + \theta, \forall i \in \mathbf{S}$ , where  $\theta = \min_{k \in \mathbf{S} \cap \mathbf{M}, q \in \bar{\mathbf{S}} \cap \mathbf{N}} c_{kq} - p_k + p_q$ .
    if  $p_i - p_j = c_{ij}, \forall i \in \mathbf{S}, j \in \bar{\mathbf{S}}$  then
       $(i, j) \in R$ .
    end if
    if  $(i, j) \in R, x_{ij} = 0, \forall i \in \mathbf{S}, j \in \bar{\mathbf{S}}$  then
      Node  $j$  is labeled with  $(i, 1), \mathbf{S} = \mathbf{S} \cup \{j\}, \bar{\mathbf{S}} = \bar{\mathbf{S}} \setminus \{j\}$ .
    end if
    if  $(j, i) \in R, x_{ji} = 1, \forall i \in \mathbf{S}, j \in \bar{\mathbf{S}}$  then
      Node  $j$  is labeled with  $(-i, 1), \mathbf{S} = \mathbf{S} \cup \{j\}, \bar{\mathbf{S}} = \bar{\mathbf{S}} \setminus \{j\}$ .
    end if
  end while
  
$$x_{ij} = \begin{cases} x_{ij} + 1, & (i, j) \in \mu^+, \\ x_{ij} - 1, & (i, j) \in \mu^-, \\ x_{ij}. & \end{cases}$$

end while

```

ALGORITHM 1: Algorithm for GAP.

*Proof.* Since  $\theta = \min_{k \in \mathbf{S} \cap \mathbf{M}, q \in \bar{\mathbf{S}} \cap \mathbf{N}} c_{kq} - p_k + p_q$ , at least an admissible will be introduced after updating node potentials in the following iteration. Thus,  $p_i - p_j = c_{i,j'}$ , and  $c_{i,j'} = \min c_{ij}, \forall i \in \mathbf{M}, j \in \mathbf{N}$ , which implies that nodes  $s, i, j'$  can be labeled. In addition, as  $c_{j't} = 0$ , then sink node  $t$  is to be labeled. Therefore, we can find an augmenting path from  $s$  to  $t$ , and there exists at least a node  $j' \in \mathbf{N}$  such that  $x_{j't} = 1$ . Furthermore, the equation  $x_{j't} = 1$  is always satisfied in the following iterations according to the details in Algorithm 1. Thus, job  $j'$  is to be processed by an agent in  $\mathbf{M}$  for optimal match. Similarly, for an optimal solution, there exists an agent  $i' \in \mathbf{M}$  to be assigned to a job in  $\mathbf{N}$ .  $\square$

From Theorem 4, we have the following corollary.

**Corollary 1.** *Let  $x_{ij}^*, i \in \mathbf{M}, j \in \mathbf{N}$  be an optimal solution for the GAP with  $k-1$  cardinality, then the agents and jobs corresponding to  $x_{ij}^* = 1$  will be assigned in GAP with  $k$  cardinality.*

**Theorem 5.** *Algorithm 1 will terminate in finite iterations.*

*Proof.* Clearly, sink node  $t$  cannot be labeled if  $x_{jt} = 1, \forall j \in \mathbf{N}$ . Otherwise, as  $c_{si} = c_{jt} = 0, \forall i \in \mathbf{M}, j \in \mathbf{N}$ ,  $t$  will be labeled while  $j$  is labeled, which implies that we can find an augmenting path from  $s$  to  $t$  and one more unit flow can be sent from  $s$  to  $t$  at first iteration. In addition, node  $i$  can be labeled if  $j$  is labeled and  $x_{ij} = 1$ . Therefore, we can find an augmenting path after updating node potential in  $l$  iterations at  $l$ -th iteration, and at least one more unit flow can be sent from  $s$  to  $t$ . For GAP with  $k$  cardinality, the total number of iterations is  $1 + 2 + \dots + k = ((k^2 + k)/2)$ , which gives Theorem 5.  $\square$

As it is described in Algorithm 1, flow can be augmented after updating nodes' potential in a finite iteration.

Therefore, the complexity of the proposed algorithm is  $O(k^2(n+m)^2)$ .

Next, we will give some numerical experiments to evaluate the efficiency of Algorithm 1.

## 4. Numerical Experiments

In this section, the performance of the proposed algorithm is demonstrated by numerical experiments and we compare its performance with a PP-Lapjv algorithm [31] and bees algorithm for GAP. The bees algorithm is proposed by Ozbakir et al. [23] for the complex integer optimization problem. The PP-Lapjv algorithm is an improved version of Lapjv algorithm [32, 33]. The GAP is transformed into a balanced AP when we use the PP-Lapjv algorithm for solving GAP. For the sake of simplicity, the PP-Lapjv algorithm is called  $A_1$ . The bees algorithm and our algorithm presented in the paper are called  $A_2$  and  $A_3$ , respectively. The presented algorithm is coded in MATLAB<sup>R</sup> R2016b on a laptop computer with an Intel<sup>TM</sup>Corei5-7300HQ CPU @2.50 GHz RAM 8 GB and the operating system is Windows 10 64 bit. As shown in Figure 1, the unit capacity network flow model of the GAP has a special structure that there is only forward  $(i, j), \forall i \in \mathbf{M}, j \in \mathbf{N}$ . Thus, it is easy to generate a random network which is connected.

A series of numerical experiments is implemented on test problems differing in input data values of  $c_{ij}$ , and the  $c_{ij}$  is drawn randomly from a uniform distribution on the intervals  $[0, 10^2], [0, 10^5]$ . This kind of generation is often encountered to evaluate algorithms for solving AP [34]. The GAP with different scales  $((m, n)$  varying from  $(20, 20)$  to  $(200, 400)$ ) is considered. And each class consists of instances with  $k$  valued  $0.2v, 0.4v, 0.6v, 0.8v, 0.9v, v$ , where  $v = \min\{m, n\}$ . The column titled  $k$  provides the coefficient of  $v$ . Also, for each class of problem, we consider the case that an agent can be assigned to  $k$  jobs when  $k$  is given. In addition, the proposed algorithm is evaluated on various

TABLE 1: Computational behavior of randomly generated GAP with various values of  $m, n, k, c$ , and a density valued 0.8 (times in seconds).

$(m, n)$	$k$	$d\% = 80\%$						$d\% = 20\%$					
		$c \in [0, 10^2]$			$c \in [0, 10^5]$			$c \in [0, 10^2]$			$c \in [0, 10^5]$		
		$A_1$	$A_2$	$A_3$	$A_1$	$A_2$	$A_3$	$A_1$	$A_2$	$A_3$	$A_1$	$A_2$	$A_3$
(20, 20)	0.2	5.620	0.046	0.040	5.940	0.047	0.014	4.971	0.040	0.012	5.837	0.043	0.006
	0.4	5.683	0.048	0.019	6.168	0.048	0.021	5.008	0.042	0.017	5.906	0.045	0.016
	0.6	5.926	0.052	0.052	6.186	0.053	0.035	5.255	0.045	0.021	6.183	0.050	0.018
	0.8	5.911	0.059	0.041	6.003	0.067	0.030	5.216	0.053	0.022	5.985	0.062	0.019
	0.9	6.098	0.060	0.050	6.064	0.072	0.040	5.343	0.053	0.029	6.090	0.066	0.024
	1.0	6.122	0.061	0.042	6.109	0.077	0.040	5.379	0.054	0.043	6.040	0.071	0.027
	Average	5.893	0.054	0.041	6.078	0.061	0.028	5.195	0.048	0.024	6.142	0.056	0.018
(20, 40)	0.2	5.855	0.108	0.040	6.187	0.111	0.025	5.168	0.095	0.011	6.007	0.108	0.015
	0.4	5.906	0.109	0.030	5.971	0.126	0.014	5.173	0.097	0.018	5.955	0.115	0.016
	0.6	6.101	0.112	0.070	6.128	0.128	0.032	5.382	0.099	0.023	5.910	0.119	0.020
	0.8	6.201	0.127	0.053	5.986	0.136	0.021	5.449	0.113	0.026	5.922	0.139	0.025
	0.9	6.440	0.135	0.063	6.030	0.139	0.045	5.705	0.123	0.030	5.917	0.145	0.025
	1.0	6.361	0.139	0.047	5.947	0.143	0.041	5.644	0.123	0.031	6.000	0.151	0.027
	Average	6.144	0.122	0.051	6.042	0.131	0.030	5.423	0.108	0.023	5.952	0.130	0.021
(50, 50)	0.2	5.951	0.551	0.178	7.246	0.667	0.083	5.247	0.567	0.050	6.977	0.850	0.045
	0.4	6.071	0.554	0.263	7.322	0.681	0.123	5.394	0.586	0.097	7.274	0.867	0.064
	0.6	6.114	0.565	0.331	7.385	0.692	0.152	5.418	0.590	0.125	7.230	0.885	0.119
	0.8	6.935	0.582	0.358	7.411	0.696	0.194	6.098	0.603	0.127	7.411	0.886	0.120
	0.9	7.311	0.584	0.358	7.272	0.701	0.206	6.417	0.618	0.210	7.232	0.952	0.128
	1.0	8.068	0.607	0.384	7.398	0.732	0.214	7.074	0.651	0.242	7.187	0.984	0.111
	Average	6.742	0.574	0.312	7.339	0.695	0.162	5.941	0.603	0.142	7.222	0.899	0.098
(50, 100)	0.2	6.737	0.846	0.344	8.624	1.105	0.064	5.896	0.746	0.080	8.305	1.082	0.042
	0.4	6.956	0.852	0.544	8.714	1.112	0.142	6.155	0.751	0.127	8.529	1.114	0.089
	0.6	6.872	0.920	0.686	8.893	1.124	0.194	6.026	0.817	0.169	8.669	1.207	0.113
	0.8	8.446	0.959	0.723	8.989	1.137	0.207	7.427	0.846	0.191	8.773	1.239	0.138
	0.9	9.252	1.051	0.803	9.057	1.146	0.214	8.153	0.925	0.234	8.667	1.304	0.203
	1.0	11.576	1.112	0.884	9.051	1.368	0.227	10.136	0.977	0.239	8.961	1.348	0.213
	Average	8.307	0.957	0.561	8.888	1.165	0.175	7.314	0.884	0.173	8.651	1.216	0.133
(100, 100)	0.2	6.760	2.264	0.782	8.940	3.038	0.168	5.944	1.988	0.322	8.604	2.843	0.183
	0.4	6.849	2.368	2.697	10.983	3.398	0.479	6.051	2.107	0.517	10.566	3.055	0.257
	0.6	6.994	2.413	3.322	11.534	3.440	0.583	6.150	2.148	0.642	11.053	3.136	0.382
	0.8	8.673	2.891	3.545	11.692	4.230	0.630	7.621	2.602	0.663	11.375	3.929	0.465
	0.9	9.286	3.107	3.714	11.943	4.661	0.854	8.181	2.796	0.866	11.744	4.018	0.587
	1.0	11.676	3.268	3.727	11.926	5.033	0.882	10.255	2.983	1.092	11.588	4.237	0.616
	Average	8.373	2.719	2.965	11.170	3.967	0.559	7.367	2.422	0.667	10.822	3.536	0.415
(100, 200)	0.2	9.555	15.103	5.256	11.826	18.724	0.343	8.411	11.700	0.769	11.597	16.087	0.212
	0.4	9.084	15.217	8.981	12.132	20.537	0.715	7.957	12.150	1.225	11.781	16.940	0.424
	0.6	9.357	16.467	11.173	12.238	23.267	0.977	8.231	13.530	1.470	11.881	19.305	0.599
	0.8	12.586	18.525	11.642	15.406	25.901	1.003	11.050	15.725	1.495	15.023	22.801	0.550
	0.9	12.845	19.637	11.963	15.714	27.588	1.099	11.283	17.052	1.758	15.302	25.578	0.744
	1.0	17.236	20.208	12.195	20.010	31.029	1.155	15.115	18.187	1.775	19.212	27.281	0.932
	Average	11.777	17.511	10.202	14.554	24.508	0.882	10.341	4.724	1.145	14.193	21.212	0.577
(200, 200)	0.2	9.670	22.612	34.092	12.301	28.250	1.826	8.498	18.316	3.673	12.087	24.907	0.847
	0.4	9.782	23.107	59.843	12.213	29.217	3.158	8.571	19.173	6.142	11.859	26.842	1.707
	0.6	9.825	24.750	70.225	12.693	34.423	4.208	8.676	21.285	7.608	12.318	30.225	2.446
	0.8	13.742	27.695	73.291	16.275	37.131	5.208	12.045	23.364	7.748	15.784	32.478	2.434
	0.9	14.158	28.733	74.686	16.692	42.157	5.658	12.420	24.952	9.039	16.286	33.167	4.569
	1.0	18.843	31.203	76.188	21.245	45.208	7.306	16.530	27.007	10.678	20.643	35.247	4.576
	Average	12.670	26.350	64.723	15.237	36.064	4.561	11.123	22.350	7.481	14.830	30.478	2.763
(200, 400)	0.2	9.529	108.531	127.211	12.686	145.807	3.993	8.375	93.961	12.338	12.139	133.425	1.628
	0.4	10.327	113.421	218.291	12.719	154.826	7.552	9.055	94.692	20.583	12.450	134.219	3.125
	0.6	10.135	120.280	266.023	12.518	164.904	9.913	8.941	102.241	24.847	12.293	143.627	4.389
	0.8	15.123	133.514	275.431	18.226	172.927	11.495	13.271	105.821	25.050	17.707	151.292	4.306
	0.9	16.541	139.770	285.422	18.824	180.619	11.870	14.486	122.327	27.784	18.184	156.443	5.852
	1.0	23.668	151.022	286.511	26.359	194.231	11.990	20.766	129.863	28.139	25.409	160.126	6.457
	Average	14.221	127.356	243.148	16.889	168.886	9.946	12.482	108.151	23.124	16.364	146.522	4.248

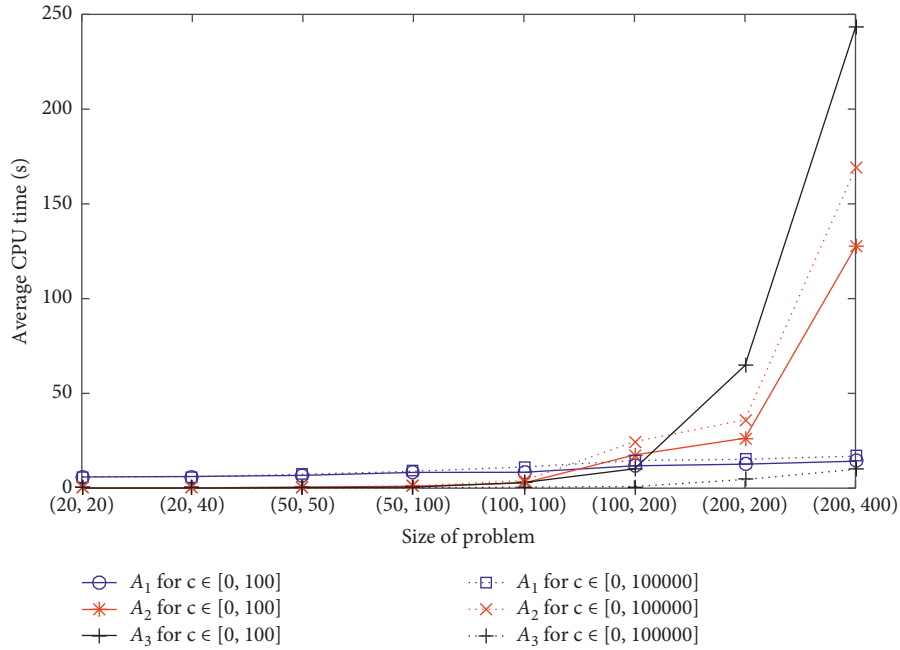


FIGURE 2: Average CPU time of algorithms  $A_1$ ,  $A_2$ , and  $A_3$  for different sizes of problem with density 80%.

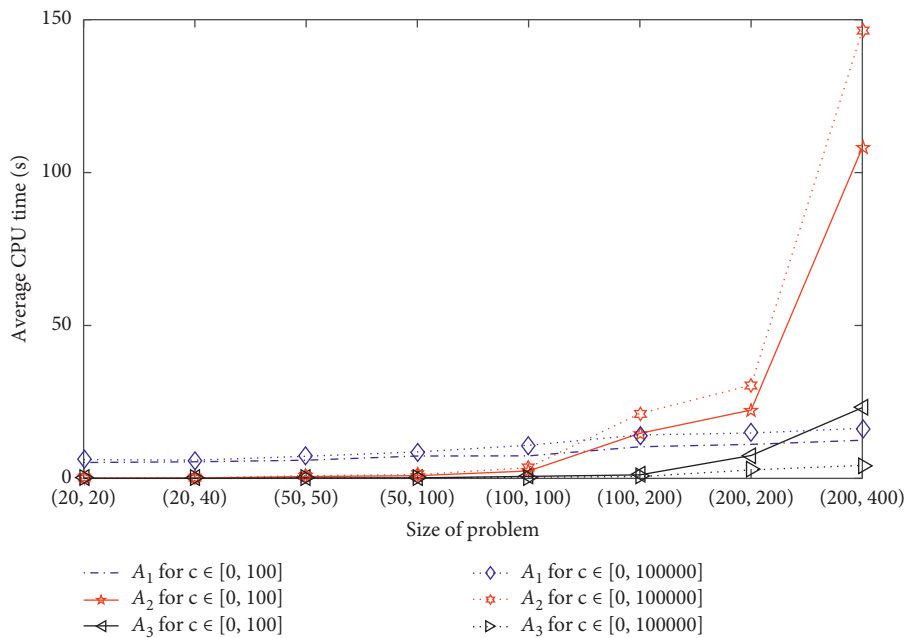


FIGURE 3: Average CPU time of algorithms  $A_1$ ,  $A_2$ , and  $A_3$  for different sizes of problem with density 20%.

degrees of density 20% and 80%. For a given density  $d\%$ , the instance is generated as follows. We first randomly generate a value  $r \in [0, 1]$ , and for each arc  $(i, j), \forall i \in M, j \in N$ , the arc  $(i, j)$  is kept in the instance with a randomly generated cost  $c_{ij}$ ; otherwise, the next arc is considered. Higher density means that each multitask agent  $i \in M$  can be assigned to more various jobs in  $M$ . Further, the parameter setting is the same as the parameters in [23].

We report the CPU time of  $A_1$ ,  $A_2$ , and  $A_3$  for solving GAP in Table 1. In each block, each entry in the 6 lines gives

the average CPU time computed over 10 instances, neglecting the input, the output, and generation of the network flow model. And the last line gives the average over all  $k$  values considered. Also, Figures 2 and 3 depict the average CPU time on the various  $k$  values for the density 80% and 20%, respectively.

$A_3$  has a high probability to find an augmenting path from  $s$  to  $t$  after updating nodes' potential in two iterations, which can be easily verified by the network model and Algorithm 1 of GAP. Furthermore, as it is shown in



Algorithm 1,  $A_3$  holds the information of node potential at each iteration. Therefore, for a larger range  $c_{ij} \in [0, 10^5]$ , the curves of the proposed algorithm are at the bottom of Figures 2 and 3. However, it will take more iterations for finding an augmenting path from  $s$  to  $t$  with a larger number of arcs  $(i, j) \in A', i \in M, j \in N$  and a smaller range  $c_{ij} \in [1, 10^2], (i, j) \in A'$  for  $A_3$ . Thus, there are exceptions for the efficiency of  $A_3$  when the size of GAP is increased (e.g.,  $(m, n)$  is greater than  $(100, 100)$  for a smaller  $c_{ij}$ ). It is understandable that  $A_3$  will find an augmenting path in much more iteration in a network with a smaller  $c_{ij}$ , especially for a larger size of problem. In addition, larger size of GAP results in a greater number of local optimal solutions. Thus, it will take much more CPU time for  $A_2$ , which is shown in Figures 2 and 3.

The CPU time of algorithm  $A_1$  consists of time for running Lapjv.m file and constructing the  $k$ -th transferred cost matrix which is mainly related to the size of the problem. Therefore, variation of CPU time of  $A_1$  is much smaller than that of  $A_3$  for the same size of problem (the size of the problem is defined by  $\max\{m, n\}$ ).

## 5. Conclusion and Further Work

This paper develops an algorithm for solving GAP in which an integer  $k$  is given and one wants to assign  $k'$  ( $k' \leq k$ ) agents to  $k$  jobs such that the sum of the corresponding cost is minimal. The problem is transformed into a network model with a special structure. And the network model of GAP can easily process the GAP in which one agent can be assigned to more than one job or a job group can be processed by more than one agent. Some important properties of GAP are derived and numerical experiments show that our proposed algorithm has some good performance for solving GAP with a larger range cost  $c_{ij}$  and a smaller number of arcs  $(i, j), \forall i \in M, j \in N$ .

In this study, the scale of GAP is not very large, and we plan to develop some techniques for processing large-scale sparse cost matrix in order to optimize the GAP in a short time. Furthermore, it is fundamental to develop an algorithm for solving the bottleneck GAP or fuzzy GAP.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

This work was supported by the Hubei Superior and Distinctive Discipline Group of Mechatronics and Automobile (no. XKQ2020004), Innovation Scientists and Technicians Troop Construction Projects of Henan Province (CXTD2016054), and Innovative Scientists and Technicians

Team of Henan Provincial High Education (no. 20IRTSTHN019).

## References

- [1] A. Haghani and S.-C. Oh, "Formulation and solution of a multi-commodity, multi-modal network flow model for disaster relief operations," *Transportation Research Part A: Policy and Practice*, vol. 30, no. 3, pp. 231–250, 1996.
- [2] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows*, Springer, Berlin, Germany, 1988.
- [3] J. Kennington and Z. Wang, "A shortest augmenting path algorithm for the semi-assignment problem," *Operations Research*, vol. 40, no. 1, pp. 178–187, 1992.
- [4] M. Dell'Amico and S. Martello, "The  $k$ -cardinality assignment problem," *Discrete Applied Mathematics*, vol. 76, no. 1–3, pp. 103–121, 1997.
- [5] C. Prins, "An overview of scheduling problems arising in satellite communications," *The Journal of the Operational Research Society*, vol. 45, no. 6, pp. 611–623, 1994.
- [6] E. Balas and P. R. Landweer, "Traffic assignment in communication satellites," *Operations Research Letters*, vol. 2, no. 4, pp. 141–147, 1983.
- [7] M. Dell'Amico, A. Lodi, and S. Martello, "Efficient algorithms and codes for  $k$ -cardinality assignment problems," *Discrete Applied Mathematics*, vol. 110, no. 1, pp. 25–40, 2001.
- [8] A. Volgenant, "Solving the  $k$ -cardinality assignment problem by transformation," *European Journal of Operational Research*, vol. 157, no. 2, pp. 322–331, 2004.
- [9] I. Belik, "The analysis of split graphs in social networks based on the  $k$ -cardinality assignment problem," *International Journal of Network Science*, vol. 1, no. 1, pp. 53–62, 2016.
- [10] M. L. Fisher and R. Jaikumar, "A generalized assignment heuristic for vehicle routing," *Networks*, vol. 11, no. 2, pp. 109–124, 1981.
- [11] P. C. Chu and J. E. Beasley, "A genetic algorithm for the generalised assignment problem," *Computers & Operations Research*, vol. 24, no. 1, pp. 17–23, 1997.
- [12] E. Munapo, M. Lesaoana, P. Nyamugure, and S. Kumar, "A transportation branch and bound algorithm for solving the generalized assignment problem," *International Journal of System Assurance Engineering and Management*, vol. 6, no. 3, pp. 217–223, 2015.
- [13] P. Avella, M. Boccia, and I. Vasilyev, "A computational study of exact knapsack separation for the generalized assignment problem," *Computational Optimization and Applications*, vol. 45, no. 3, pp. 543–555, 2010.
- [14] M. L. Fisher, R. Jaikumar, and L. N. Van Wassenhove, "A multiplier adjustment method for the generalized assignment problem," *Management Science*, vol. 32, no. 9, pp. 1095–1103, 1986.
- [15] R. M. Naus, "Solving the generalized assignment problem: an optimizing and heuristic approach," *Informatics Journal on Computing*, vol. 15, no. 3, pp. 249–266, 2003.
- [16] S. Martin, "A branch-and-price algorithm for the generalized assignment problem," *Operations Research*, vol. 45, no. 6, pp. 831–841, 1997.
- [17] A. J. Woodcock and J. M. Wilson, "A hybrid tabu search/branch & bound approach to solving the generalized assignment problem," *European Journal of Operational Research*, vol. 207, no. 2, pp. 566–578, 2010.
- [18] J. A. Diaz and E. Fernandez, "A tabu search heuristic for the generalized assignment problem," *European Journal of Operational Research*, vol. 132, no. 1, pp. 22–38, 2001.

- [19] I. H. Osman, "Heuristics for the generalised assignment problem: simulated annealing and tabu search approaches," *Operations-Research-Spektrum*, vol. 17, no. 4, pp. 211–225, 1995.
- [20] J. Majumdar and A. K. Bhunia, "An alternative approach for unbalanced assignment problem via genetic algorithm," *Applied Mathematics and Computation*, vol. 218, no. 12, pp. 6934–6941, 2012.
- [21] M. Fatih Tasgetiren, P. N. Suganthan, T. J. Chua, and A. Al-Hajri, "Differential evolution algorithms for the generalized assignment problem," in *Proceedings of the 2009 IEEE Congress on Evolutionary Computation*, pp. 2606–2613, IEEE, Trondheim, Norway, May 2009.
- [22] K. Sethanan and R. Pitakaso, "Improved differential evolution algorithms for solving generalized assignment problem," *Expert Systems with Applications*, vol. 45, pp. 450–459, 2016.
- [23] L. Özbakir, A. Baykasoglu, and P. Tapkan, "Bees algorithm for generalized assignment problem," *Applied Mathematics and Computation*, vol. 215, no. 11, pp. 3782–3795, 2010.
- [24] S. K. Dubey, A. Kumar, and V. Upadhyay, "The average sum method for the unbalanced assignment problems," *International Journal of Mathematics Trends and Technology*, vol. 55, no. 2, pp. 89–100, 2018.
- [25] Q. Rabbani, A. Khan, and A. Quddoos, "Modified Hungarian method for unbalanced assignment problem with multiple jobs," *Applied Mathematics and Computation*, vol. 361, pp. 493–498, 2019.
- [26] A. Khandelwal, "An amalgamated approach for solving unbalanced assignment problem," *Malaya Journal of Matematik*, vol. 6, no. 2, pp. 321–325, 2018.
- [27] M. Posta, J. A. Ferland, and P. Michelon, "An exact method with variable fixing for solving the generalized assignment problem," *Computational Optimization and Applications*, vol. 52, no. 3, pp. 629–644, 2012.
- [28] H. Zhu, D. Liu, S. Zhang, Yu Zhu, L. Teng, and S. Teng, "Solving the many to many assignment problem by improving the Kuhn-Munkres algorithm with backtracking," *Theoretical Computer Science*, vol. 618, pp. 30–41, 2016.
- [29] Y. Hu, X. Zhao, J. Liu, B. Liang, and C. Ma, "An efficient algorithm for solving minimum cost flow problem with complementarity slack conditions," *Mathematical Problems in Engineering*, vol. 2020, Article ID 2439265, 2020.
- [30] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms and Applications*, Prentice-Hall, Upper Saddle River, NJ, USA, 1993.
- [31] W. Boonphakdee and P. Charnsethikul, "Solving large scale assignment problem using the successive complementary slackness conditions," in *Proceedings of the 2nd International Conference on Mathematics, Engineering and Industrial Applications 2016*, Songkhla, Thailand, August 2016.
- [32] R. Jonker and A. Volgenant, "A shortest augmenting path algorithm for dense and sparse linear assignment problems," *Computing*, vol. 38, no. 4, pp. 325–340, 1987.
- [33] R. Jonker and A. Volgenant, "Linear assignment procedures," *European Journal of Operational Research*, vol. 116, no. 1, pp. 233–234, 1999.
- [34] D. Knuth, "The art of computer programming," in *Semimerical Algorithms* vol. 2, 2nd edition, 1981.

## Research Article

# Three-Dimensional Numerical Calculation Model for Static Behavior Simulation of Cross-Laminated Timber Plates under Thermal Environment

Wenliang Hu <sup>1</sup>, Wei Hou,<sup>1</sup> Zhao Zhu <sup>1</sup>, and Xuhui Huang<sup>2</sup>

<sup>1</sup>School of Highway, Chang'an University, Xi'an 710064, China

<sup>2</sup>School of Civil Engineering, Anhui Jianzhu University, Anhui 230601, China

Correspondence should be addressed to Wenliang Hu; 2015021022@chd.edu.cn

Received 1 August 2020; Revised 1 October 2020; Accepted 26 December 2020; Published 12 January 2021

Academic Editor: Yuqiang Xu

Copyright © 2021 Wenliang Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cross-laminated timber (CLT) is well known as an interesting technical and economical product for modern wood structures. The use of CLT for modern construction industry has become increasingly popular in particular for residential timber buildings. Analyzing the CLT behavior in high thermal environment has attracted scholars' attention. Thermal environment greatly influences the CLT properties and load bearing capacity of CLT, and the investigation can form the basis for predicting the structural response of such CLT-based structures. In the present work, the finite element method (FEM) is employed to analyze the thermal influence on the deformation of CLT. Furthermore, several factors were taken into consideration, including board layer number, hole conformation, and hole position, respectively. In order to determine the influence, several numerical models for different calculation were established. The calculation process was validated by comparing with published data. The performance is quantified by demonstrating the temperature distribution and structural deformation.

## 1. Introduction

Wood possesses excellent mechanical and physical properties which is well known as an environmentally sustainable building material with a pleasant appearance and, thus, is particularly suitable for residential buildings, public buildings, and industrial buildings [1, 2]. In recent years, timber is undergoing a revival and recapturing market shares from other building materials, and its world market was witnessed a considerable growth in the total production [3, 4].

The revival of timber is largely beneficial from the development of the innovative laminar timber product cross-laminated timber (CLT). This timber product commonly composed of an uncertain number of layers, each made of boards placed side-by-side, which were arranged crosswise to each other at an angle of 90° [5]. Comparing with the traditional wood, CLT possesses a higher structural capacity, which makes it suitable for walls, floor slabs, and roofs, and

is also a wood material which is suitable for large-scale and high-rise building constructions [6]. CLT can be constructed into large solid timber panels, and this allows the transfer of high vertical loads and guarantees a high building stiffness and robustness. Another advantage of CLT is an excellent thermal insulation. There is a majority of researches investigating the fire behavior of CLT panels. Eurocode gives a simplified method for the calculation of the mechanical resistance of timber beams and columns (reduced cross-sectional model) [7]. Frangi et al. studied the fire resistance of CLT panels located out of the plane [8]. Klippel et al. investigated the fire safety of CLT, and the fire design has been done [9, 10]. Schmid et al. has modeled and tested the fire-exposed CLT, and they also proposed results from fire tests of beam strips cut from CLT with adequate side protection in bending in order to achieve 1D heat transfer [11–13]. Frangi et al. presented the experimental and numerical analyses of the fire behavior of CLT panels, and they

further compared the fire behavior of CLT panels with homogeneous timber panels [14, 15]. Experiment was performed, and 3D finite element model was developed to predict the fire behavior of loaded CLT panels [16]. The performance of CLT and hybrid structure was presented [17, 18]. From numerous literature surveys, we can see that the experimental test was expensive and random.

As the fire tests of CLT are crucial to ensure the safety standards in the timber buildings, the investigation on the CLT behavior in high thermal environment has drawn our attention. The experimental fire tests were costly, and it is of great importance to develop numerical models that can be used to extend the few experimental results available to different configurations and different load conditions. Researchers also presented the homogenization method to build the CLT parameters, and results were compared with the experiments to reveal the potential predictive capabilities of the multiscale modeling for the analysis of wood material [3, 19]. Researchers made efforts to propose numerical models, and they were implemented in existing software [20–22]. The finite element method (FEM) was commonly utilized to study the architectural engineering issues [23–25]. Takeda and Mehaffey have proposed a 2D computer model for predicting heat transfer through uninsulated wood-stud walls [26]. Franssen et al. employed finite element method to test the heat transfer of 3D beam [27]. A high-order FEM for vibration analysis of CLT structures and the suitable ranges for the stiffness parameters of the finite elements which represent the junctions were identified [28]. Michele Betti et al. proposed a new test method to assess the quality of face bonding in CLT [29]. The present literatures were focused on analysis of the structural influence on the CLT performance. And the literatures reveal that the FEM was the most widely used method. As a result, we choose to use FEM to analyze the behavior of the structure with change in temperature.

As is known from all other timber products with crosswise layering, CLT possesses a high dimensional stability in-plane. But, along thickness direction, swelling and shrinking have to be considered equal to solid timber. This work presented in this paper is concerned with numerical simulations of CLT plates. We made the first attempt to fill the gap in illustrating the static behavior of CLT plates subject to thermal loads. The finite element method was employed, and the influence of board layer number, hole dimension, hole conformation, and hole position on the CLT structural performance was taken into account. The influence was quantified by demonstrating the temperature distribution and structural deformation of CLT structure. Several numerical examples were established to investigate these influences.

## 2. Finite Element Method

The geometric equations of a CLT structure are basically listed below:

$$\begin{aligned}\varepsilon_x &= -E_x^{-1}(\sigma_x - \mu_{xz}\sigma_z - \mu_{xy}\sigma_y), \\ \varepsilon_y &= -E_y^{-1}(-\mu_{yx}\sigma_x - \mu_{yz}\sigma_z + \sigma_y), \\ \varepsilon_z &= -E_z^{-1}(-\mu_{zx}\sigma_x + \sigma_z - \mu_{zy}\sigma_y), \\ \gamma_{yz} &= \frac{\tau_{yz}}{G_{yz}}, \\ \gamma_{zx} &= \frac{\tau_{zx}}{G_{zx}}, \\ \gamma_{xy} &= \frac{\tau_{xy}}{G_{xy}},\end{aligned}\quad (1)$$

where  $\varepsilon_x$ ,  $\varepsilon_y$ , and  $\varepsilon_z$  represent the strain components in the  $x$ ,  $y$ , and  $z$  directions;  $E_x$ ,  $E_y$ , and  $E_z$  represent the elastic modulus in the  $x$ ,  $y$ , and  $z$  directions;  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  represent the stress components in the  $x$ ,  $y$ , and  $z$  directions;  $\mu_{xz}$ ,  $\mu_{yz}$ , and  $\mu_{xy}$  represent Poisson's ratio and are expressed as  $\mu_{xz} = \varepsilon_z/\varepsilon_x$ ,  $\mu_{yz} = \varepsilon_z/\varepsilon_y$ , and  $\mu_{xy} = \varepsilon_y/\varepsilon_x$ ;  $\gamma_{yz}$ ,  $\tau_{zx}$ , and  $G_{yx}$  represent the shear strain, shear stress, and shear modulus, respectively.

The heat transfer for the CLT plate is assembled by the partial differential equation:

$$-D\nabla T = Q, \quad (2)$$

where  $T$  represents the temperature,  $D$  represents the thermal conductivity parameter matrix, and  $Q$  is heat flux matrix:

$$\nabla T = \begin{Bmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \\ \frac{\partial T}{\partial z} \end{Bmatrix}, \quad (3)$$

$$D = \begin{bmatrix} k_{xx}(T) & 0 & 0 \\ 0 & k_{yy}(T) & 0 \\ 0 & 0 & k_{zz}(T) \end{bmatrix},$$

where  $k_{xx}(T)$ ,  $k_{yy}(T)$ , and  $k_{zz}(T)$  represent the thermal conductivities in the  $x$ ,  $y$ , and  $z$  directions.

The thermal expansion coefficient of wood parallel was ranging from about 3.1 to  $4.5 \times 10^{-6} \text{K}^{-1}$ . And the radial thermal expansion  $\alpha_y$  and tangential thermal expansion  $\alpha_z$  are the coefficients approximated by the following equations, where  $G_0$  is the specific gravity:

$$\begin{aligned}\alpha_y &= (32.4G_0 + 9.9) \times 10^{-6} \text{K}^{-1}, \\ \alpha_z &= (32.4G_0 + 18.4) \times 10^{-6} \text{K}^{-1}.\end{aligned}\quad (4)$$



### 3. The FEM for Analyzing CLT under High-Temperature Environment

The motion equations for the thermal structural coupling analysis can be written as

$$\begin{bmatrix} \mathbf{K}_{uu} & \mathbf{K}_{ut} \\ 0 & \mathbf{K}_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{T} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{Q} \end{bmatrix},$$

$$\begin{aligned} \mathbf{K}_{uu} &= \int_{\Omega} \bar{\mathbf{B}}_u^T [\mathbf{C}] \bar{\mathbf{B}}_u d\Omega, \\ \mathbf{K}_{ut} &= \int_{\Omega} \bar{\mathbf{B}}_u^T [\mathbf{C}] [\alpha] \bar{\mathbf{B}}_t d\Omega, \\ \mathbf{K}_{tt} &= \int_{\Omega} \bar{\mathbf{B}}_t^T [\boldsymbol{\varepsilon}] \bar{\mathbf{B}}_t d\Omega, \\ \mathbf{F} &= \int_S \mathbf{N}^T \mathbf{F} dS, \\ \mathbf{Q} &= \int_S \mathbf{N}^T \mathbf{Q} dS, \end{aligned} \quad (5)$$

where  $\mathbf{K}$ ,  $\mathbf{u}$ , and  $\mathbf{T}$  represent the stiffness, displacement, and temperature matrix;  $\mathbf{F}$  and  $\mathbf{Q}$  represent the mechanical and thermal load matrix;  $\mathbf{C}$ ,  $\boldsymbol{\varepsilon}$ , and  $\alpha$  denote the elastic, conductivity, and thermal expansion coefficients matrix.

The FEM in calculating the CLT in thermal environment was verified with a comparison with research [30]. A five layer CLT plate was analyzed, and the thickness direction temperature distribution at point A ( $x = 218.75$ ,  $y = 737.5$ ) was plotted (Figure 1).

The FEM result matched well with the reference result. The relative error is around 5%, and the maximum relative error does not exceed 15%.

### 4. Numerical Examples

In this section, we established several numerical examples to investigate the static characteristics of CLT. The numerical examples are all on the basis of a rectangle plate, the basic configuration of the CLT plate is shown in Figure 2. The radiata pine was applied for analysis, and the parameters are listed in Table 1. The thermal boundary condition is given in Figure 2, the ambient temperature is  $0^\circ\text{C}$ , and the corner of the plate was fixed ( $u_x = u_y = u_z = 0$ , at points A, B, D, and H). We take into account of the impact of several factors on the performance of the CLT. The board layer number influence, hole conformation influences, and hole position influence were analyzed by demonstrating the temperature distribution in layer at depth positions  $h = 40$  mm, 80 mm, 120 mm, and 160 mm, and the contour plots of plate deformation in  $x$ ,  $y$ , and  $z$  coordinate directions. Furthermore, the  $x$ -,  $y$ -, and  $z$ -direction displacements were also given along the plate borderlines AB and BD, also along the thickness directions BC, FG, and DE.

**4.1. CLT Board Layer Influence.** The correctness of the FEM in calculating the CLT deformation under high-temperature environment was validated previously. Then, we illustrated the board layer effects on the CLT behavior, three kinds of

CLT were analyzed, and they were composed of 3 board layers, 5 board layers, and 7 board layers (Figure 3). In Figure 4, the temperature distribution of layers at depth position  $h = 40$  mm, 80 mm, 120 mm, and 160 mm was demonstrated. In Figure 5, the contour plots of plate deformation in  $x$ ,  $y$ , and  $z$  coordinate directions are given. The displacement along the CLT plate borderlines (AB line  $u_x$ , BD line  $u_y$ , and BD line  $u_z$ ) and the displacement along the thickness direction (BC line  $u_x$ , DE line  $u_y$ , FG line  $u_z$ ) are demonstrated in Figures 6 and 7.

From Figure 4, we can see that 5 board layers have a different temperature distribution form from the other two CLT plates. From Figures 5–7, effect of the board layer on the CLT behavior is shown. For  $u_x$  displacement component, the result increases with the increase in board layer. For  $u_y$  and  $u_z$  displacement components, the result decreases with the increase in board layer increase. Furthermore, the displacement distributions along plate borderlines and thickness reveal that the board layer plays a strong effect on  $u_z$  displacement components than the other two. Hence, the  $u_z$  displacement component was more sensitive to the thermal load, and then the load bearing capacity of CLT plates will be influenced.

**4.2. Hole Conformation Influence.** A square hole was framed in the centre of the 5-board layer rectangular plate. When investigating the hole conformation influence on the CLT static behavior, particular attention should be paid to ensure that the results of the different models remain comparable. And the conformation and dimension coupling influence were further investigated. Five kinds of square hole length dimensions were developed  $l = d = 177.2$  mm,  $l = d = 354.4$  mm,  $l = d = 531.6$  mm,  $l = 502.4$  mm  $d = 250$  mm, and  $l = 420$  mm  $d = 300$  mm (Figure 8). The temperature distribution and the contour plots of plate deformation were given (Figures 9–14).

Figures 9 and 11 demonstrate the temperature distribution in different configurations, from which we can get the same conclusion that the temperature regularly changes along radial direction. In Figures 10 and 12, the square hole dimension effects on the CLT deformation are shown, and it brings a slight effect on the deformation of the plate. In Figures 13 and 14, we can see that the  $u_z$  displacement component is the most sensitive to the hole radius variation.

**4.3. Hole Position Influence.** The hole position also has an influence on the CLT structural performance. When investigating the holes' position influence on the CLT behavior, we constructed four CLT plate models. The concrete structural forms and dimensions are shown in Figure 15. Figures 16–19 demonstrate the temperature distribution and the contour plots of plate deformation.

Figure 16 demonstrate the temperature distribution in models, from which we can see the difference among these models. In Figures 17–19, the hole position effects on the CLT deformation are shown. The hole position brings a strong effect on the deformation of the plate, and the effect was obviously shown in the displacement distribution.

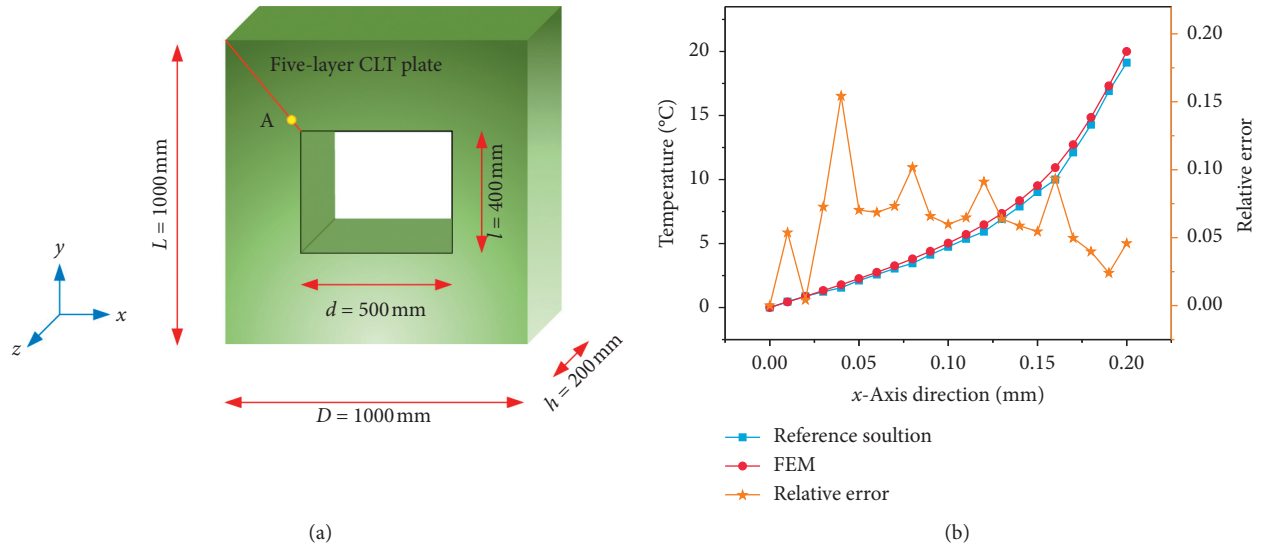


FIGURE 1: (a) Schematic diagram of the CLT plate in [30]. (b) Temperature distribution at point A and the relative error of the FEM.

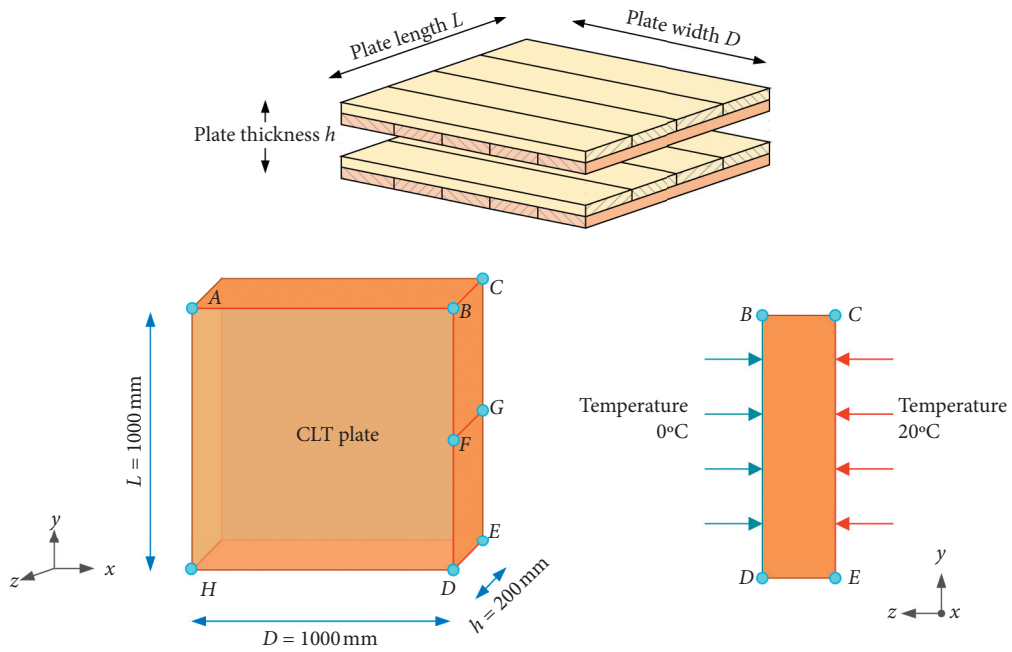


FIGURE 2: Schematic diagram of a CLT plate.

TABLE 1: Material parameters of radiata pine.

Material	Density ( $\text{g/cm}^3$ )	Moisture content (%)		$E_x$ (MPa)	$E_z$ (MPa)	$E_y$ (MPa)	$G_{xy}$ (MPa)	$G_{xz}$ (MPa)	$G_{yz}$ (MPa)
Radiata pine	0.550	10		16272	1103	573	1172	676	66
	$a_x$ ( $\text{K}^{-1}$ ) $4 \times 10^{-6}$	$a_y$ ( $\text{K}^{-1}$ ) $27.7 \times 10^{-6}$	$a_z$ ( $\text{K}^{-1}$ ) $36.2 \times 10^{-6}$	$k_{xx}$ (W/mK) 0.35	$k_{yy}$ (W/mK) 0.16	$k_{zz}$ (W/mK) 0.16	$M_{xy}$ 0.42	$M_{xz}$ 0.51	$M_{yz}$ 0.68

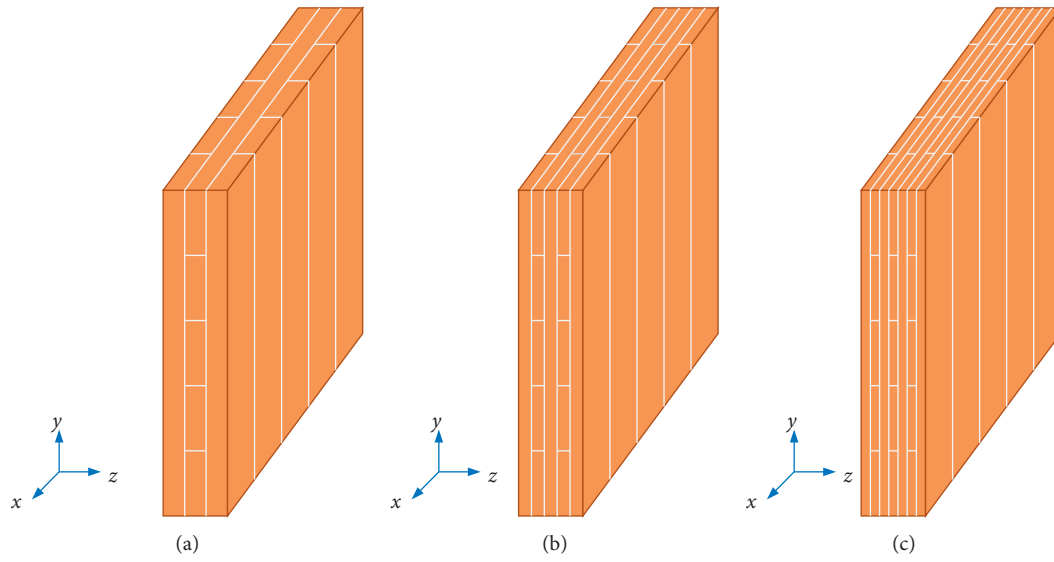


FIGURE 3: The CLT plate with three different board layers. (a) 3-layer CLT plate. (b) 5-layer CLT plate. (c) 7-layer CLT plate.

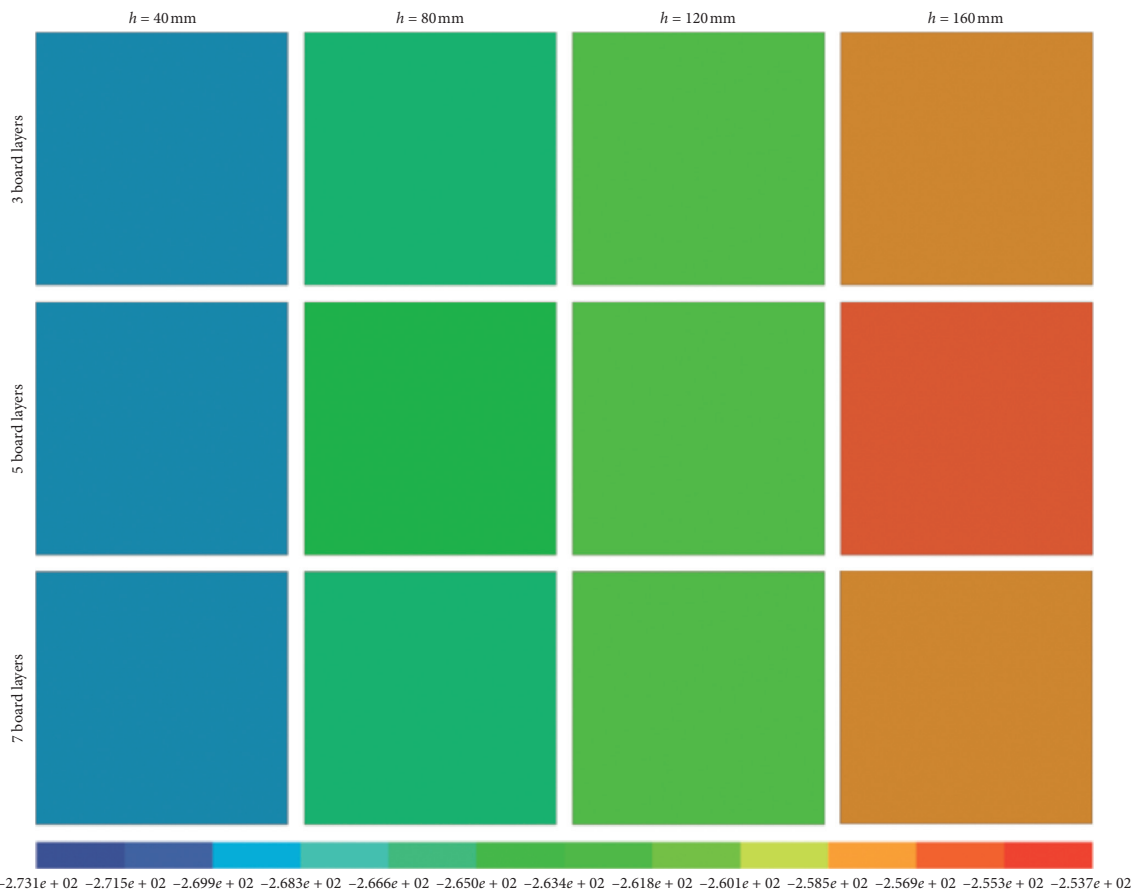
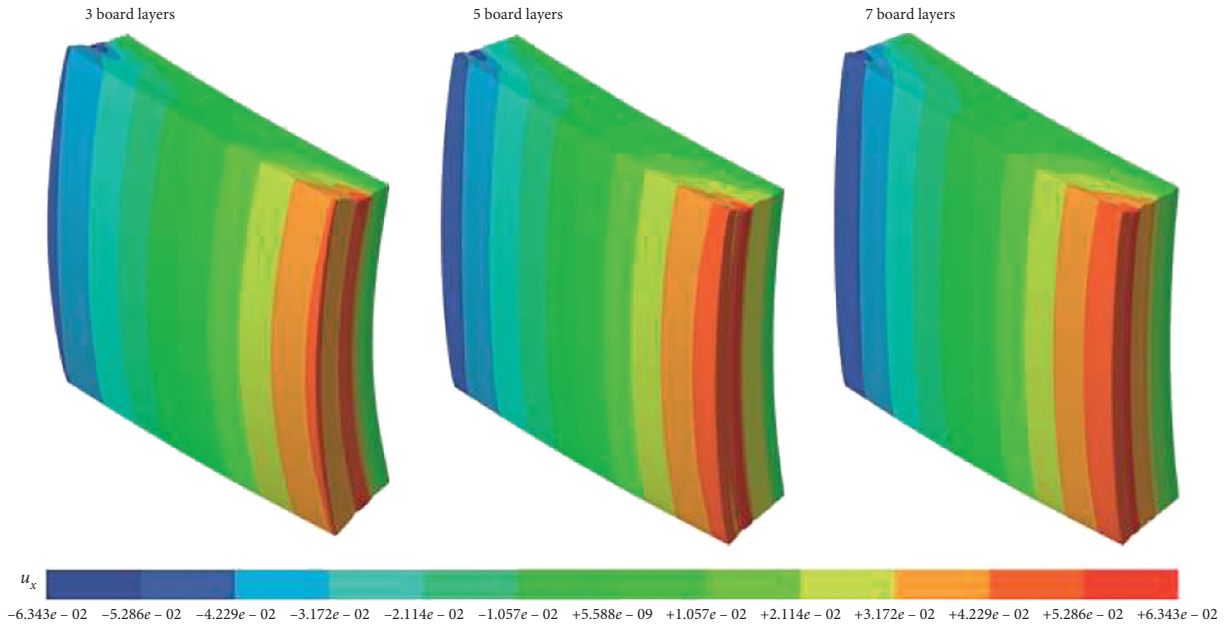
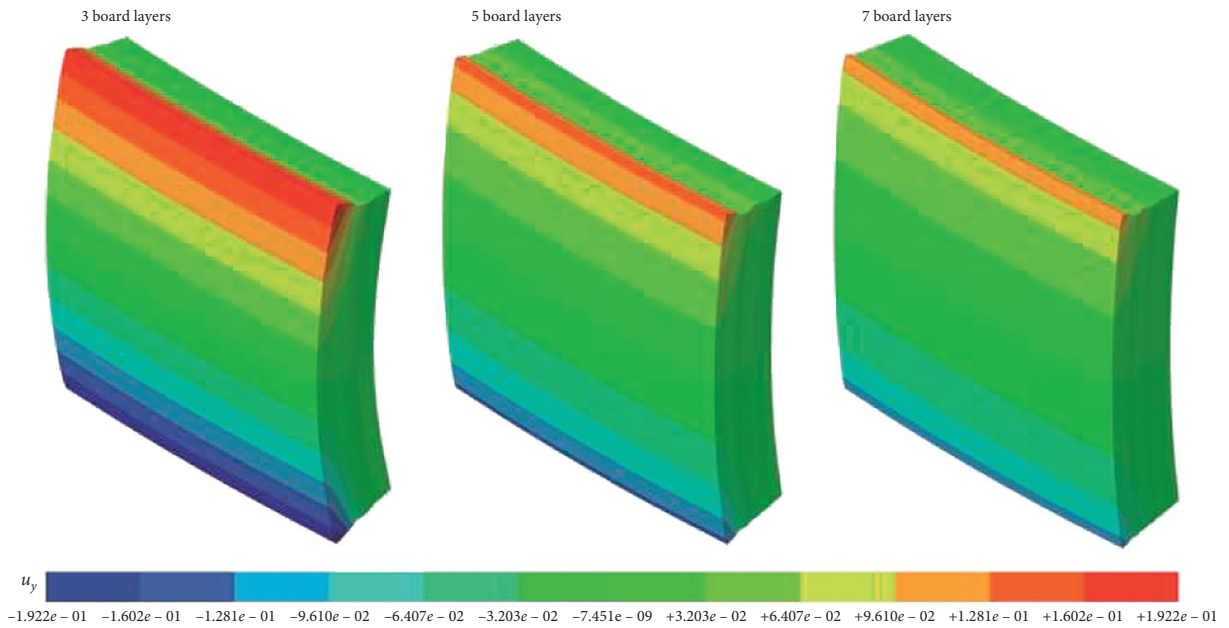


FIGURE 4: The temperature distribution of the CLT plate with various layer components at depth positions  $h = 40\text{ mm}$ ,  $80\text{ mm}$ ,  $120\text{ mm}$ , and  $160\text{ mm}$ .





(a)



(b)

FIGURE 5: Continued.

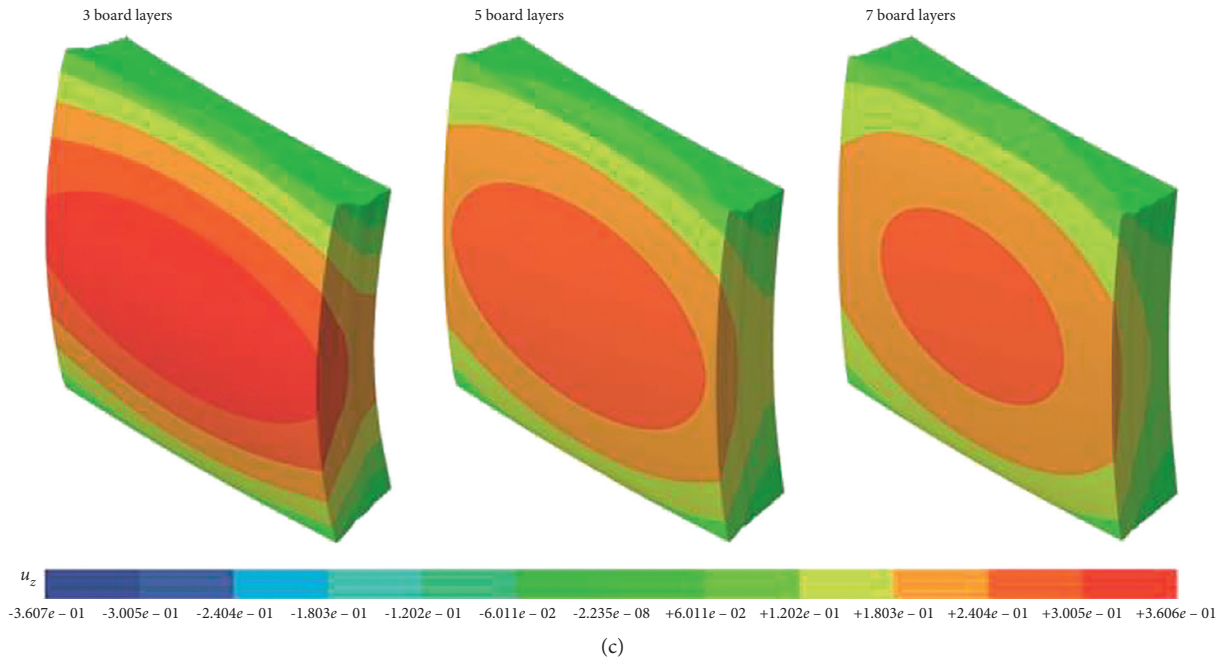


FIGURE 5: The deformations of the CLT plate with various layer components in  $x$ ,  $y$ , and  $z$  coordinate directions. (a)  $u_x$  displacement; (b)  $u_y$  displacement; (c)  $u_z$  displacement.

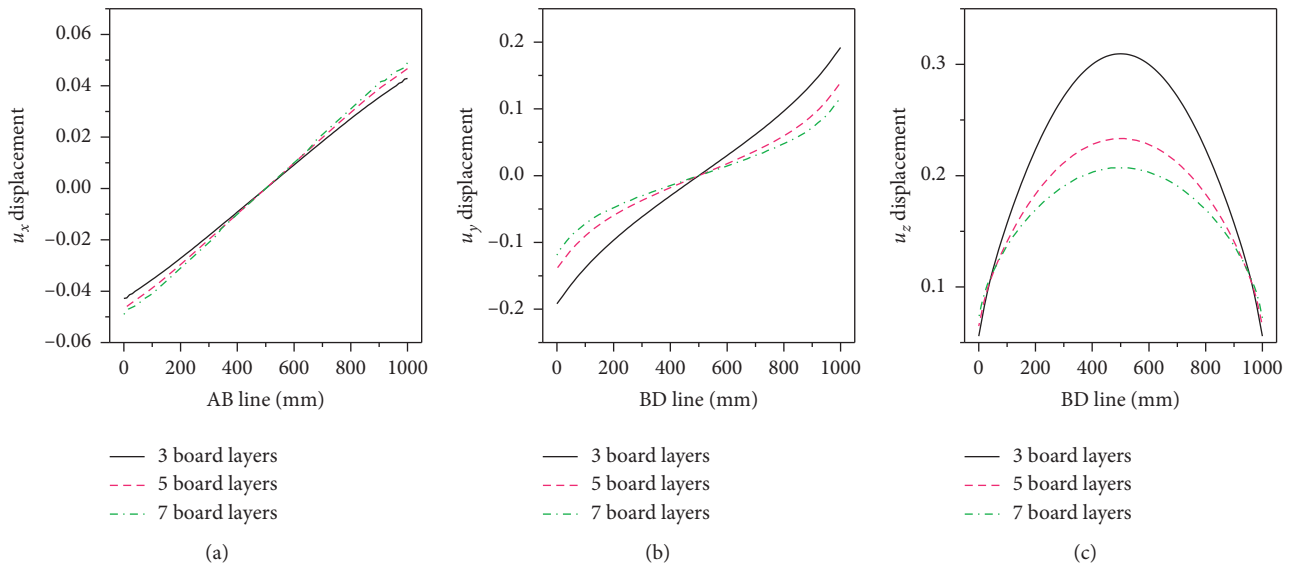


FIGURE 6: The displacement distributions along the CLT plate borderlines. (a)  $u_x$  displacement of  $AB$  line; (b)  $u_y$  displacement of  $BD$  line; (c)  $u_z$  displacement of  $BD$  line.

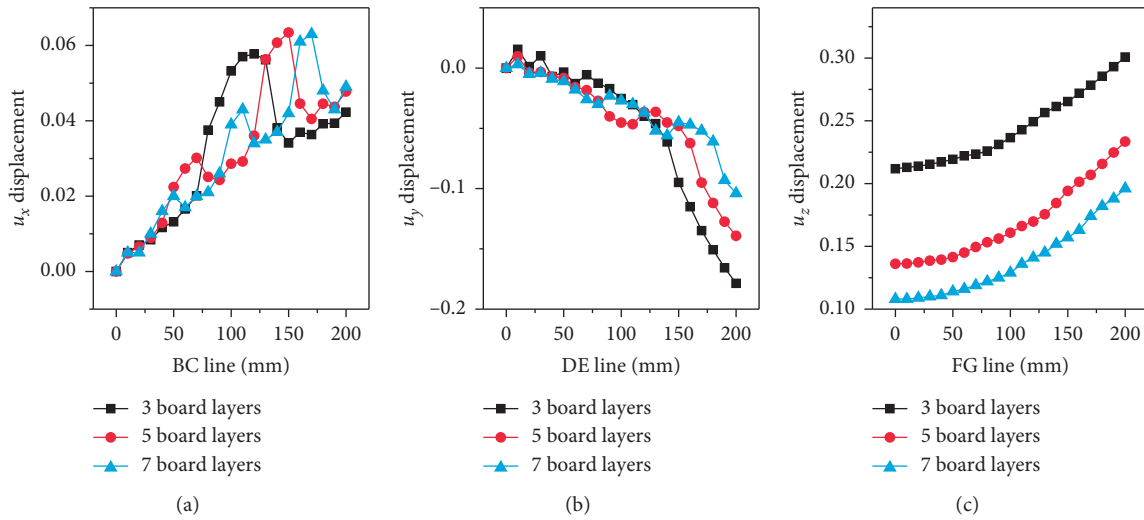


FIGURE 7: The displacement distributions along the CLT plate thickness direction. (a)  $u_x$  displacement of BC line; (b)  $u_y$  displacement of DE line; (c)  $u_z$  displacement of FG line.

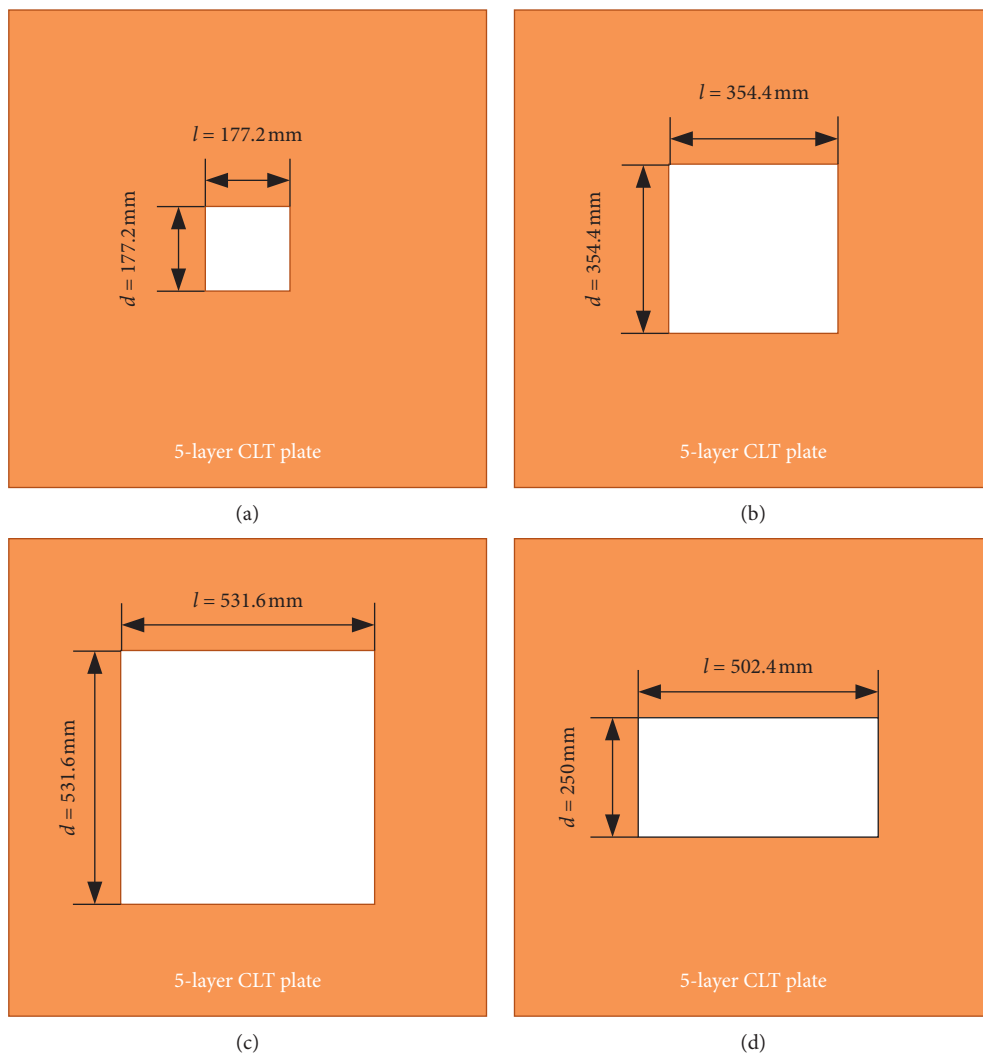
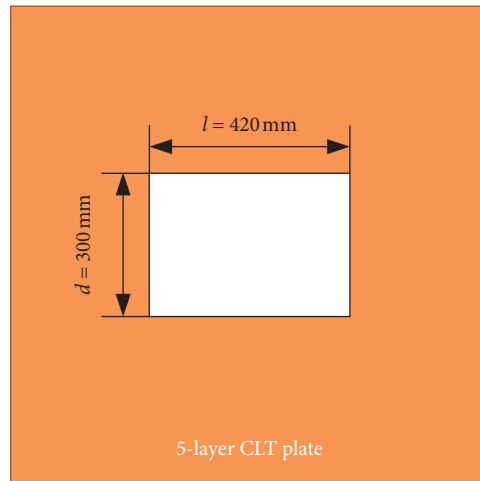


FIGURE 8: Continued.



(e)

FIGURE 8: The CLT plate with five kinds of square hole dimension. (a)  $l = d = 77.2$  mm; (b)  $l = d = 354.4$  mm; (c)  $l = d = 531.6$  mm; (d)  $l = 502.4$  mm,  $d = 250$  mm; (e)  $l = 420$  mm,  $d = 300$  mm.

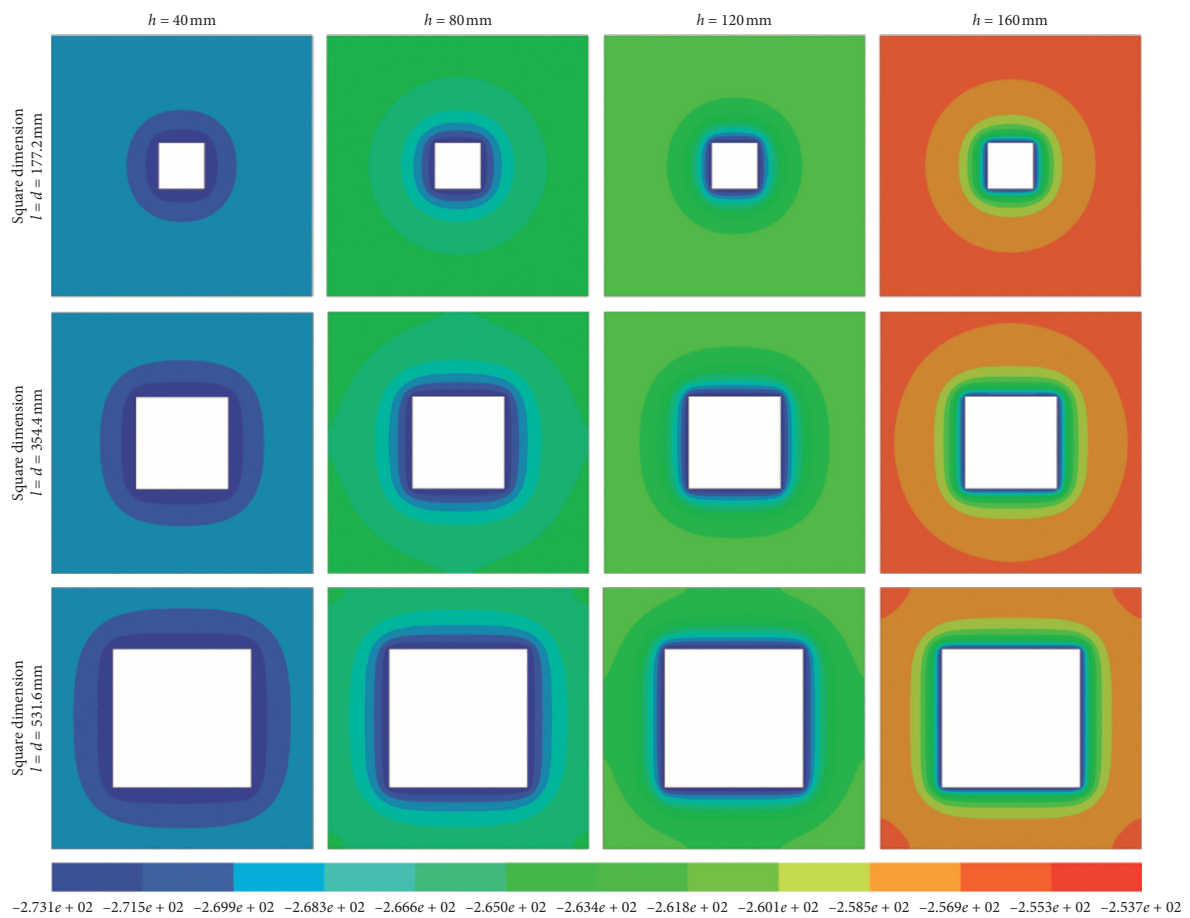
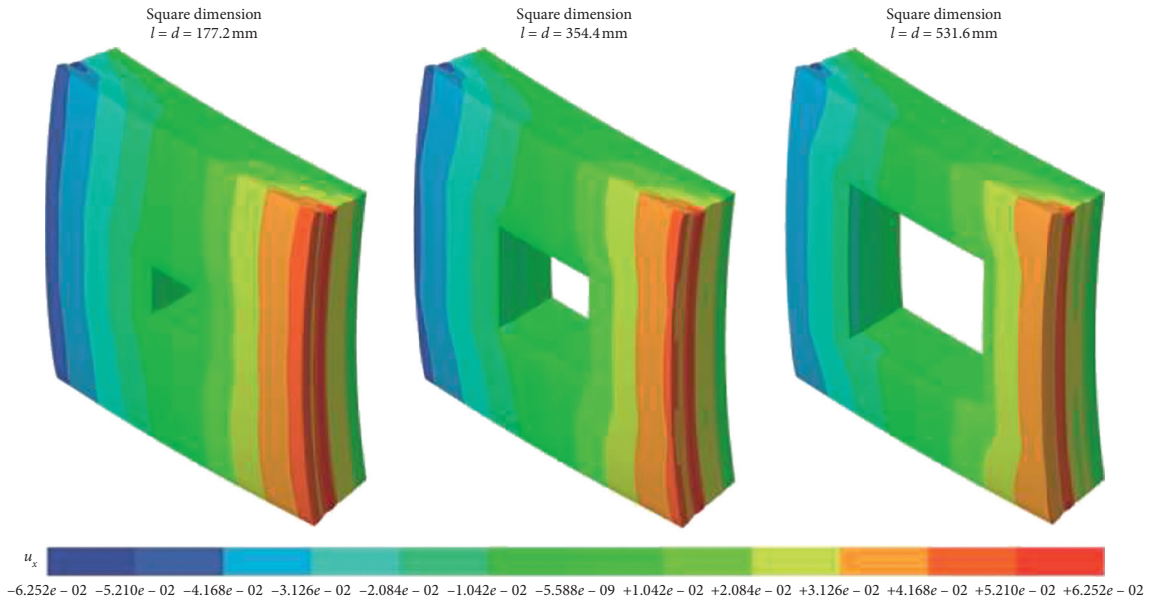
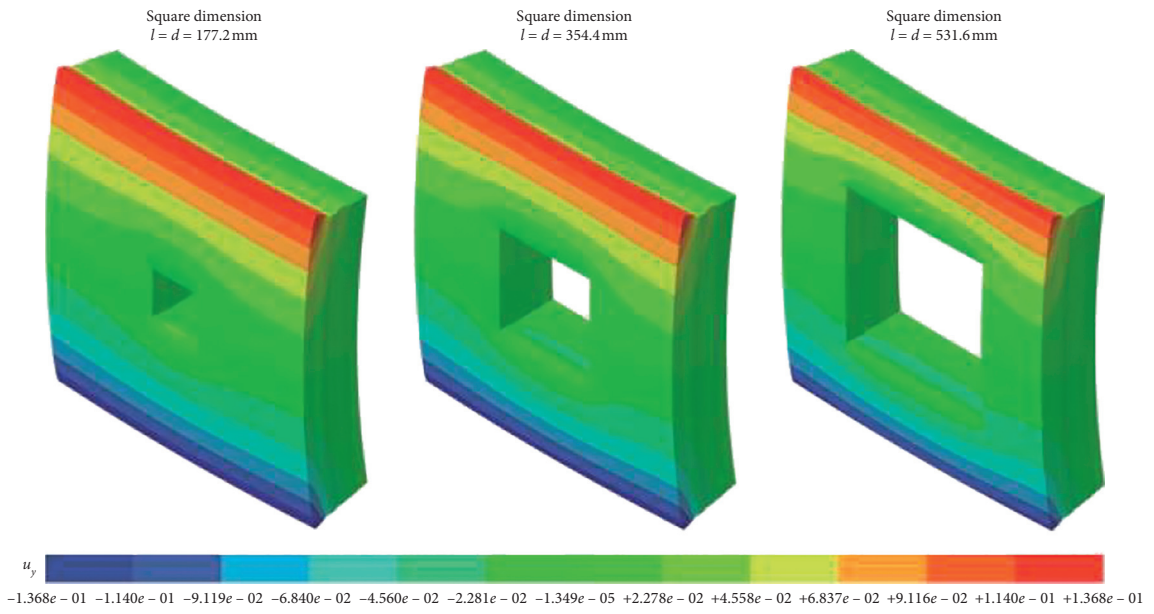


FIGURE 9: The temperature distribution of the CLT plate with various central hole dimension at depth positions  $h = 40$  mm,  $80$  mm,  $120$  mm, and  $160$  mm.



(a)



(b)

FIGURE 10: Continued.

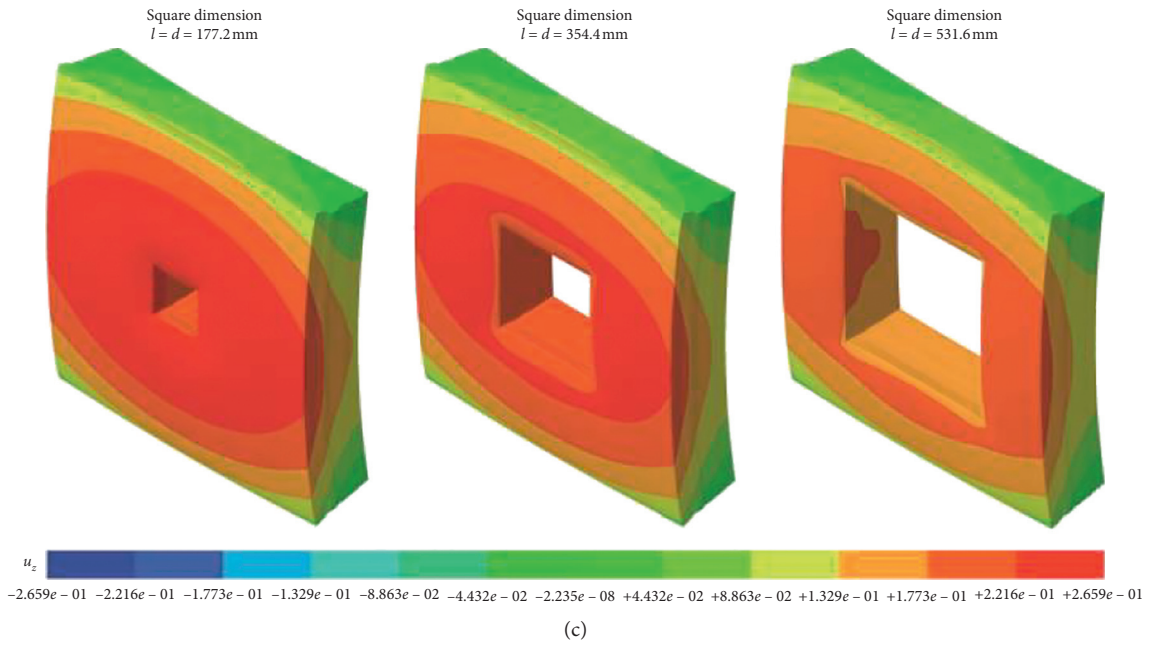


FIGURE 10: The deformations of the CLT plate with various central hole dimensions in  $x$ ,  $y$ , and  $z$  coordinate directions. (a)  $u_x$  displacement; (b)  $u_y$  displacement; (c)  $u_z$  displacement.

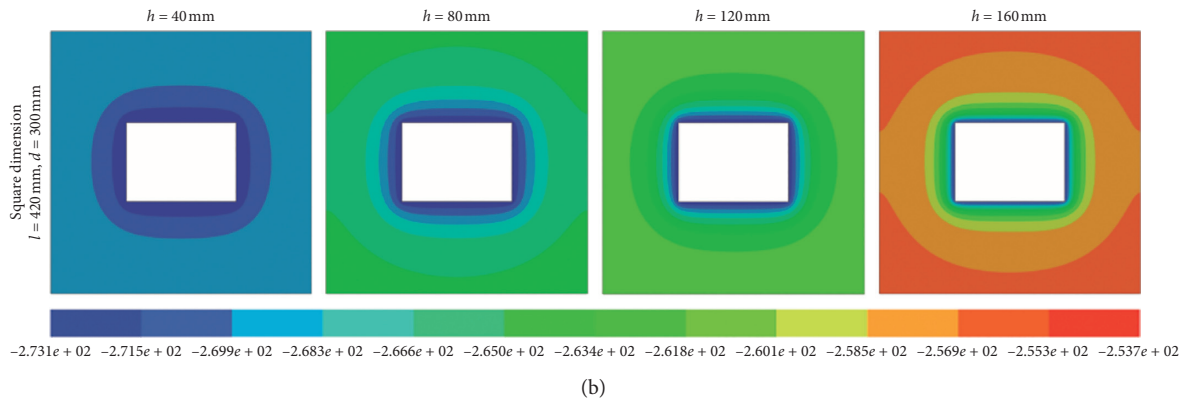
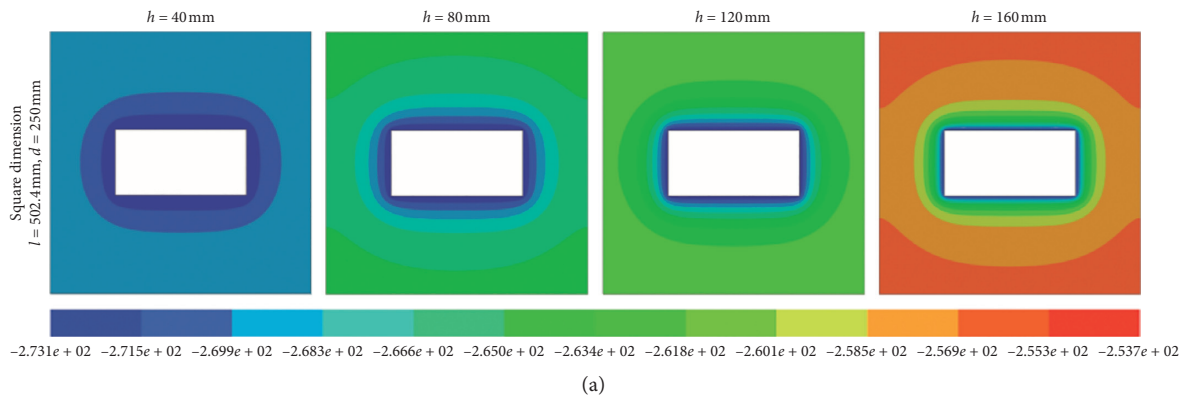


FIGURE 11: The temperature distribution of the CLT plate with two different square hole dimensions at depth positions  $h = 40$  mm,  $80$  mm,  $120$  mm, and  $160$  mm.



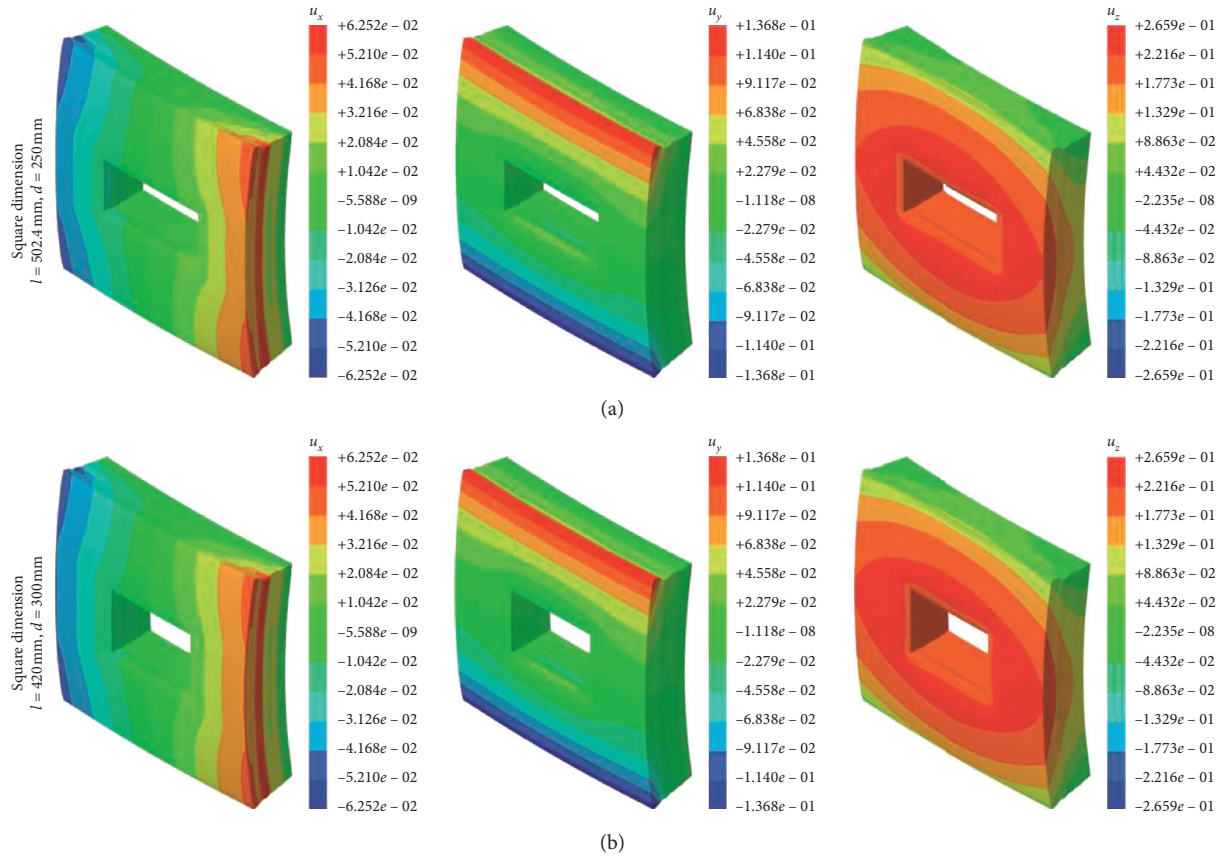


FIGURE 12: The deformations of the CLT plate with two different square hole dimensions in  $x$ ,  $y$ , and  $z$  coordinate directions. (a)  $u_x$  displacement; (b)  $u_y$  displacement; (c)  $u_z$  displacement.

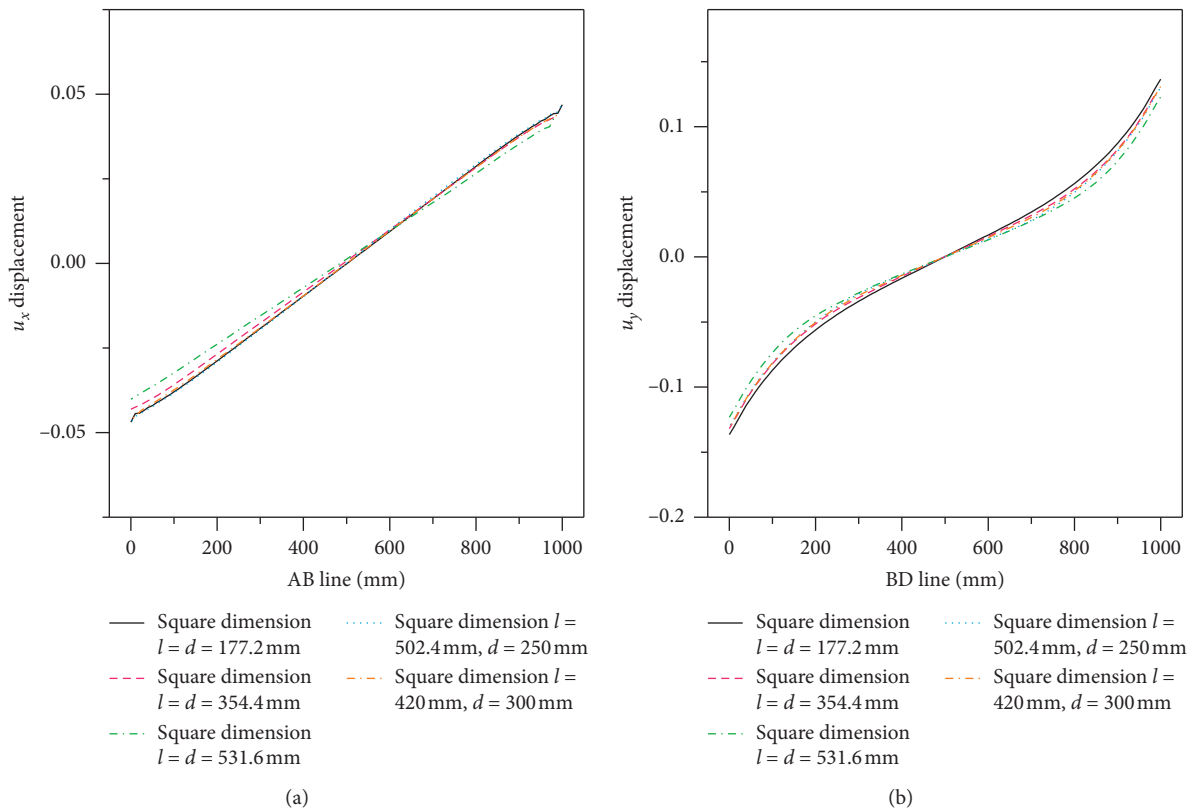


FIGURE 13: Continued.



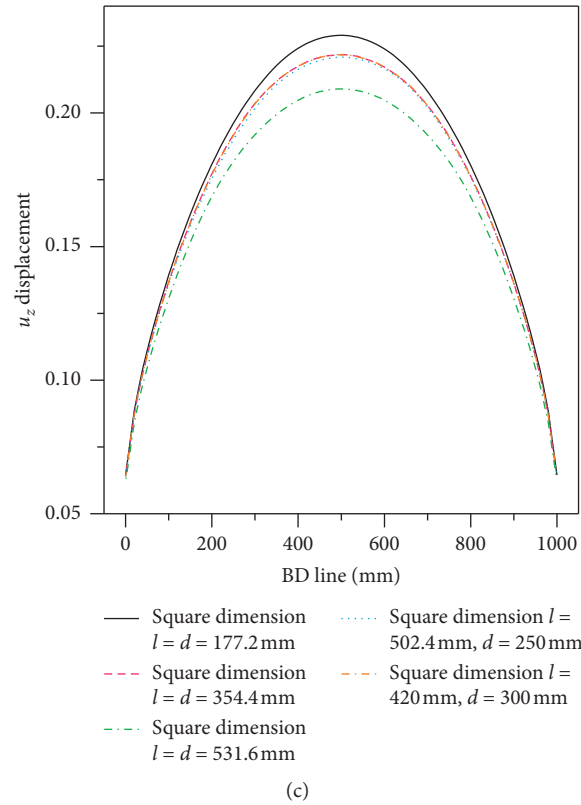


FIGURE 13: The displacement distributions along the CLT plate borderlines. (a)  $u_x$  displacement of  $AB$  line; (b)  $u_y$  displacement of  $BD$  line; (c)  $u_z$  displacement of  $BD$  line.

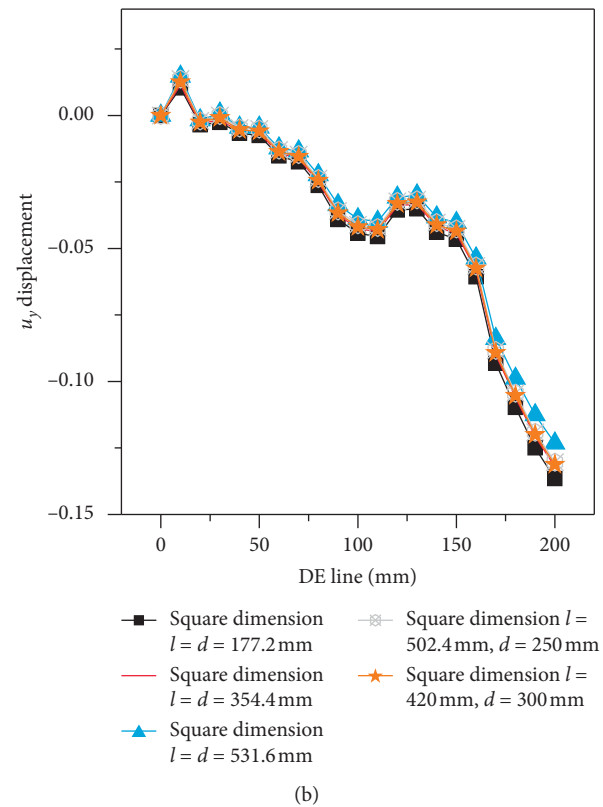
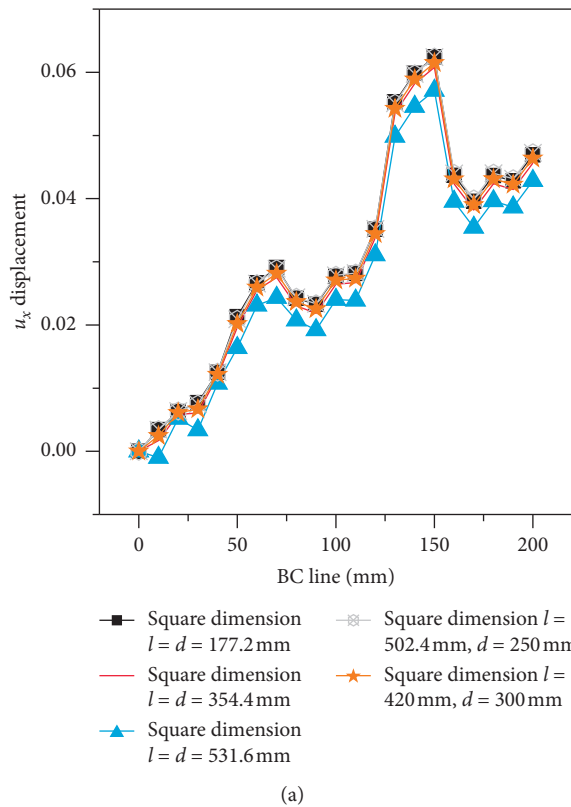


FIGURE 14: Continued.

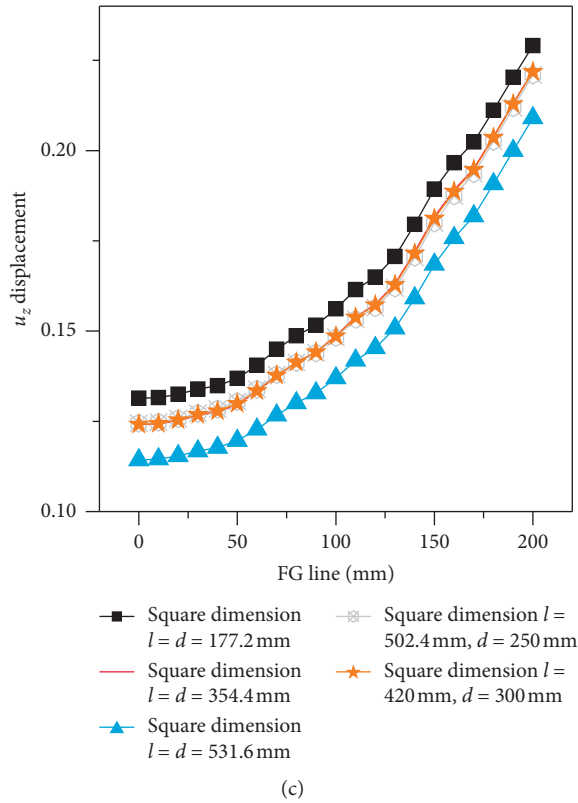


FIGURE 14: The displacement distributions along the CLT plate thickness direction. (a)  $u_x$  displacement of BC line; (b)  $u_y$  displacement of DE line; (c)  $u_z$  displacement of FG line.

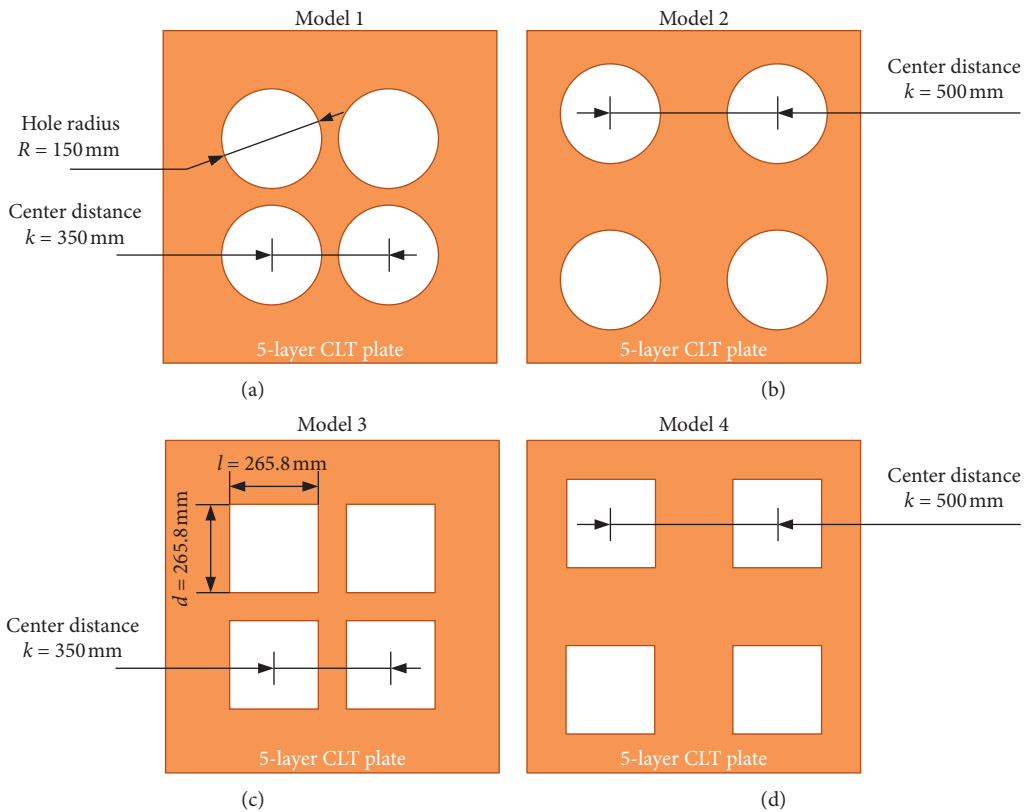


FIGURE 15: The CLT plate in four model compositions, the holes were located in different central distances and forms. (a) Four round holes in central distance  $k = 350$  mm; (b) four round hole in central distance  $k = 500$  mm; (c) four square hole in central distance  $k = 350$  mm; (d) four square hole in central distance  $k = 500$  mm.

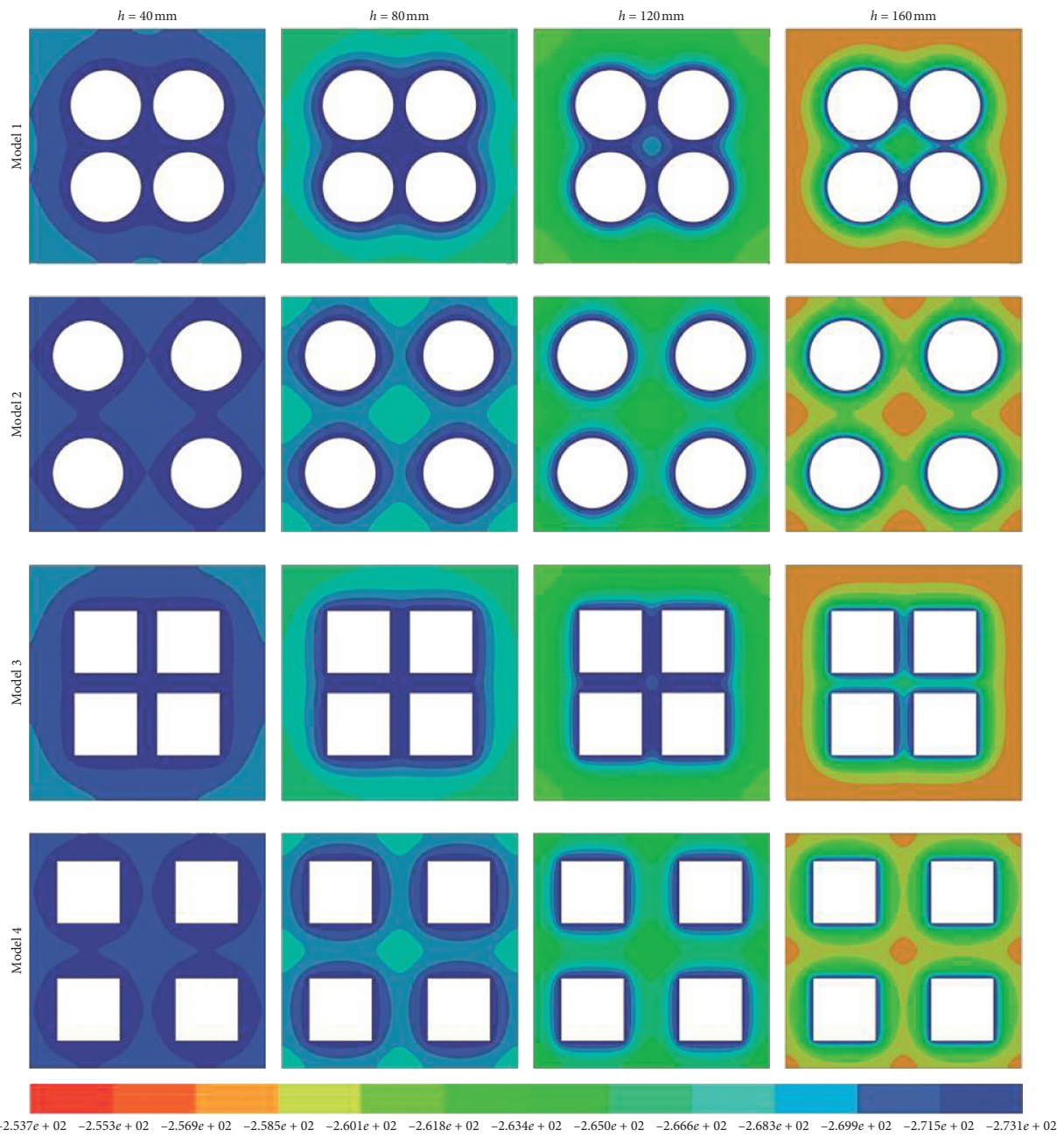
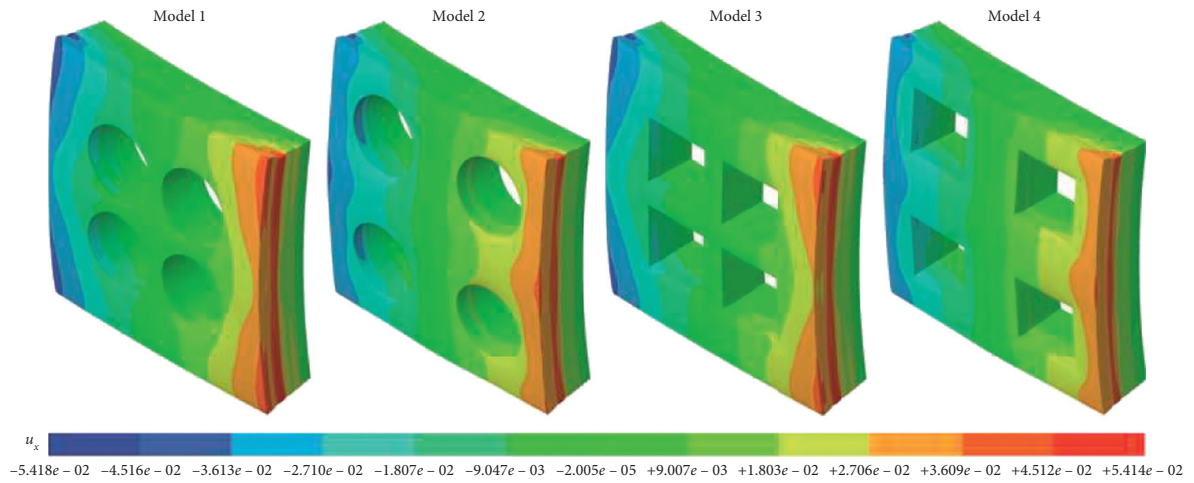
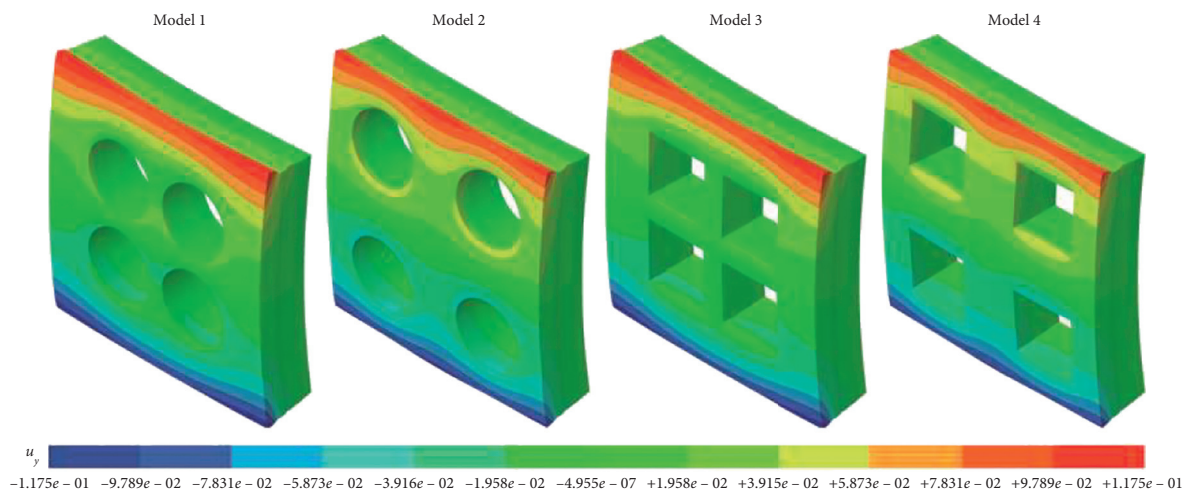


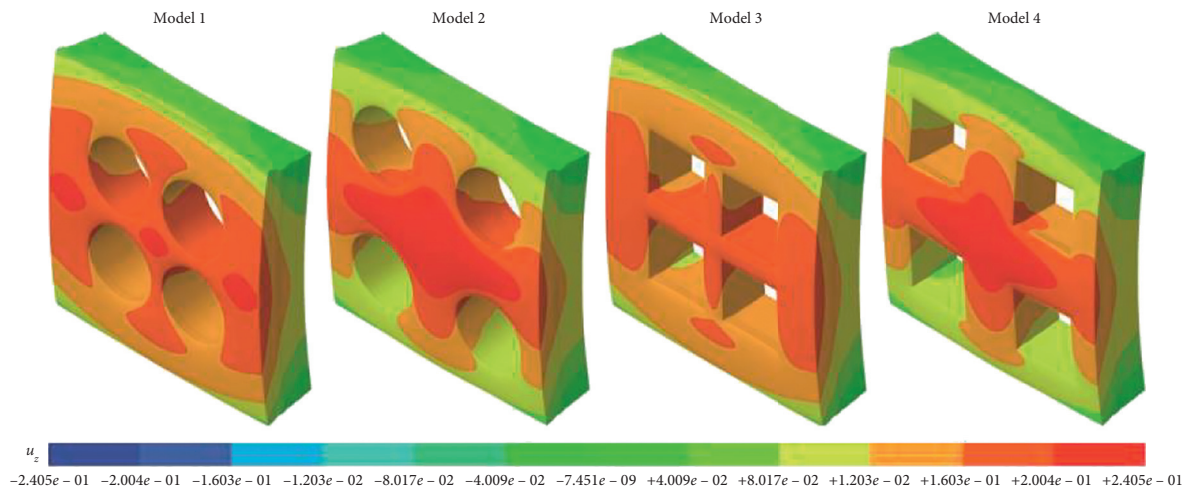
FIGURE 16: The temperature distribution of the CLT plate in four model compositions at depth positions  $h = 40\text{ mm}$ ,  $80\text{ mm}$ ,  $120\text{ mm}$ , and  $160\text{ mm}$ .



(a)



(b)



(c)

FIGURE 17: The deformations of the CLT plate in four model compositions in  $x$ ,  $y$ , and  $z$  coordinate directions. (a)  $u_x$  displacement; (b)  $u_y$  displacement; (c)  $u_z$  displacement.

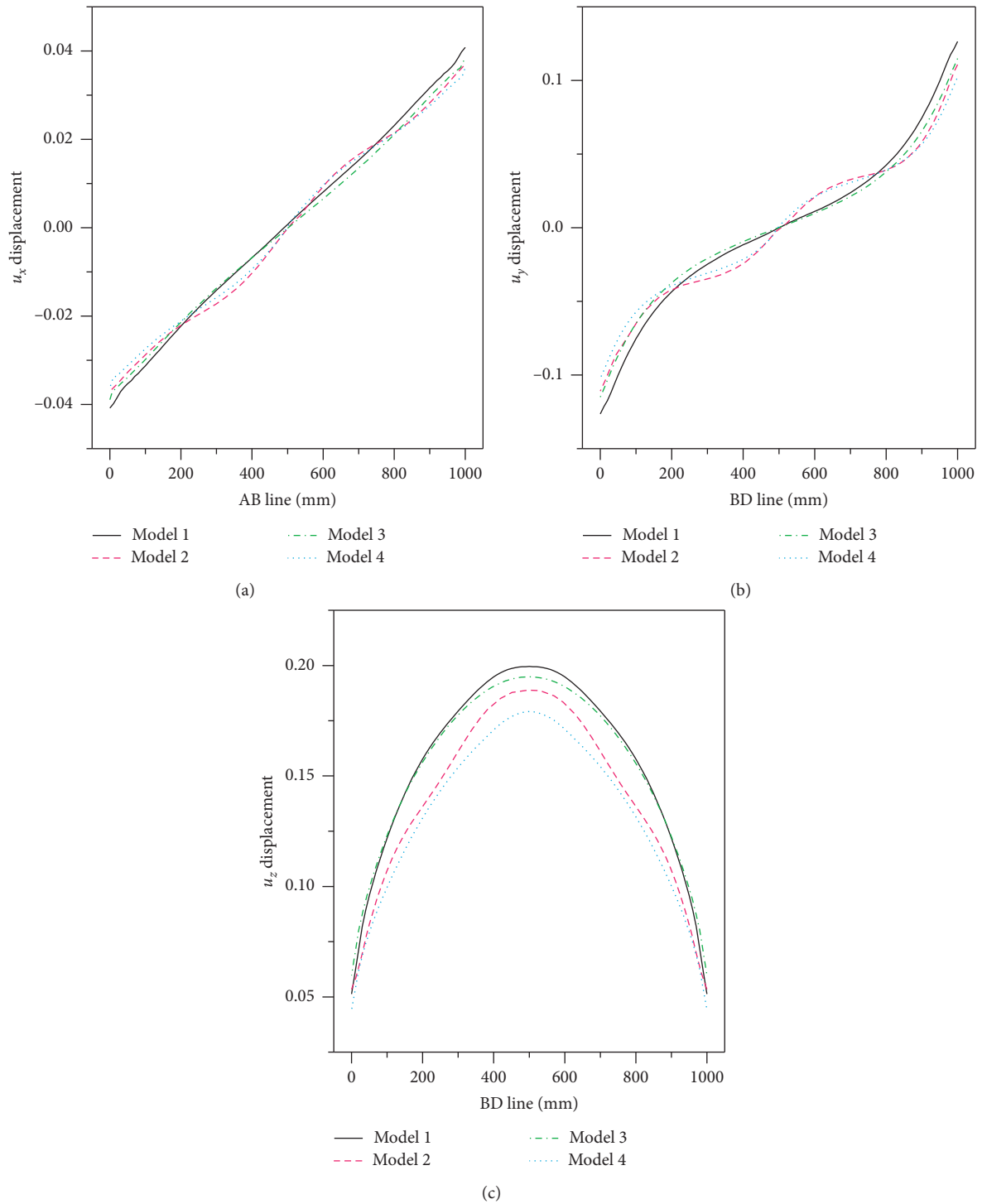


FIGURE 18: The displacement distributions along the CLT plate borderlines. (a)  $u_x$  displacement of AB line; (b)  $u_y$  displacement of BD line; (c)  $u_z$  displacement of BD line.

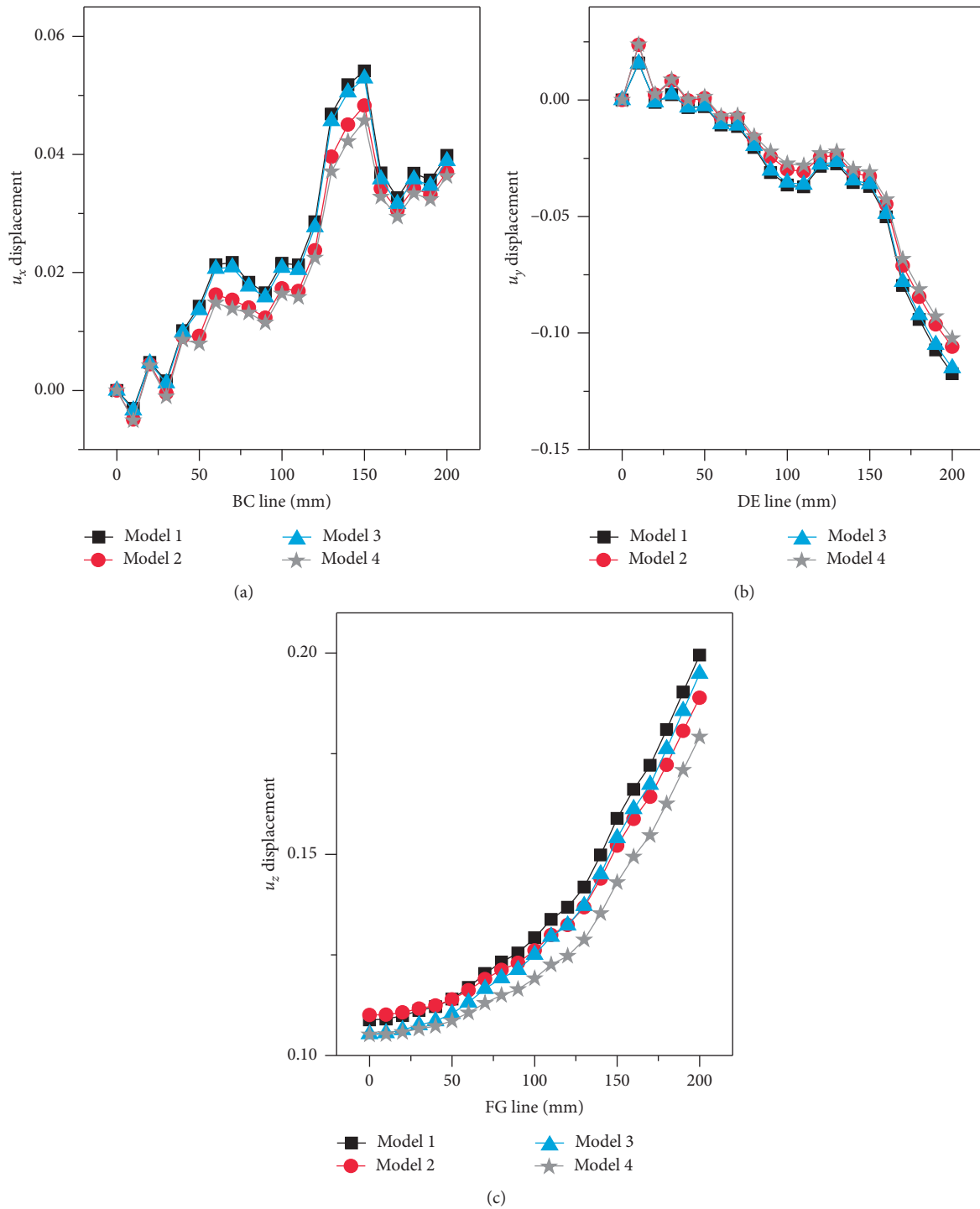


FIGURE 19: The displacement distributions along the CLT plate thickness direction. (a)  $u_x$  displacement of BC line; (b)  $u_y$  displacement of DE line; (c)  $u_z$  displacement of FG line.



## 5. Conclusions

This paper presented the CLT structural behavior in high-temperature environment via the well-known FEM. The basic equations for the CLT material were given. The temperature distribution and the structural deformation were demonstrated. Several numerical models were built for analyzing different impact factors. A CLT plate was established to investigate the board layer number influence, central hole dimension, conformation influences and hole position influence.

We have compared the results with the reference paper, and the great agreement verified the correctness of the FEM calculation process. Board layer plays a significant influence on the performance of CLT plate. With the change of layer, the temperature distributions of the plate and displacement in all coordinate directions have changed dramatically. In terms of other impact factors, from the contour plots of the temperature and displacements, we can notice these factors actually bring effect on the CLT plate, but from the displacement distribution in all coordinate directions along the borderline and thickness, these factors are not as obvious as the number of plate layers. And the  $u_z$  displacement component was more sensitive to the influence factors. The investigation of CLT structures brings reference data for the application of CLT material.

## Data Availability

The data supporting the conclusions of the study are present in the manuscript.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors would like to acknowledge the financial support from Chinese Universities Scientific Fund, no. 300102218214 and Chinese Universities Scientific Fund, no. 300102219211.

## References

- [1] A. J. Barotto, S. Monteoliva, J. Gyenge, A. Martinez-Meier, and M. E. Fernandez, "Functional relationships between wood structure and vulnerability to xylem cavitation in races of *Eucalyptus globulus* differing in wood density," *Tree Physiology*, vol. 38, no. 2, pp. 243–251, 2018.
- [2] C. Chen, D. Tu, Q. Zhou et al., "Development and evaluation of a surface-densified wood composite with an asymmetric structure," *Construction and Building Materials*, vol. 242, Article ID 118007, 2020.
- [3] E. S. Flores, I. Dayyani, R. M. Ajaj et al., "Analysis of cross-laminated timber by computational homogenisation and experimental validation," *Composite Structures*, vol. 121, pp. 386–394, 2015.
- [4] Y. Jiang and R. Crocetti, "CLT-concrete composite floors with notched shear connectors," *Construction and Building Materials*, vol. 195, pp. 127–139, 2019.
- [5] A. Hassanieh, H. R. Valipour, and M. A. Bradford, "Composite connections between CLT slab and steel beam: experiments and empirical models," *Journal of Constructional Steel Research*, vol. 138, pp. 823–836, 2017.
- [6] L.-M. Ottenhaus, M. Li, and T. Smith, "Structural performance of large-scale dowelled CLT connections under monotonic and cyclic loading," *Engineering Structures*, vol. 176, pp. 41–48, 2018.
- [7] E. Eurocode, 2: *Design of Concrete Structures-Part 1-2: General Rules-Structural Fire Design*, European Concrete Platform, Brussels, Belgium, 2004.
- [8] M. Fragiaco, A. Menis, I. Clemente, G. Bochicchio, and A. Ceccotti, "Fire resistance of cross-laminated timber panels loaded out of plane," *Journal of Structural Engineering*, vol. 139, no. 12, Article ID 04013018, 2013.
- [9] M. Klippel, J. Schmid, A. Frangi, and G. Fink, "Fire Design of glued-laminated timber beams with regard to the adhesive performance using the reduced cross-section method," vol. 2014, pp. 339–348, in *Proceedings of the 1st Meeting of the International Network on Timber Engineering Research (INTER 2014)*, vol. 2014, pp. 339–348, Timber Scientific Publishing, KIT Holzbau und Baukonstruktionen, Bath, UK, September 2004.
- [10] M. Klippel, *Fire Safety of Bonded Structural Timber Elements*, ETH Zurich, Zürich, Switzerland, 2014.
- [11] J. Schmid, J. König, and J. Köhler, "Fire-exposed cross-laminated timber—modelling and tests," in *Proceedings of the World Conference on Timber Engineering*, Trentino, Italy, June 2010.
- [12] J. Schmid, M. Klippel, A. Just, A. Frangi, and M. Tiso, "Simulation of the fire resistance of cross-laminated timber (CLT)," *Fire Technology*, vol. 54, no. 5, pp. 1113–1148, 2018.
- [13] J. Schmid, N. Werther, M. Klippel, and A. Frangi, "Structural fire design-statement on the design of cross-laminated timber (CLT)," *Civil Engineering Research Journal*, vol. 7, no. 5, pp. 113–117, 2019.
- [14] A. Frangi, M. Fontana, E. Hugli, and R. Jübstl, "Experimental analysis of cross-laminated timber panels in fire," *Fire Safety Journal*, vol. 44, no. 8, pp. 1078–1087, 2009.
- [15] A. Frangi, M. Fontana, M. Knobloch, and G. Bochicchio, "Fire behaviour of cross-laminated solid timber panels," *Fire Safety Science*, vol. 9, pp. 1279–1290, 2008.
- [16] Y. Wang, J. Zhang, F. Mei, J. Liao, and W. Li, "Experimental and numerical analysis on fire behaviour of loaded cross-laminated timber panels," *Advances in Structural Engineering*, vol. 23, no. 1, pp. 22–36, 2020.
- [17] C. Bedon and M. Fragiaco, "Fire resistance of in-plane compressed log-house timber walls with partial thermal insulation," *Buildings*, vol. 8, no. 10, p. 131, 2018.
- [18] C. H. Luce and B. C. Wemple, "Introduction to special issue on hydrologic and geomorphic effects of forest roads," *Earth Surface Processes and Landforms*, vol. 26, no. 2, pp. 111–113, 2001.
- [19] E. I. S. Flores, K. Saavedra, J. Hinojosa, Y. Chandra, and R. Das, "Multi-scale modelling of rolling shear failure in cross-laminated timber structures by homogenisation and cohesive zone models," *International Journal of Solids and Structures*, vol. 81, pp. 219–232, 2016.
- [20] D. Stobacz, *Behavior of Wood in Case of Fire: Proposal for a Stochastic Dimensioning of Structural Elements*, Müller, Marshfield, WI, USA, 2008.
- [21] Z. H. Wang, L. Li, and Y. X. Zhang, "Bond-slip model considering freeze-thaw damage effect of concrete and its application," *Engineering Structures*, vol. 201, Article ID 109831, 2019.



- [22] Z. H. Wang, L. Li, Y. X. Zhang et al., "Reinforcement model considering slip effect," *Engineering Structures*, vol. 198, Article ID 109493, 2019.
- [23] Y. Song, "Finite-element implementation of piezoelectric energy harvesting system from vibrations of railway bridge," *Journal of Energy Engineering*, vol. 145, Article ID 04018076, 2019.
- [24] J. Wu, Q. S. Yan, S. P. Huang, C. Zou, J. T. Zhong, and W. F. Wang, "Finite element model updating in bridge structures using Kriging model and Latin hypercube sampling method," *Advances in Civil Engineering*, vol. 2018, Article ID 8980756, 11 pages, 2018.
- [25] S. Nasrin and A. Ibrahim, "Finite-element modeling of UHPC hybrid bridge deck connections," *International Journal of Advanced Structural Engineering*, vol. 10, no. 3, pp. 199–210, 2018.
- [26] H. Takeda and J. R. Mehoff, "WALL2D: a model for predicting heat transfer through wood-stud walls exposed to fire," *Fire and Materials*, vol. 22, no. 4, pp. 133–140, 1998.
- [27] J. M. Franssen, "SAFIR: a thermal/structural program for modeling structures under fire," *Engineering Journal-American Institute of Steel Construction Inc*, vol. 42, no. 3, pp. 143–158, 2005.
- [28] A. Paolini, S. Kollmannsberger, C. Winter et al., "A high-order finite element model for vibration analysis of cross-laminated timber assemblies," *Building Acoustics*, vol. 24, no. 3, pp. 135–158, 2017.
- [29] M. Betti, M. Brunetti, M. P. Lauriola, M. Nocetti, F. Ravalli, and B. Pizzo, "Comparison of newly proposed test methods to evaluate the bonding quality of Cross-Laminated Timber (CLT) panels by means of experimental data and finite element (FE) analysis," *Construction and Building Materials*, vol. 125, pp. 952–963, 2016.
- [30] A. R. Díaz, E. I. Saavedra Flores, S. J. Yanez, D. A. Vasco, J. C. Pina, and C. F. Guzmán, "Multiscale modeling of the thermal conductivity of wood and its application to cross-laminated timber," *International Journal of Thermal Sciences*, vol. 144, pp. 79–92, 2019.

## Research Article

# On the Geodesic Identification of Vertices in Convex Plane Graphs

Fawaz E. Alsaadi,<sup>1</sup> Muhammad Salman ,<sup>2</sup> Masood Ur Rehman ,<sup>3</sup> Abdul Rauf Khan,<sup>4</sup> Jinde Cao ,<sup>5,6</sup> and Madini Obad Alassafi<sup>1</sup>

<sup>1</sup>Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

<sup>2</sup>Department of Mathematics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

<sup>3</sup>Department of Basic Sciences, Balochistan University of Engineering and Technology Khuzdar, Khuzdar 89100, Pakistan

<sup>4</sup>Department of Mathematics, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan

<sup>5</sup>School of Mathematics, Southeast University, Nanjing 210096, China

<sup>6</sup>Yonsei Frontier Lab, Yonsei University, Seoul 03722, Republic of Korea

Correspondence should be addressed to Masood Ur Rehman; masoodqau27@gmail.com

Received 20 August 2020; Accepted 1 December 2020; Published 28 December 2020

Academic Editor: Muhammad Javaid

Copyright © 2020 Fawaz E. Alsaadi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A shortest path between two vertices  $u$  and  $v$  in a connected graph  $G$  is a  $u - v$  geodesic. A vertex  $w$  of  $G$  performs the geodesic identification for the vertices in a pair  $(u, v)$  if either  $v$  belongs to a  $u - w$  geodesic or  $u$  belongs to a  $v - w$  geodesic. The minimum number of vertices performing the geodesic identification for each pair of vertices in  $G$  is called the strong metric dimension of  $G$ . In this paper, we solve the strong metric dimension problem for three convex plane graphs by performing the geodesic identification of their vertices.

## 1. Introduction

The identification of vertices of a graph using various graph parameters is a fascinating problem for researchers. In the literature, almost over 1000 articles are contributed to explore the identification of vertices in a remarkable way by using various graph theory concepts including graph coloring, labeling of vertices, domination in graphs, vertex covering, graph automorphisms with symmetry breaking technique, independence of vertices, and by defining the metric on graphs, to name a few.

By defining the metric on a graph, the problem of identification of vertices attracted many researchers due to its significant applications in several extents including verification, security, and discovery of networks [1], the chemistry of pharmaceuticals for drug designing [2], mastermind game strategies [3], navigation of robots [4], connected joins in graphs [5], and solution of coin weighing

problems [6]. Because of these practical significances of this problem, from the last two decades, numerous researchers identified vertices of a graph by considering the metric-related well-known concept of the metric dimension [2, 7, 8]. Later on, many researchers extended the study of this concept by defining its several variations including the fractional metric dimension [9], the resolving domination [10], the doubly metric dimension [11], the independent metric dimension [12], the weighted metric dimension [13], the  $k$ -metric dimension [14], the solid metric dimension [15], the mixed metric dimension [16], the local metric dimension [17], the simultaneous metric dimension [18], the strong metric dimension [5], and the connected metric dimension [19].

This paper is aimed to identify the vertices of three convex plane graphs by using geodesics between them, which provides the strong metric dimension of these graphs.

## 2. Geodesic Identification: The Strong Metric Dimension

Let  $G$  be a simple connected graph. A shortest path between two vertices  $x$  and  $y$  in  $G$  is known as a  $x - y$  geodesic. The geodesic identification of vertices is equivalently the strong resolvability of vertices, which is defined as follows: a vertex  $w$  of  $G$  performs the geodesic identification for the vertices in a pair  $(x, y)$  (i.e.,  $w$  strongly resolves the vertices in a pair  $(x, y)$ ) if either  $x$  belongs to (lies on) a  $y - w$  geodesic or  $y$  belongs to (lies on) a  $x - w$  geodesic. A set  $S$  of vertices in  $G$  is said to be a strong metric generator for  $G$  if for every pair of vertices of  $G$ , there is always a vertex in  $S$  which performs the geodesic identification for the vertices in the pair. The cardinality of such a smallest set  $S$  is called the strong metric dimension of  $G$ , denoted by  $\text{sdim}(G)$  [5].

Sebö and Tannier initiated the study of geodesic identification of vertices by defining the concepts of strong metric generator and strong metric dimension in 2004 [5]. Later on, this study was extended by many researchers, and they contributed to the literature with a variety of remarkable research work. To develop the readers interest, we shortly survey the strong metric dimension problem as follows:

- (i) The strong metric dimension problem has been solved for Sierpiński graph in [20], for hamming graphs in [21], for some convex polytopes in [22, 23], for wheel related graphs (including  $n$ -fold wheel, sunflower, helm, and friendship graphs) in [24], for path, cycle, complete, complete bipartite, and tree graphs in [25], for Cayley graphs in [26], for Cartesian sum graphs in [27], for the power graph of a finite group in [28], for distance-hereditary graphs in [29], for generalized butterfly and starbarbell graphs in [30], for antiprism and king graphs in [31], for sun, windmill, and Möbius ladder graphs in [32], and for crossed prism in [33].
- (ii) The strong metric dimension of various products of graphs including Cartesian product, direct product, strong product, lexicographic product, rooted product, and corona product has been supplied through the articles in [26, 27, 30, 31, 33–39].
- (iii) The fractional version of the strong metric dimension problem has been introduced in [40] and further studied for various graphs and graph products in [41, 42].
- (iv) The technique of the computation of strong metric dimension with the concept of the vertex cover number has been provided in [38] by proposing the construction of strong resolving graph of a connected graph. Furthermore, the article in [25] supplied some fundamental realizations and characterizations of the strong metric dimension problem in connection with the strong resolving graph.
- (v) To solve the strong metric dimension problem, the genetic algorithmic approach is used in [43], and the variable neighborhood search method is used in [44].
- (vi) The article in [45] supplied the Nordhaus–Gaddum type results whenever the strong metric dimension problem was solved for graphs and their compliments.
- (vii) To compute the strong metric dimension of graphs using optimization techniques, the integer linear programming model for the strong metric dimension problem was formulated in [22].
- (viii) The complexity and the optimal approximability in the computation of the strong metric dimension problem of graphs have been discussed in [38, 46].
- (ix) Furthermore, we refer a survey in [47] and the Ph.D. thesis in [48] to the readers having interest in the computation of the strong metric dimension of graphs.

With this paper, we extend the study of the identification of vertices using the concept of strong metric dimension. We consider three convex plane graphs and investigate their strong metric dimension by performing the geodesic identification of vertices in the graphs.

## 3. Basic Works

Let  $G$  be a simple and connected graph with vertex set  $V(G)$  and edge set  $E(G)$ . We denote the adjacent vertices  $u$  and  $v$  by  $u \sim v$  and nonadjacent vertices by  $u \not\sim v$  in  $G$ . The neighborhood of a vertex  $v$  in  $G$  is  $N(v) = \{u \in V(G) : u \sim v \text{ in } G\}$ . The number of vertices adjacent with a vertex  $v$  is called its degree and is denoted by  $d(v)$ . The metric on  $G$  is a mapping  $d: V(G) \times V(G) \rightarrow \mathbb{Z}^+ \cup \{0\}$  defined by  $d(x, y) = l$ , where  $l$  is the length of (the number of edges in) a  $x - y$  geodesic. Accordingly, a vertex  $w$  of  $G$  performs the geodesic identification for the vertices in a pair  $(x, y)$  if and only if either  $d(x, w) = d(x, y) + d(y, w)$  or  $d(y, w) = d(y, x) + d(x, w)$ .

Kratika et al. in [22] supplied the following two results, which are useful tools for the geodesic identification of vertices.

**Lemma 1** (see [22]). *Let  $(u, v)$  be a pair of distinct vertices in a connected graph  $G$  such that*

$$\begin{aligned} d(u, v) &\geq d(w, v), & \forall w \in N(u), \\ d(u, v) &\geq d(w, u), & \forall w \in N(v). \end{aligned} \quad (1)$$

Then, there is no vertex in  $V(G) - \{u, v\}$  which performs the geodesic identification for the pair  $(u, v)$ .

**Proposition 1** (see [22]). *If  $S$  is a strong metric generator for a connected graphs  $G$ , then for every pair  $(u, v)$  of distinct vertices in  $G$  satisfying both the conditions in (1), either  $u \in S$  or  $v \in S$ .*

For a connected graph  $G$ , the number  $\text{diam}(G) = \max\{\text{ecc}(v) = \max_{u \in V(G)} d(v, u); v \in V(G)\}$  is called the diameter of  $G$ , where  $\text{ecc}(v)$  is the eccentricity of a vertex  $v$ . The following result, provided by Kratica et al. in [22], is also a useful tool for the geodesic identification of vertices.

**Proposition 2** (see [22]). *If  $S$  is a strong metric generator for a connected graphs  $G$ , then for every pair  $(u, v)$  of distinct vertices in  $G$  such that  $d(u, v) = \text{diam}(G)$ , either  $u \in S$  or  $v \in S$ .*

A vertex  $v$  of  $G$  for which  $d(v) = 1$  is known as a leaf. The following result describes the geodesic identification of leaves in a connected graph.

**Lemma 2.** *If  $S$  is a strong metric generator for a connected graphs  $G$  and  $(u, v)$  is a pair of distinct leaves in  $G$ , then either  $u \in S$  or  $v \in S$ .*

*Proof.* Since  $u$  and  $v$  are leaves,  $d(u) = 1 = d(v)$ . Without loss of generality, let  $N(u) = \{x\}$  and  $N(v) = \{y\}$ . Now, if  $x = y$ , then

$$\begin{aligned} d(u, v) &= 2 > d(x, v), \\ d(u, v) &= 2 > d(x, u). \end{aligned} \quad (2)$$

If  $x \neq y$ , then

$$d(u, v) = d(u, x) + d(x, y) + d(y, v) = d(x, y) + 2, \quad (3)$$

$$d(x, v) = d(x, y) + d(y, v) = d(x, y) + 1 < d(u, v), \quad (4)$$

$$d(y, u) = d(y, x) + d(x, u) = d(x, y) + 1 < d(u, v). \quad (5)$$

The identities (2)–(5) provide that the pair  $(u, v)$  satisfies both the conditions in (1), and hence, Proposition 1 yields the required result.

Lemma 2 supplies the following straightforward proposition.  $\square$

**Proposition 3.** *If  $S$  is a strong metric generator for a connected graph  $G$  and  $L$  is the set of  $m$  leaves in  $G$ , then  $S$  must contain at least  $m - 1$  leaves from  $L$ .*

#### 4. Convex Plane Graphs

A graph  $G$  is convex if a straight line joining any two vertices of  $G$  entirely lies within the region occupied by  $G$ . A graph  $G$  is a plane graph if it is free from crossing of edges. Throughout this section, it should be helpful to keep in mind the following points:

- (i) The right arrow ( $\longrightarrow$ ) in the super script of the notation of geodesics will indicate forward nature of geodesics (see Figure 1).
- (ii) The left arrow ( $\longleftarrow$ ) in the super script of the notation of geodesics will indicate backward nature of geodesics (see Figure 1).

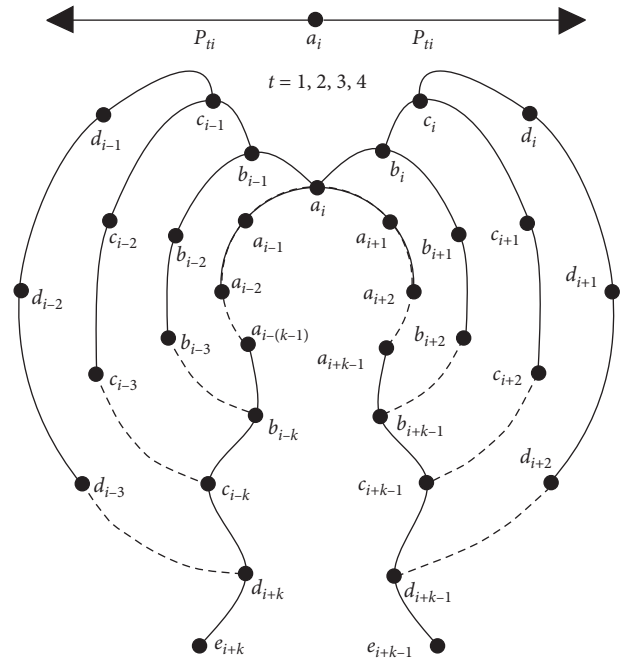


FIGURE 1: Illustrating forward and backward natures of  $a_i - a_{i+k}$  and  $a_i - a_{i+k-1}$  geodesics.

- (iii) The indices greater than  $n$  or less than 1 will be taken modulo  $n$ .

In the next sections, we investigate the strong metric dimension of three convex plane graphs.

**4.1. Convex Plane Graph  $S_n^p$ .** The graph of a convex polytope  $S_n$  is defined in [49] for  $n \geq 3$ . The vertex set of  $S_n$  is  $V(S_n) = \{a_i, b_i, c_i, d_i; 1 \leq i \leq n\}$ , and its edge set is

$$\begin{aligned} E(S_n) &= \{a_i \sim a_{i+1}, a_i \sim b_i, b_i \sim b_{i+1}, b_i \sim a_{i+1}; 1 \leq i \leq n\} \\ &\cup \{b_i \sim c_i, c_i \sim c_{i+1}, c_i \sim d_i, d_i \sim d_{i+1}; 1 \leq i \leq n\}. \end{aligned} \quad (6)$$

The convex plane graph  $S_n^p$  ( $p$  for pendant) can be obtained from the graph of a convex polytope  $S_n$  by attaching one pendant vertex (leaf)  $e_i$  to the vertex  $d_i$  of  $S_n$  for each  $1 \leq i \leq n$  (see Figure 2) [50]. Thus, the vertex and edge sets of  $S_n^p$  are

$$\begin{aligned} V(S_n^p) &= V(S_n) \cup \{e_i; 1 \leq i \leq n\}, \\ E(S_n^p) &= E(S_n) \cup \{d_i \sim e_i; 1 \leq i \leq n\}. \end{aligned} \quad (7)$$

We investigate the strong matrix dimension of  $S_n^p$  by supplying the following main result.

**Theorem 1.** *For  $n \geq 3$ , let  $S_n^p$  be a convex plane graph. Then,*

$$\text{sdim}(S_n^p) = \begin{cases} n, & \text{when } n \text{ is odd,} \\ \frac{3n}{2}, & \text{when } n \text{ is even.} \end{cases} \quad (8)$$

The next four lemmas will lead the proof of Theorem 1.

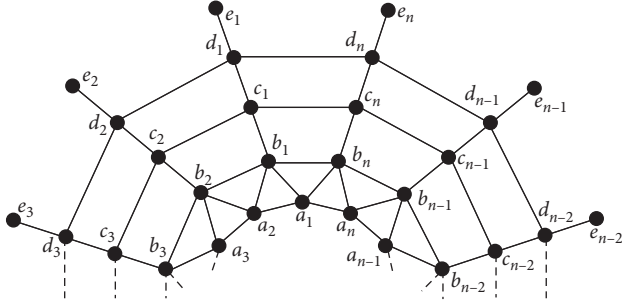


FIGURE 2: The general half view of the convex plane graph  $S_n^p$ .

**Lemma 3.** For odd values of  $n \geq 3$ , if  $S$  is a strong metric generator for  $S_n^p$ , then  $|S| \geq n$ .

*Proof.* Let  $n = 2k + 1$  with  $k \geq 1$ , and consider Table 1 for  $1 \leq i \leq n$ .

In  $S_n^p$ , note the following points:

P1:  $L = \{e_i; 1 \leq i \leq n\}$  is the set of leaves.

P2: for each  $1 \leq i \leq n$ ,  $d(a_i, e_{i+k}) = k + 4$ . It follows, according to Table 1, that the pair  $(a_i, e_{i+k})$  satisfies both the conditions in (1).

P3:  $\text{diam}(S_n^p) = k + 4 = d(a_i, e_{i+k})$  for each  $1 \leq i \leq n$ .

From P1, without loss of generality, we suppose that  $e_2, e_3, \dots, e_n \in S$  by Proposition 3. Then, since the pair  $(a_{k+1}, e_1)$  satisfies P2 and P3, we have either  $a_{k+1} \in S$  or  $e_1 \in S$ , by Propositions 1 and 2. Hence,  $|S| \geq n - 1 + 1 = n$ .  $\square$

**Lemma 4.** For even values of  $n \geq 4$ , if  $S$  is a strong metric generator for  $S_n^p$ , then  $|S| \geq (3n/2)$ .

*Proof.* Let  $n = 2k$  with  $k \geq 2$ , and consider Table 2 for  $1 \leq i \leq n$ .

In  $S_n^p$ , the following points hold:

$$\overrightarrow{P_{1i}}: a_i \sim a_{i+1} \sim a_{i+2} \sim \dots \sim a_{i+k} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (9)$$

$$\overrightarrow{P_{2i}}: a_i \sim b_i \sim b_{i+1} \sim b_{i+2} \sim \dots \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (10)$$

$$\overrightarrow{P_{3i}}: a_i \sim b_i \sim c_i \sim c_{i+1} \sim c_{i+2} \sim \dots \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (11)$$

$$\overrightarrow{P_{4i}}: a_i \sim b_i \sim c_i \sim d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k} \sim e_{i+k}, \quad (12)$$

$$\overleftarrow{P_{1i}}: a_i \sim a_{i-1} \sim a_{i-2} \sim \dots \sim \{a_{i-k} = a_{i+k+1}\} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (13)$$

$$\overleftarrow{P_{2i}}: a_i \sim b_{i-1} \sim b_{i-2} \sim \dots \sim \{b_{i-k} = b_{i+k+1}\} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (14)$$

$$\overleftarrow{P_{3i}}: a_i \sim b_{i-1} \sim c_{i-1} \sim c_{i-2} \sim \dots \sim \{c_{i-k} = c_{i+k+1}\} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (15)$$

$$\overleftarrow{P_{4i}}: a_i \sim b_{i-1} \sim c_{i-1} \sim d_{i-1} \sim d_{i-2} \sim \dots \sim \{d_{i-k} = d_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}. \quad (16)$$

TABLE 1: Vertices willing to perform the geodesic identification in  $S_n^p$ .

Vertex $v$	$N(v)$	$d(w, \cdot)$ for all $w \in N(v)$
$a_i$	$\{a_{i+1}, a_{i-1}, b_{i-1}, b_i\}$	$d(w, e_{i+k}) = k + 3$
$e_{i+k}$	$\{d_{i+k}\}$	$d(w, a_i) = k + 3$

P1:  $L = \{e_i; 1 \leq i \leq n\}$  is the set of leaves.

P2: for each  $1 \leq i \leq n$ ,  $d(a_i, e_{i+k-1}) = d(a_i, e_{i+k}) = d(b_i, e_{i+k}) = k + 3$ . Therefore, Table 2 yields that the pairs  $(a_i, e_{i+k-1})$ ,  $(a_i, e_{i+k})$  and  $(b_i, e_{i+k})$  satisfy both the conditions in (1) for each  $1 \leq i \leq n$ . Table 2 further implies that the pair  $(a_j, a_{j+k})$  satisfies both the conditions in (1) because  $d(a_j, a_{j+k}) = k$  for each  $1 \leq j \leq k$ .

P3:  $\text{diam}(S_n^p) = k + 3 = d(a_i, e_{i+k-1}) = d(a_i, e_{i+k}) = d(b_i, e_{i+k})$  for each  $1 \leq i \leq n$ .

Due to P1, without loss of generality, we let  $e_2, e_3, \dots, e_n \in S$ , by Proposition 3. Then, as the pairs  $(a_{k+2}, e_1)$ ,  $(a_{k+1}, e_1)$  and  $(b_{k+1}, e_1)$  satisfy P2 and P3, we must have either  $a_{k+2} \in S$  or  $a_{k+1} \in S$  or  $b_{k+1} \in S$  or  $e_1 \in S$ , by Propositions 1 and 2. For the geodesic identification of these three pairs, it is enough to consider  $e_1$  in  $S$ . Furthermore, for each  $1 \leq j \leq k$ , either  $a_j \in S$  or  $a_{j+k} \in S$ , due to P2 and Proposition 1. Hence,  $|S| \geq n - 1 + 1 + k = (3n/2)$ .  $\square$

**Lemma 5.** If  $n = 2k + 1$  with  $k \geq 1$ , then the set  $S = \{e_1, e_2, \dots, e_n\} \subset V(S_n^p)$  is a strong metric generator for  $S_n^p$ .

*Proof.* For any  $s \in S$  and any  $u \in V(S_n^p) - \{s\}$ , since the vertex  $s$  performs the geodesic identification for the pair  $(s, u)$ , we have to perform the geodesic identification for every pair  $(x, y)$  of distinct vertices with  $x, y \in V(S_n^p) - S$ . For each  $1 \leq i \leq n$ , consider the following  $a_i - e_{i+k}$  geodesics of length  $k + 4$ :

TABLE 2: Vertices willing to perform the geodesic identification in  $S_n^p$ .

Vertex $v$	$N(v)$	$d(w, \cdot)$ for all $w \in N(v)$
$a_i$	$\{a_{i-1}, a_{i+1}, b_{i-1}, b_i\}$	$d(w, a_{i+k}) \in \{k-1, k\}$ $d(w, e_{i+k-1}) \in \{k+2, k+3\}$ $d(w, e_{i+k}) \in \{k+2, k+3\}$
$a_{i+k}$	$\{a_{i+k-1}, a_{i+k+1}, b_{i+k-1}, b_{i+k}\}$	$d(w, a_i) \in \{k-1, k\}$
$b_i$	$\{a_i, a_{i+1}, b_{i-1}, b_{i+1}, c_i\}$	$d(w, e_{i+k}) \in \{k+2, k+3\}$
$e_{i+k-1}$	$\{d_{i+k-1}\}$	$d(w, a_i) = k+2$
$e_{i+k}$	$\{d_{i+k}\}$	$d(w, a_i) = k+2$ $d(w, b_i) = k+2$

Each geodesic, listed from (9) to (16), provides that the vertex  $e_{i+k} \in S$  performs the geodesic identification for every pair  $(x, y)$  with  $x, y \in V(S_n^p) - S$ . It follows that  $S$  is a strong metric generator for  $S_n^p$ .  $\square$

**Lemma 6.** *If  $n = 2k$  with  $k \geq 2$ , then the set  $S = \{a_1, a_2, \dots, a_k, e_1, e_2, \dots, e_n\} \subset V(S_n^p)$  is a strong metric generator for  $S_n^p$ .*

*Proof.* We have to perform the geodesic identification each pair  $(u, v)$  of distinct vertices for  $u, v \in V(S_n^p) - S$  because for any  $s \in S$  and any  $v \in V(S_n^p) - \{s\}$ , the pair  $(s, v)$  possesses the geodesic identification by the vertex  $s$ . For each  $1 \leq i \leq n$ , consider the following  $a_i - e_{i+k-1}$  geodesics and  $a_i - e_{i+k}$  geodesics of length  $k+3$ :

$$\vec{P}_{1i}: a_i \sim a_{i+1} \sim a_{i+2} \sim \dots \sim a_{i+k-1} \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (17)$$

$$\vec{P}_{2i}: a_i \sim b_i \sim b_{i+1} \sim b_{i+2} \sim \dots \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (18)$$

$$\vec{P}_{3i}: a_i \sim b_i \sim c_i \sim c_{i+1} \sim c_{i+2} \sim \dots \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (19)$$

$$\vec{P}_{4i}: a_i \sim b_i \sim c_i \sim d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k-1} \sim e_{i+k-1}, \quad (20)$$

$$\overleftarrow{P}_{1i}: a_i \sim a_{i-1} \sim a_{i-2} \sim \dots \sim \{a_{i-(k-1)} = a_{i+k+1}\} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (21)$$

$$\overleftarrow{P}_{2i}: a_i \sim b_{i-1} \sim b_{i-2} \sim \dots \sim \{b_{i-(k-1)} = b_{i+k+1}\} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (22)$$

$$\overleftarrow{P}_{3i}: a_i \sim b_{i-1} \sim c_{i-1} \sim c_{i-2} \sim \dots \sim \{c_{i-(k-1)} = c_{i+k+1}\} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (23)$$

$$\overleftarrow{P}_{4i}: a_i \sim b_{i-1} \sim c_{i-1} \sim d_{i-1} \sim d_{i-2} \sim \dots \sim \{d_{i-(k-1)} = d_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}. \quad (24)$$

Except the pairs  $(a_i, a_j)$  for  $k \leq i \leq n-1$  and  $i+1 \leq j \leq n$ , all the pairs of vertices owned the geodesic identification by the vertices  $e_{i+k}$  and  $e_{i+k-1}$  due to the geodesics listed from (17) to (24). By considering the following  $a_k - a_n$  geodesic

$$\vec{P}: a_k \sim a_{k+1} \sim a_{k+2} \sim \dots \sim a_{n-1} \sim a_n, \quad (25)$$

we get that the vertex  $a_k$  performs the geodesic identification for all the remaining pairs  $(a_i, a_j)$ . Hence,  $S$  is a strong metric generator for  $S_n^p$ .  $\square$

*Proof* (Theorem 1). The lower bound for the strong metric dimension of  $S_n^p$  is provided by Lemmas 3 and 4, whereas the upper bound is due to Lemmas 5 and 6.  $\square$

**4.2. Convex Plane Graph  $T_n^p$ .** The graph of a convex polytope  $T_n$  is defined in [49] for  $n \geq 3$ . The vertex set of  $T_n$  is  $V(T_n) = \{a_i, b_i, c_i, d_i; 1 \leq i \leq n\}$ , and its edge set is



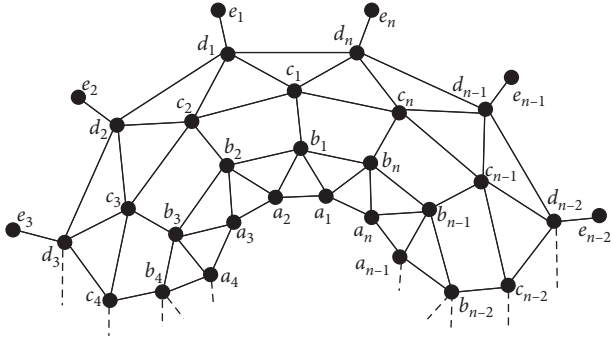


FIGURE 3: The general half view of the convex plane graph  $T_n^p$ .

$$\begin{aligned} E(T_n) = & \{a_i \sim a_{i+1}, a_i \sim b_i, b_i \sim b_{i+1}, b_i \sim a_{i+1}, b_i \sim c_i; 1 \leq i \leq n\} \\ & \cup \{c_i \sim c_{i+1}, c_i \sim d_i, c_i \sim d_{i-1}, d_i \sim d_{i+1}; 1 \leq i \leq n\}. \end{aligned} \quad (26)$$

The convex plane graph  $T_n^p$  ( $p$  for pendant) can be obtained from the graph of a convex polytope  $T_n$  by attaching one pendant vertex (leaf)  $e_i$  to the vertex  $d_i$  of  $T_n$  for each  $1 \leq i \leq n$  (see Figure 3) [50]. Accordingly, the vertex and edge sets of  $T_n^p$  are

$$\begin{aligned} V(T_n^p) &= V(T_n) \cup \{e_i; 1 \leq i \leq n\}, \\ E(T_n^p) &= E(T_n) \cup \{d_i \sim e_i; 1 \leq i \leq n\}. \end{aligned} \quad (27)$$

Mainly, we have the following result for the strong metric dimension of  $T_n^p$ .

**Theorem 2.** For  $n \geq 3$ , let  $T_n^p$  be a convex plane graph. Then,

$$\text{sdim}(T_n^p) = \begin{cases} 2n, & \text{when } n \text{ is odd,} \\ \frac{5n}{2}, & \text{when } n \text{ is even.} \end{cases} \quad (28)$$

The proof of Theorem 2 will be followed after the next four lemmas.

**Lemma 7.** For odd values of  $n \geq 3$ , if  $S$  is a strong metric generator for  $T_n^p$ , then  $|S| \geq 2n$ .

*Proof.* Let  $n = 2k + 1$  with  $k \geq 1$ , and consider Table 3 for  $1 \leq i \leq n$ .

Note the following three points in  $T_n^p$ :

P1:  $L = \{e_i; 1 \leq i \leq n\}$  is the set of leaves.

P2:  $d(a_i, e_{i+k-1}) = d(a_i, e_{i+k}) = d(b_i, e_{i+k}) = k + 3$  and  $d(a_i, c_{i+k}) = k + 2$  for each  $1 \leq i \leq n$ . Hence, Table 3 ensures that pairs of vertices in  $T_n^p$  satisfying both the conditions in (1) are  $(a_i, e_{i+k-1})$ ,  $(a_i, e_{i+k})$ ,  $(b_i, e_{i+k})$ , and  $(a_i, c_{i+k})$  for each  $1 \leq i \leq n$ .

P3:  $\text{diam}(T_n^p) = k + 3 = d(a_i, e_{i+k-1}) = d(a_i, e_{i+k}) = d(b_i, e_{i+k})$  for each  $1 \leq i \leq n$ .

By Proposition 3, let us take  $e_2, e_3, \dots, e_n \in S$  without loss of generality due to P1. Then, from the remaining pairs  $(a_{k+3}, e_1)$ ,  $(a_{k+2}, e_1)$ , and  $(b_{k+2}, e_1)$ , we must have either  $a_{k+3} \in S$  or  $a_{k+2} \in S$  or  $b_{k+2} \in S$  or  $e_1 \in S$ , by Propositions 1 and 2, because these pairs are satisfying P2 and P3. Here, it can be seen, by letting  $e_1 \in S$  only, that all these pairs possess the geodesic identification by the vertex  $e_1$ . Furthermore, for each  $1 \leq i \leq n$ , from the pair  $(a_i, c_{i+k})$  satisfying P2, either  $a_i \in S$  or  $c_{i+k} \in S$ , by Proposition 1. Hence,  $|S| \geq n - 1 + 1 + n = 2n$ .  $\square$

**Lemma 8.** For even values of  $n \geq 4$ , if  $S$  is a strong metric generator for  $T_n^p$ , then  $|S| \geq (5n/2)$ .

*Proof.* Let  $n = 2k$  with  $k \geq 2$ , and consider Table 4 for  $1 \leq i \leq n$ .

The following points are considerable in  $T_n^p$ :

P1:  $L = \{e_i; 1 \leq i \leq n\}$  is the set of leaves.

P2:  $d(a_i, e_{i-1}) = 4$ ,  $d(a_i, e_{i+k-1}) = k + 3$  and  $d(b_i, c_{i+k}) = k + 1$ , which implies according to Table 4, that the pairs  $(a_i, e_{i-1})$ ,  $(a_i, e_{i+k-1})$  and  $(b_i, c_{i+k})$  satisfy both the conditions in (1) for each  $1 \leq i \leq n$ . Moreover, for each  $1 \leq j \leq k$ ,  $d(a_j, a_{j+k}) = k$ , and so Table 4 concludes that the pair  $(a_j, a_{j+k})$  also satisfies both the conditions in (1).

P3:  $\text{diam}(T_n^p) = k + 3 = d(a_i, e_{i+k-1})$  for all  $1 \leq i \leq n$ .

Now, without loss of generality, let us take  $e_2, e_3, \dots, e_n$  in  $S$ , by Proposition 3 and due to P1. Then, from the remaining pairs  $(a_{k+2}, e_1)$  and  $(a_2, e_1)$  satisfying P2 and P3, respectively, we must have either  $a_{k+2} \in S$  or  $e_1 \in S$  by Proposition 2, and either  $a_2 \in S$  or  $e_1 \in S$  by Proposition 1. If we take  $e_1$  in  $S$ , then this vertex performs the geodesic identification for both the remaining pairs. Also, Proposition 1 yields that either  $b_i \in S$  or  $c_{i+k} \in S$  for each  $1 \leq i \leq n$  and either  $a_j \in S$  or  $a_{j+k} \in S$  for each  $1 \leq j \leq k$  because of the point P2. Therefore,  $|S| \geq n - 1 + 1 + n + k = (5n/2)$ .  $\square$

**Lemma 9.** For  $n = 2k + 1$  with  $k \geq 1$ , the set  $S = \{a_1, a_2, \dots, a_n, e_1, e_2, \dots, e_n\} \subset V(T_n^p)$  is a strong metric generator for  $T_n^p$ .

*Proof.* Note that every pair  $(v, s)$  with  $s \in S$  and  $v \in V(T_n^p) - \{s\}$  possesses the geodesic identification by the vertex  $s$ . It follows that we should perform the geodesic identification for every pair  $(u, v)$  of distinct vertices by the vertices in  $S$  whenever  $u, v \in V(T_n^p) - S$ . For each  $1 \leq i \leq n$ , except the pairs of vertices,

$$\begin{aligned} & (v, b_{i-k}), (v, b_{i+k}), (v, c_{i-k}), (v, c_{i+k}), (v, d_{i-k}), (v, d_{i+k}), \\ & \text{for } v \in \{b_i, c_i, d_i\}, \end{aligned} \quad (29)$$

All other pairs of vertices possess the geodesic identification by the vertices  $e_{i+k-1}$  and  $e_{i+k}$  using  $a_i - e_{i+k-1}$  geodesics and  $a_i - e_{i+k}$  geodesics of length  $k + 3$ , listed from the following equations:



TABLE 3: Vertices willing to perform the geodesic identification in  $T_n^p$ .

Vertex $v$	$N(v)$	$d(w, \cdot)$ for all $w \in N(v)$
$a_i$	$\{a_{i-1}, a_{i+1}, b_{i-1}, b_i\}$	$d(w, c_{i+k}) = k + 1$ $d(w, e_{i+k-1}) \in \{k + 2, k + 3\}$ $d(w, e_{i+k}) \in \{k + 2, k + 3\}$
$b_i$	$\{a_i, a_{i+1}, b_{i-1}, b_{i+1}, c_i\}$	$d(w, e_{i+k}) \in \{k + 2, k + 3\}$
$c_{i+k}$	$\{b_{i+k}, c_{i+k-1}, c_{i+k+1}, d_{i+k-1}, d_{i+k}\}$	$d(w, a_i) \in \{k + 1, k + 2\}$
$e_{i+k-1}$	$\{d_{i+k-1}\}$	$d(w, a_i) = k + 2$
$e_{i+k}$	$\{d_{i+k}\}$	$d(w, a_i) = k + 2$ $d(w, b_i) = k + 2$

TABLE 4: Vertices willing to perform the geodesic identification in  $T_n^p$ .

Vertex $v$	$N(v)$	$d(w, \cdot)$ for all $w \in N(v)$
$a_i$	$\{a_{i-1}, a_{i+1}, b_{i-1}, b_i\}$	$d(w, a_{i+k}) \in \{k - 1, k\}$ $d(w, e_{i+k-1}) = k + 2$ $d(w, e_{i-1}) \in \{3, 4\}$
$a_{i+k}$	$\{a_{i+k-1}, a_{i+k+1}, b_{i+k-1}, b_{i+k}\}$	$d(w, a_i) \in \{k - 1, k\}$
$b_i$	$\{a_i, a_{i+1}, b_{i-1}, b_{i+1}, c_i\}$	$d(w, c_{i+k}) \in \{k, k + 1\}$
$c_{i+k}$	$\{b_{i+k}, c_{i+k-1}, c_{i+k+1}, d_{i+k-1}, d_{i+k}\}$	$d(w, b_i) \in \{k, k + 1\}$
$e_{i-1}$	$\{d_{i-1}\}$	$d(w, a_i) = 3$
$e_{i+k-1}$	$\{d_{i+k-1}\}$	$d(w, a_i) = k + 2$

$$\overrightarrow{P_{1i}}: a_i \sim a_{i+1} \sim a_{i+2} \sim \dots \sim a_{i+k-1} \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (30)$$

$$\overrightarrow{P_{2i}}: a_i \sim b_i \sim b_{i+1} \sim b_{i+2} \sim \dots \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (31)$$

$$\overrightarrow{P_{3i}}: a_i \sim b_i \sim c_i \sim c_{i+1} \sim c_{i+2} \sim \dots \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (32)$$

$$\overrightarrow{P_{4i}}: a_i \sim b_i \sim c_i \sim d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k-1} \sim e_{i+k-1}, \quad (33)$$

$$\overleftarrow{P_{1i}}: a_i \sim a_{i-1} \sim a_{i-2} \sim \dots \sim a_{i-(k-1)} \sim b_{i-k} \sim \{c_{i-k} = c_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}, \quad (34)$$

$$\overleftarrow{P_{2i}}: a_i \sim b_{i-1} \sim b_{i-2} \sim \dots \sim b_{i-(k-1)} \sim b_{i-k} \sim \{c_{i-k} = c_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}, \quad (35)$$

$$\overleftarrow{P_{3i}}: a_i \sim b_{i-1} \sim c_{i-1} \sim c_{i-2} \sim \dots \sim c_{i-(k-1)} \sim \{c_{i-k} = c_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}, \quad (36)$$

$$\overleftarrow{P_{4i}}: a_i \sim b_{i-1} \sim c_{i-1} \sim d_{i-2} \sim d_{i-3} \sim \dots \sim d_{i-(k-1)} \sim \{d_{i-k} = d_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}. \quad (37)$$

Now, for each  $1 \leq i \leq n$ , by considering  $b_i - e_{i+k}$  geodesics, listed from (38) to (43), the vertex  $e_{i+k}$  performs the

geodesic identification for all the pairs of vertices listed in (29):

$$\overrightarrow{Q_{1i}}: b_i \sim b_{i+1} \sim b_{i+2} \sim \dots \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (38)$$

$$\overrightarrow{Q_{2i}}: b_i \sim c_i \sim c_{i+1} \sim c_{i+2} \sim \dots \sim c_{i+k} \sim d_{i+k} \sim e_{i+k}, \quad (39)$$

$$\overrightarrow{Q_{3i}}: b_i \sim c_i \sim d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k} \sim e_{i+k}, \quad (40)$$

$$\overleftarrow{Q}_{1i} : b_i \sim b_{i-1} \sim b_{i-2} \sim \dots \sim b_{i-k} \sim \{c_{i-k} = c_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}, \quad (41)$$

$$\overleftarrow{Q}_{2i} : b_i \sim c_i \sim c_{i-1} \sim c_{i-2} \sim \dots \sim \{c_{i-k} = c_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}, \quad (42)$$

$$\overleftarrow{Q}_{3i} : b_i \sim c_i \sim d_{i-1} \sim d_{i-2} \sim \dots \sim \{d_{i-k} = d_{i+k+1}\} \sim d_{i+k} \sim e_{i+k}. \quad (43)$$

It concludes that  $S$  is a strong metric generator for  $T_n^p$ .  $\square$

**Lemma 10.** *If  $n = 2k$  with  $k \geq 2$ , then the set*

$$S = \{a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_n, e_1, e_2, \dots, e_n\} \subset V(T_n^p), \quad (44)$$

*is a strong metric generator for  $T_n^p$ .*

*Proof.* Since the vertex  $s \in S$  performs the geodesic identification for every pair of vertices including  $s$ , it implies that we have to perform the geodesic identification by the vertices in  $S$  for all the pairs of vertices from  $V(T_n^p) - S$ . For this, consider the following  $a_i - e_{i+k-1}$  geodesics of length  $k + 3$ :

$$\overrightarrow{P}_{1i} : a_i \sim a_{i+1} \sim a_{i+2} \sim \dots \sim a_{i+k-1} \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (45)$$

$$\overrightarrow{P}_{2i} : a_i \sim b_i \sim c_i \sim c_{i+1} \sim c_{i+2} \sim \dots \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (46)$$

$$\overrightarrow{P}_{3i} : a_i \sim b_i \sim c_i \sim d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k-1} \sim e_{i+k-1}, \quad (47)$$

$$\overleftarrow{P}_{1i} : a_i \sim a_{i-1} \sim a_{i-2} \sim \dots \sim \{a_{i-(k-1)} = a_{i+k+1}\} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (48)$$

$$\overleftarrow{P}_{2i} : a_i \sim b_{i-1} \sim c_{i-1} \sim c_{i-2} \sim \dots \sim \{c_{i-(k-1)} = c_{i+k+1}\} \sim c_{i+k} \sim d_{i+k-1} \sim e_{i+k-1}, \quad (49)$$

$$\overleftarrow{P}_{3i} : a_i \sim b_{i-1} \sim c_{i-1} \sim d_{i-2} \sim d_{i-3} \sim \dots \sim \{d_{i-(k-1)} = d_{i+k+1}\} \sim d_{i+k} \sim d_{i+k-1} \sim e_{i+k-1}. \quad (50)$$

Due to geodesics, listed from (45) to (50), the vertex  $e_{i+k-1}$  performs the geodesic identification for all the pairs of vertices except the pairs  $(c_i, c_{i+k})$ ,  $(d_i, c_{i+k})$ ,  $(d_i, d_{i+k})$  for all  $1 \leq i \leq n$  and the pair  $(a_j, d_{j-1})$  for each  $k + 1 \leq j \leq n$ . These pairs are identified as follows:

- (i) For each  $1 \leq i \leq n$ , the vertex  $b_i$  performs the geodesic identification for the pair  $(c_i, c_{i+k})$  due to the geodesic  $\overrightarrow{Q}_i : b_i \sim c_i \sim c_{i+1} \sim c_{i+2} \sim \dots \sim c_{i+k-1} \sim c_{i+k}$ .
- (ii) For each  $1 \leq i \leq n$ , the vertex  $b_{i+k}$  performs the geodesic identification for the pair  $(d_i, c_{i+k})$  due to the geodesic  $\overrightarrow{R}_i : d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k-1} \sim c_{i+k} \sim b_{i+k}$ .
- (iii) For each  $1 \leq i \leq n$ , the vertex  $e_{i+k}$  performs the geodesic identification for the pair  $(d_i, d_{i+k})$  due to the geodesic  $\overrightarrow{S}_i : d_i \sim d_{i+1} \sim d_{i+2} \sim \dots \sim d_{i+k-1} \sim d_{i+k} \sim e_{i+k}$ .
- (iv) For each  $k + 1 \leq j \leq n$ , the vertex  $e_{j-1}$  performs the geodesic identification for the pair  $(a_j, d_{j-1})$  due to the geodesic  $T_i : a_j \sim b_{j-1} \sim d_{j-1} \sim e_{j-1}$ .

Hence, we conclude that  $S$  is a strong metric generator for  $T_n^p$ .  $\square$

*Proof.* (Theorem 2). The lower bound for the strong metric dimension of  $T_n^p$  is provided by Lemmas 7 and 8, whereas the upper bound is due to Lemmas 9 and 10.  $\square$

**4.3. Convex Plane Graph  $U_n^p$ .** The graph of a convex polytope  $U_n$  is defined in [49] for  $n \geq 3$ . The vertex set of  $U_n$  is  $V(U_n) = \{a_i, b_i, c_i, d_i, e_i; 1 \leq i \leq n\}$ , and its edge set is

$$E(U_n) = \{a_i \sim a_{i+1}, a_i \sim b_i, b_i \sim b_{i+1}, b_i \sim c_i; 1 \leq i \leq n\} \cup \{c_i \sim d_i, c_i \sim d_{i-1}, d_i \sim e_i, e_i \sim e_{i+1}; 1 \leq i \leq n\}. \quad (51)$$

The convex plane graph  $U_n^p$  ( $p$  for pendant) is obtained from the graph of a convex polytope  $U_n$  by attaching one pendant vertex (leaf)  $f_i$  to the vertex  $e_i$  of  $U_n$  for each  $1 \leq i \leq n$  (see Figure 4) [50]. Thus, we have the followings vertex and edge sets for  $U_n^p$ :

$$V(U_n^p) = V(U_n) \cup \{f_i; 1 \leq i \leq n\}, \quad (52)$$

$$E(U_n^p) = E(U_n) \cup \{e_i \sim f_i; 1 \leq i \leq n\}.$$

The following main result provides the strong metric dimension of  $U_n^p$ .

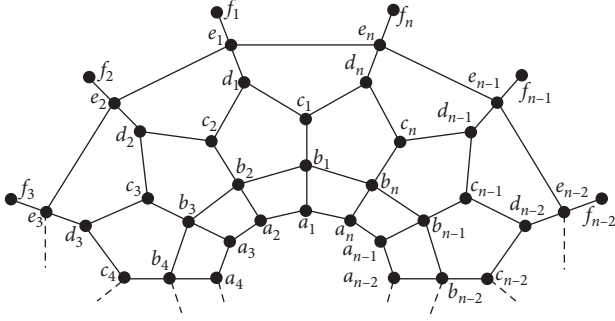

 FIGURE 4: The general half view of the convex plane graph  $U_n^p$ .

 TABLE 5: Vertices willing to perform the geodesic identification in  $U_n^p$ .

Vertex $v$	$N(v)$	$d(w, \cdot)$ for all $w \in N(v)$
$a_i$	$\{a_{i-1}, a_{i+1}, b_i\}$	$d(w, f_{i+k}) = k + 4$
$c_i$	$\{b_i, d_{i-1}, d_i\}$	$d(w, d_{i+k}) = k + 2$
$d_{i+k}$	$\{c_{i+k}, c_{i+k+1}, e_{i+k}\}$	$d(w, c_i) = k + 2$
$f_{i+k}$	$\{e_{i+k}\}$	$d(w, a_i) = k + 4$

**Theorem 3.** For  $n \geq 3$ , let  $U_n^p$  be a convex plane graph. Then,

$$\text{sdim}(U_n^p) = \begin{cases} 2n, & \text{when } n \text{ is odd,} \\ \frac{5n}{2}, & \text{when } n \text{ is even.} \end{cases} \quad (53)$$

The proof of Theorem 3 will be followed after the next four lemmas.

**Lemma 11.** For odd values of  $n \geq 3$ , if  $S$  is a strong metric generator for  $U_n^p$ , then  $|S| \geq 2n$ .

*Proof.* Let  $n = 2k + 1$  with  $k \geq 1$ , and consider Table 5 for  $1 \leq i \leq n$ .

Note the following points in  $U_n^p$ :

P1:  $L = \{f_i; 1 \leq i \leq n\}$  is the set of leaves.

P2: for each  $1 \leq i \leq n$ , the pairs of vertices  $(a_i, f_{i+k})$  and  $(c_i, d_{i+k})$  satisfy both the conditions in (1) because of Table 5 and since  $d(a_i, f_{i+k}) = k + 5$ ,  $d(c_i, d_{i+k}) = k + 3$ .

P3:  $\text{diam}(U_n^p) = k + 5 = d(a_i, f_{i+k})$  for each  $1 \leq i \leq n$ .

Due to P1 and by Proposition 3, we let  $f_2, f_3, \dots, f_n \in S$  without loss of generality. Then, the remaining pair  $(a_{k+2}, f_1)$  satisfies both P2 and P3. So, we must have either  $a_{k+2} \in S$  or  $f_1 \in S$ , by Propositions 1 and 2. Furthermore, Proposition 1 yields that, from the pair  $(c_i, d_{i+k})$ , either  $c_i \in S$  or  $d_{i+k} \in S$  for all  $1 \leq i \leq n$  due to P2. Hence,  $|S| \geq n - 1 + 1 + n = 2n$ .  $\square$

**Lemma 12.** For even values of  $n \geq 4$ , if  $S$  is a strong metric generator for  $U_n^p$ , then  $|S| \geq (5n/2)$ .

*Proof.* Let  $n = 2k$  with  $k \geq 2$ , and consider Table 6 for  $1 \leq i \leq n$ .

The following points hold in  $U_n^p$ :

 TABLE 6: Vertices willing to perform the geodesic identification in  $U_n^p$ .

Vertex $v$	$N(v)$	$d(w, \cdot)$ for all $w \in N(v)$
$a_i$	$\{a_{i-1}, a_{i+1}, b_i\}$	$d(w, c_{i+k}) = k + 1$ $d(w, f_{i+k-1}) \in \{k + 3, k + 4\}$ $d(w, f_{i+k}) \in \{k + 3, k + 4\}$
$c_i$	$\{b_i, d_{i-1}, d_i\}$	$d(w, c_{i+k}) \in \{k + 1, k + 2\}$ $d(w, d_{i+k}) \in \{k + 1, k + 2\}$
$c_{i+k}$	$\{b_{i+k}, d_{i+k-1}, d_{i+k}\}$	$d(w, a_i) \in \{k + 1, k + 2\}$ $d(w, c_i) \in \{k + 1, k + 2\}$
$d_i$	$\{c_i, c_{i+1}, e_i\}$	$d(w, d_{i+k}) \in \{k + 1, k + 2\}$ $d(w, f_{i+k}) \in \{k + 1, k + 2\}$
$d_{i+k}$	$\{c_{i+k}, c_{i+k+1}, e_{i+k}\}$	$d(w, c_i) \in \{k + 1, k + 2\}$ $d(w, d_i) \in \{k + 1, k + 2\}$
$f_{i+k}$	$\{e_{i+k}\}$	$d(w, a_i) = k + 3$ $d(w, d_i) = k + 1$
$f_{i+k-1}$	$\{e_{i+k-1}\}$	$d(w, a_i) = k + 3$

P1:  $L = \{f_i; 1 \leq i \leq n\}$  is the set of leaves.

P2: for each  $1 \leq i \leq n$ ,  $d(a_i, f_{i+k-1}) = k + 4 = d(a_i, f_{i+k})$  and  $d(a_i, c_{i+k}) = d(c_i, d_{i+k}) = d(d_i, f_{i+k}) = k + 2$ . It follows, by Table 6, that the pairs of vertices  $(a_i, f_{i+k-1})$ ,  $(a_i, f_{i+k})$ ,  $(a_i, c_{i+k})$ ,  $(c_i, d_{i+k})$ , and  $(d_i, f_{i+k})$  satisfy both the conditions in (1) for all  $1 \leq i \leq n$ . Moreover, for each  $1 \leq j \leq k$ ,  $d(c_j, c_{j+k}) = k + 2 = d(d_j, d_{j+k})$ . Therefore, Table 6 again yields that the pairs of vertices  $(c_j, c_{j+k})$  and  $(d_j, d_{j+k})$  satisfy both the conditions in (1) for all  $1 \leq j \leq k$ .

P3:  $\text{diam}(U_n^p) = k + 4 = d(a_i, f_{i+k-1}) = d(a_i, f_{i+k})$  for all  $1 \leq i \leq n$ .

Without loss of generality, taking  $f_2, f_3, \dots, f_n$  in  $S$ , by Proposition 3 and due to P1, we get three pairs of vertices  $(a_{k+2}, f_1)$ ,  $(a_{k+1}, f_1)$ , and  $(d_{k+1}, f_1)$ , which are still satisfying P2 and P3. For these pairs, Propositions 1 and 2 allow us to take either  $a_{k+2} \in S$  or  $a_{k+1} \in S$  or  $f_1 \in S$ . It can be seen that the best suitable choice, to perform the geodesic identification for these three pairs, is to take  $f_1$  in  $S$ . For otherwise, we will get  $|S| \geq 3n$ . Furthermore, the point P3 together with Proposition 2 yields that, from the pairs of vertices  $(a_i, c_{i+k})$  and  $(c_i, d_{i+k})$ , either  $a_i \in S$  or  $c_{i+k} \in S$  and  $c_i \in S$  or  $d_{i+k} \in S$  for all  $1 \leq i \leq n$ . Accordingly, we discuss the following four cases:

- (i) If  $a_i \in S$  and  $c_i \in S$  for all  $1 \leq i \leq n$ , then we must have either  $d_j \in S$  or  $d_{j+k} \in S$  for all  $1 \leq j \leq k$ , by the point P2 and Proposition 1. For otherwise, we get the pair of vertices  $(d_j, d_{j+k})$ , for all  $1 \leq j \leq k$ , which is left unidentified by the elements of  $S$ , but we have  $|S| \geq n - 1 + 1 + n + n + k = (7n/2)$ .
- (ii) If  $a_i \in S$  and  $d_{i+k} \in S$  for all  $1 \leq i \leq n$ , then we must have either  $c_j \in S$  or  $c_{j+k} \in S$  for all  $1 \leq j \leq k$ , by the point P2 and Proposition 1. Otherwise, we get the pair of vertices  $(c_j, c_{j+k})$ , for all  $1 \leq j \leq k$ , which is left unidentified by the elements of  $S$ , but we have  $|S| \geq n - 1 + 1 + n + n + k = (7n/2)$ .
- (iii) If  $c_{i+k} \in S$ , then of course  $c_i \in S$  for all  $1 \leq i \leq n$ . Moreover, we must have either  $d_j \in S$  or  $d_{j+k} \in S$  for

all  $1 \leq j \leq k$ , by the point P2 and Proposition 1. For otherwise, we get the pair of vertices  $(d_j, d_{j+k})$ , for all  $1 \leq j \leq k$ , which is left unidentified by the elements of  $S$ . Then, we have  $|S| \geq n - 1 + 1 + n + k = (5n/2)$ .

- (iv) If  $c_{i+k} \in S$  and  $d_{i+k} \in S$  for all  $1 \leq i \leq n$ , then we have  $|S| \geq n - 1 + 1 + n + n = 3n$ .

It can be concluded from these four cases that the most suitable choice to construct a strong metric generator  $S$  with minimum cardinality is found in the Case 3, which yields that  $|S| \geq (5n/2)$ .  $\square$

**Lemma 13.** *If  $n = 2k + 1$  with  $k \geq 1$ , then the set  $S = \{c_1, c_2, \dots, c_n, f_1, f_2, \dots, f_n\} \subset V(U_n^p)$  is a strong metric generator for  $U_n^p$ .*

*Proof.* It is enough to perform the geodesic identification for those pairs of vertices of  $U_n^p$  having no element from the set  $S$  because every pair of vertices having one element  $s$  from  $S$  possesses the geodesic identification by the element  $s$ . Let us consider the following  $a_i - f_{i+k}$  geodesics of length  $k + 5$ , for each  $1 \leq i \leq n$ :

$$\overrightarrow{P_{1i}}: a_i \sim a_{i+1} \sim a_{i+2} \sim \dots \sim a_{i+k} \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k} \sim f_{i+k}, \quad (54)$$

$$\overrightarrow{P_{2i}}: a_i \sim b_i \sim b_{i+1} \sim b_{i+2} \sim \dots \sim b_{i+k} \sim c_{i+k} \sim d_{i+k} \sim e_{i+k} \sim f_{i+k}, \quad (55)$$

$$\overrightarrow{P_{3i}}: a_i \sim b_i \sim c_i \sim d_i \sim e_i \sim e_{i+1} \sim e_{i+2} \sim \dots \sim e_{i+k} \sim f_{i+k}, \quad (56)$$

$$\overleftarrow{P_{1i}}: a_i \sim a_{i-1} \sim a_{i-2} \sim \dots \sim \{a_{i-k} = a_{i+k+1}\} \sim b_{i+k+1} \sim c_{i+k+1} \sim d_{i+k} \sim e_{i+k} \sim f_{i+k}, \quad (57)$$

$$\overleftarrow{P_{2i}}: a_i \sim b_i \sim b_{i-1} \sim b_{i-2} \sim \dots \sim \{b_{i-k} = b_{i+k+1}\} \sim c_{i+k+1} \sim d_{i+k} \sim e_{i+k} \sim f_{i+k}, \quad (58)$$

$$\overleftarrow{P_{3i}}: a_i \sim b_i \sim c_i \sim d_{i-1} \sim e_{i-1} \sim e_{i-2} \sim \dots \sim \{e_{i-k} = e_{i+k+1}\} \sim e_{i+k} \sim f_{i+k}. \quad (59)$$

The vertex  $f_{i+k}$  performs the geodesic identification for all the pairs of vertices due to geodesics listed from (54) to (59), except the following pairs of vertices:

$$(d_i, d_{i+1}), (d_i, d_{i-1}), \quad \text{for } 1 \leq i \leq n, \quad (60)$$

$$(d_i, d_{i+j}), (d_i, d_{i-j}), \quad \text{for } 1 \leq i \leq n, 2 \leq j \leq k. \quad (61)$$

For  $1 \leq i \leq n, 2 \leq j \leq k$ , the following geodesics ensure the geodesic identification of the pairs, given in (60) and (61), by some vertices of  $S$  as described in Table 7:

$$\overrightarrow{Q_i}: c_i \sim d_i \sim c_{i+1} \sim d_{i+1}, \quad (62)$$

$$\overleftarrow{Q_i}: c_{i+1} \sim d_i \sim c_i \sim d_{i-1}, \quad (63)$$

$$\overrightarrow{R_{ji}}: c_i \sim d_i \sim e_i \sim e_{i+1} \sim e_{i+2} \sim \dots \sim e_{i+j-1} \sim e_{i+j} \sim d_{i+j}, \quad (64)$$

$$\overleftarrow{R_{ji}}: c_{i+1} \sim d_i \sim e_i \sim e_{i-1} \sim e_{i-2} \sim \dots \sim e_{i-(j-1)} \sim e_{i-j} \sim d_{i-j}. \quad (65)$$

Hence, we conclude that  $S$  is a strong metric generator for  $U_n^p$ .  $\square$

**Lemma 14.** *If  $n = 2k$  with  $k \geq 2$ , then the set*

$$S = \{c_1, c_2, \dots, c_n, d_1, d_2, \dots, d_k, f_1, f_2, \dots, f_n\} \subset V(U_n^p) \quad (66)$$

*is a strong metric generator for  $U_n^p$ .*

*Proof.* For  $s \in S$ , a pair of vertices  $(v, s)$  with  $v \in V(U_n^p) - \{s\}$  possesses the geodesic identification by the vertex  $s$ . Therefore, we should perform the geodesic identification for each pair  $(x, y)$  of vertices of  $U_n^p$  by the vertices in  $S$  for both  $x, y \in S$ . For each  $1 \leq i \leq n$ , the following  $a_i - f_{i+k-1}$  and  $a_i - f_{i+k}$  geodesics of length  $k + 4$  are useful for our purpose:

$$\overrightarrow{P_{1i}}: a_i \sim a_{i+1} \sim a_{i+2} \sim \dots \sim a_{i+k-1} \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1} \sim f_{i+k-1}, \quad (67)$$

$$\overrightarrow{P_{2i}}: a_i \sim b_i \sim b_{i+1} \sim b_{i+2} \sim \dots \sim b_{i+k-1} \sim c_{i+k-1} \sim d_{i+k-1} \sim e_{i+k-1} \sim f_{i+k-1}, \quad (68)$$

$$\overrightarrow{P_{3i}}: a_i \sim b_i \sim c_i \sim d_i \sim e_i \sim e_{i+1} \sim e_{i+2} \sim \dots \sim e_{i+k-1} \sim f_{i+k-1}, \quad (69)$$

TABLE 7: Geodesic identification for the pairs given in (60) and (61).

Pairs	Vertex performing the geodesic identification	Due to the geodesic in
$(d_i, d_{i+1})$ $1 \leq i \leq n$	$c_i$	(62)
$(d_i, d_{i-1})$ $1 \leq i \leq n$	$c_{i+1}$	(63)
$(d_i, d_{i+j})$ $1 \leq i \leq n$ $2 \leq j \leq k$	$c_i$	(64)
$(d_i, d_{i-j})$ $1 \leq i \leq n$ $2 \leq j \leq k$	$c_{i+1}$	(65)

TABLE 8: Geodesic identification for the pairs given in (73)–(76).

Pairs	Vertex performing the geodesic identification	Due to the geodesic in
$(a_i, a_{i+k}), (a_i, b_{i+k})$ $1 \leq i \leq k$	$c_{i+k}$	(77)
$(b_i, b_{i+k})$ $1 \leq i \leq k$	$c_{i+k}$	(78)
$(e_j, d_{j+k}), (e_j, e_{j+k})$ $1 \leq j \leq k$	$d_j$	(79)
$(d_r, d_{r+1})$ $k+1 \leq r \leq n-1$	$c_r$	(80)
$(d_l, d_{l+m})$ $k+1 \leq l \leq n-2$ $2 \leq m \leq k-1$	$c_l$	(81)

$$\overleftarrow{P}_{1i} : a_i \sim a_{i-1} \sim a_{i-2} \sim \cdots \sim a_{i-(k-1)} \sim b_{i-(k-1)} \sim c_{i-(k-1)} \sim \{d_{i-k} = d_{i+k}\} \sim e_{i+k} \sim f_{i+k}, \quad (70)$$

$$\overleftarrow{P}_{2i} : a_i \sim b_i \sim b_{i-1} \sim b_{i-2} \sim \cdots \sim b_{i-(k-1)} \sim c_{i-(k-1)} \sim \{d_{i-k} = d_{i+k}\} \sim e_{i+k} \sim f_{i+k}, \quad (71)$$

$$\overleftarrow{P}_{3i} : a_i \sim b_i \sim c_i \sim d_{i-1} \sim e_{i-1} \sim e_{i-2} \sim \cdots \sim e_{i-(k-1)} \sim \{e_{i-k} = e_{i+k}\} \sim f_{i+k}. \quad (72)$$

$$\overrightarrow{Q}_{1i} : a_i \sim a_{i+1} \sim a_{i+2} \sim \cdots \sim a_{i+k} \sim b_{i+k} \sim c_{i+k}, \quad (77)$$

$$\overrightarrow{Q}_{2i} : a_i \sim b_i \sim b_{i+1} \sim b_{i+2} \sim \cdots \sim b_{i+k} \sim c_{i+k}, \quad (78)$$

$$\overrightarrow{R}_j : d_j \sim e_j \sim e_{j+1} \sim e_{j+2} \sim \cdots \sim e_{j+k-1} \sim e_{j+k} \sim d_{j+k}, \quad (79)$$

$$\overrightarrow{S}_j : c_r \sim d_r \sim c_{r+1} \sim d_{r+1}, \quad (80)$$

$$\overrightarrow{T}_{ml} : c_l \sim d_l \sim e_l \sim e_{l+1} \sim e_{l+2} \sim \cdots \sim e_{l+m-1} \sim e_{l+m} \sim d_{l+m}. \quad (81)$$

Thus,  $S$  is a strong metric generator for  $U_n^p$ .  $\square$

It can be seen that, for some pairs of vertices, the vertex  $f_{i+k-1}$  performs the geodesic identification due to geodesics in (67)–(69), and for some pairs of vertices, the vertex  $f_{i+k}$  performs the geodesic identification due to geodesics in (70)–(72). The pairs of vertices, which are left unidentified, are as follows:

$$(a_i, a_{i+k}), (a_i, b_{i+k}), (b_i, b_{i+k}), \quad \text{for } 1 \leq i \leq k, \quad (73)$$

$$(e_j, d_{j+k}), (e_j, e_{j+k}), \quad \text{for } 1 \leq j \leq k, \quad (74)$$

$$(d_r, d_{r+1}), \quad \text{for } k+1 \leq r \leq n-1, \quad (75)$$

$$(d_l, d_{l+m}), \quad \text{for } k+1 \leq l \leq n-2, 2 \leq m \leq k-1. \quad (76)$$

All these pairs of vertices possess the geodesic identification, as described in Table 8, with the help of the following geodesics, for  $1 \leq i \leq k$ ,  $1 \leq j \leq k$ ,  $k+1 \leq r \leq n-1$ ,  $k+1 \leq l \leq n-2$ , and  $2 \leq m \leq k-1$ :

*Proof* (Theorem 3). The lower bound for the strong metric dimension of  $U_n^p$  is provided by Lemmas 11 and 12, whereas the upper bound is due to Lemmas 13 and 14.  $\square$

## 5. Concluding Remarks

The geodesic identification (strong resolvability) for each pair of vertices has been performed in three families  $S_n^p$ ,  $T_n^p$ , and  $U_n^p$  of convex plane graphs for all  $n \geq 3$ . In fact, we extended the study of the identification of vertices by solving the strong metric dimension problem for three more families of graphs. The problem is solved with the following investigations:

- (i) In  $S_n^p$ , total  $\binom{5n}{2}$  pairs of vertices are identified with  $n$  vertices by using  $8n$  geodesics when  $n$  is odd and with  $(3n/2)$  vertices by using  $8n + 1$  geodesics when  $n$  is even.
- (ii) In  $T_n^p$ , total  $\binom{5n}{2}$  pairs of vertices are identified with  $2n$  vertices by using  $14n$  geodesics when  $n$  is odd and with  $(5n/2)$  vertices by using  $(19n/2)$  geodesics when  $n$  is even.
- (iii) In  $U_n^p$ , total  $\binom{6n}{2}$  pairs of vertices are identified with  $2n$  vertices by using  $n(n + 3)$  geodesics when  $n$  is odd and with  $(5n/2)$  vertices by using  $(n^2/4) + 6n + 3$  geodesics when  $n$  is even.

## Data Availability

The data used to support the findings of this study are available from the corresponding author.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under grant no. RG-19-135-38. The authors therefore acknowledge with thanks DSR for technical and financial support.

## References

- [1] Z. Beerliova, F. Eberhard, T. Erlebach et al., "Network discovery and verification," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 12, pp. 2168–2181, 2006.
- [2] G. Chartrand, L. Eroh, M. A. Johnson, and O. R. Oellermann, "Resolvability in graphs and the metric dimension of a graph," *Discrete Applied Mathematics*, vol. 105, no. 1–3, pp. 99–113, 2000.
- [3] V. Chvátal, "Mastermind," *Combinatorica*, vol. 3, no. 3–4, pp. 325–329, 1983.
- [4] S. Khuller, B. Raghavachari, and A. Rosenfeld, "Landmarks in graphs," *Discrete Applied Mathematics*, vol. 70, no. 3, pp. 217–229, 1996.
- [5] A. Sebö and E. Tannier, "On metric generators of graphs," *Mathematics of Operations Research*, vol. 29, pp. 383–393, 2004.
- [6] H. Shapiro and S. Soderberg, "A combinatory detection problem," *American Mathematical Monthly*, vol. 70, pp. 1066–1070, 1963.
- [7] F. Harary and R. A. Melter, "On the metric dimension of a graph," *Ars Combinatoria*, vol. 2, pp. 191–195, 1976.
- [8] P. J. Slater, "Leaves of trees," *Congressus Numerantium*, vol. 14, pp. 549–559, 1975.
- [9] S. Arumugam and V. Mathew, "The fractional metric dimension of graphs," *Discrete Mathematics*, vol. 312, no. 9, pp. 1584–1590, 2012.
- [10] R. C. Brigham, G. Chartrand, R. D. Dutton, and P. Zhang, "Resolving domination in graphs," *Mathematica Bohemica*, vol. 128, no. 1, pp. 25–36, 2003.
- [11] J. Cáceres, C. Hernando, M. Mora et al., "On the metric dimension of cartesian products of graphs," *SIAM Journal on Discrete Mathematics*, vol. 21, no. 2, pp. 423–441, 2007.
- [12] G. Chartrand, V. Saenpholphat, and P. Zhang, "The independent resolving number of a graph," *Mathematica Bohemica*, vol. 128, no. 4, pp. 379–393, 2003.
- [13] L. Epstein, A. Levin, and G. J. Woeginger, "The (weighted) metric dimension of graphs: hard and easy cases," *Algorithmica*, vol. 72, no. 4, pp. 1130–1171, 2015.
- [14] A. Estrada-Moreno, J. A. Rodríguez-Velázquez, and I. G. Yero, "The  $k$ -metric dimension of a graph," *Applied Mathematics*, vol. 9, no. 6, pp. 2829–2840, 2015.
- [15] A. Hakanen, V. Junnila, and T. Laihonen, "The solid-metric dimension," *Theoretical Computer Science*, vol. 806, pp. 156–170, 2020.
- [16] A. Kelenc, D. Kuziak, A. Taranenko, and I. G. Yero, "Mixed metric dimension of graphs," *Applied Mathematics and Computation*, vol. 314, pp. 429–438, 2017.
- [17] F. Okamoto, B. Phinezy, and P. Zhang, "The local metric dimension of a graph," *Mathematica Bohemica*, vol. 135, no. 3, pp. 239–255, 2010.
- [18] Y. Ramírez-Cruz, O. R. Oellermann, and J. A. Rodríguez-Velázquez, "Simultaneous resolvability in graph families," *Electronic Notes in Discrete Mathematics*, vol. 46, pp. 241–248, 2014.
- [19] V. Saenpholphat and P. Zhang, "Connected resolving sets in graphs," *Ars Combinatoria*, vol. 68, pp. 3–16, 2003.
- [20] E. Estaji and J. A. Rodríguez-Velázquez, "The strong metric dimension of generalized Sierpiński graphs with pendant vertices," *Ars Mathematica Contemporanea*, vol. 12, no. 1, pp. 127–134, 2017.
- [21] J. Kratica, V. Kovacevic-Vujcic, M. Cangalovic, and M. Stojanovic, "Minimal doubly resolving sets and the strong metric dimension of Hamming graphs," *Applicable Analysis and Discrete Mathematics*, vol. 6, no. 1, pp. 63–71, 2012.
- [22] J. Kratica, V. Kovačević-Vučić, M. Čangalović, and M. Stojanović, "Minimal doubly resolving sets and the strong metric dimension of some convex polytopes," *Applied Mathematics and Computation*, vol. 218, no. 19, pp. 9790–9801, 2012.
- [23] M. Salman, I. Javaid, and M. A. Cahudhry, "Minimum fault-tolerant, local and strong metric dimension of graphs," *Ars Combinatoria*, vol. 138, pp. 333–353, 2018.
- [24] T. A. Kusmayadi, S. Kuntari, D. Rahmadi, and F. A. Lathifah, "On the strong metric dimension of some related wheel graph," *Far East Journal of Mathematical Sciences (FJMS)*, vol. 99, no. 9, pp. 1325–1334, 2016.
- [25] D. Kuziak, M. L. Puertas, J. A. Rodríguez-Velázquez, and I. G. Yero, "Strong resolving graphs: the realization and the



- characterization problems,” *Discrete Applied Mathematics*, vol. 236, pp. 270–287, 2018.
- [26] D. Kuziak, I. Peterin, and I. G. Yero, “Resolvability and strong resolvability in the direct product of graphs,” *Results in Mathematics*, vol. 71, no. 1-2, pp. 509–526, 2017.
- [27] D. Kuziak, I. G. Yero, and J. A. Rodríguez-Velázquez, “On the strong metric dimension of Cartesian sum graphs,” *Fundamenta Informaticae*, vol. 141, no. 1, pp. 57–69, 2015.
- [28] X. Ma, M. Feng, and K. Wang, “The strong metric dimension of the power graph of a finite group,” *Discrete Applied Mathematics*, vol. 239, pp. 159–164, 2018.
- [29] T. R. May and O. R. Oellermann, “The strong dimension of distance-hereditary graphs,” *Journal of Combinatorial Mathematics and Combinatorial Computing*, vol. 76, pp. 59–73, 2011.
- [30] Y. R. Mayasari and T. A. Kusmayadi, “On the strong metric dimension of generalized butterfly graph, starbarbell graph, and  $C_m \odot P_n$  graph,” *Journal of Physics: Conference Series*, vol. 1008, Article ID 012048, 2018.
- [31] Y. Mintarsih and T. A. Kusmayadi, “On the strong metric dimension of antiprism graph, king graph, and  $K_m \odot K_n$  graph,” *Journal of Physics: Conference Series*, vol. 1008, Article ID 012051, 2018.
- [32] M. Widyaningrum and T. A. Kusmayadi, “On the strong metric dimension of sun graph, windmill graph, and Möbius ladder graph,” *Journal of Physics: Conference Series*, vol. 1008, Article ID 012032, 2018.
- [33] A. Yunitasari and T. A. Kusmayadi, “On the strong metric dimension of crossed prism graph and edge corona of cycle with path graph,” *Journal of Physics: Conference Series*, vol. 1306, Article ID 012018, 2019.
- [34] D. Kuziak, I. G. Yero, and J. A. Rodríguez-Velázquez, “Strong metric dimension of rooted product graphs,” *International Journal of Computer Mathematics*, vol. 93, no. 8, pp. 1265–1280, 2016.
- [35] D. Kuziak, J. A. Rodríguez-Velázquez, and G. I. Yero, “Closed formulae for the strong metric dimension of lexicographic product graphs,” *Discussiones Mathematicae Graph Theory*, vol. 36, no. 4, pp. 1051–1064, 2016.
- [36] D. Kuziak, I. G. Yero, and J. A. Rodríguez-Velázquez, “On the strong metric dimension of the strong products of graphs,” *Open Mathematics*, vol. 13, pp. 64–74, 2015.
- [37] D. Kuziak, I. G. Yero, and J. A. Rodríguez-Velázquez, “On the strong metric dimension of corona product graphs and join graphs,” *Discrete Applied Mathematics*, vol. 161, no. 7-8, pp. 1022–1027, 2013.
- [38] O. R. Oellermann and J. Peters-Fransen, “The strong metric dimension of graphs and digraphs,” *Discrete Applied Mathematics*, vol. 155, no. 3, pp. 356–364, 2007.
- [39] J. A. Rodríguez-Velázquez, I. G. Yero, D. Kuziak, and O. R. Oellermann, “On the strong metric dimension of Cartesian and direct products of graphs,” *Discrete Mathematics*, vol. 335, pp. 8–19, 2014.
- [40] C. X. Kang and E. Yi, “The fractional strong metric dimension of graphs,” in *Combinatorial Optimization and Applications*, pp. 84–95, Springer, Berlin, Germany, 2013.
- [41] C. X. Kang, I. G. Yero, and E. Yi, “The fractional strong metric dimension in three graph products,” *Discrete Applied Mathematics*, vol. 251, pp. 190–203, 2018.
- [42] C. X. Kang, “On the fractional strong metric dimension of graphs,” *Discrete Applied Mathematics*, vol. 213, pp. 153–161, 2016.
- [43] J. Kratica, V. Kovacevic-Vujcic, and M. Cangalovic, “Computing strong metric dimension of some special classes of graphs by genetic algorithms,” *Yugoslav Journal of Operations Research*, vol. 18, no. 2, pp. 143–151, 2008.
- [44] N. Mladenović, J. Kratica, V. Kovačević-Vučić, and M. Čangalović, “Variable neighborhood search for the strong metric dimension problem,” *Electronic Notes in Discrete Mathematics*, vol. 39, pp. 51–57, 2012.
- [45] E. Yi, “On strong metric dimension of graphs and their complements,” *Acta Mathematica Sinica, English Series*, vol. 29, no. 8, pp. 1479–1492, 2013.
- [46] B. DasGupta and N. Mobasher, “On optimal approximability results for computing the strong metric dimension,” *Discrete Applied Mathematics*, vol. 221, pp. 18–24, 2017.
- [47] J. Kratica, V. Kovacevic-Vujcic, M. Cangalovic, and N. Mladenovic, “Strong metric dimension: a survey,” *Yugoslav Journal of Operations Research*, vol. 24, no. 2, pp. 187–198, 2014.
- [48] D. Kuziak, *Strong resolvability in product graphs*, Ph.D. thesis, Universitat Rovira I Virgili, Tarragona, Spain, 2014.
- [49] M. Imran, S. A. Ul Haq Bokhary, and A. Q. Baig, “On families of convex polytopes with constant metric dimension,” *Computers & Mathematics with Applications*, vol. 60, no. 9, pp. 2629–2638, 2010.
- [50] M. Imran, S. A. Bokhary, A. Q. Baig, and I. Tomescu, “On metric dimension of convex polytopes with pendant edges,” *Ars Combinatoria*, vol. 125, pp. 433–447, 2016.



## Research Article

# Algorithm Design of Early Warning Seatbelt Intelligent Adjustment System Based on Neural Network and Big Data Analysis

Chunxu Zhou 

*Department of Ship and Ocean Engineering, Jiangsu Shipping Vocational and Technical College, Nantong, Jiangsu, China*

Correspondence should be addressed to Chunxu Zhou; [zhoucx@jssc.edu.cn](mailto:zhoucx@jssc.edu.cn)

Received 13 August 2020; Revised 19 October 2020; Accepted 8 November 2020; Published 21 November 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Chunxu Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the purpose to guarantee the safety of drivers and passengers as well as lower the death rate collision, the early warning seatbelt intelligent adjustment system is designed by using big data analysis technology based on the aspects of hardware equipment, database, and software program. In the hardware system, microcontroller AT89C52 is applied as the control core. By means of the sensor detection and drive control, the early warning safety belt tightening, locking and lifting, and other functions are realized. Meanwhile, various components of the hardware system are coordinated through debugging several modules in the hardware system and using the modified circuit to connect them together. We determine the relational rules of the database and create the corresponding database table, to provide sufficient data support for the realization of software functions. Using the big data analysis method to process the real-time detection data received by the sensor, the software functions such as timely tightening of safety belt, humidity relaxation, and over-rolling prevention can be realized according to different driving conditions of drivers and vehicles, respectively. The conclusion is drawn through the system test experiment: compared with the traditional regulation system, the design system has a higher degree of regulation, and the application of the design results to the actual vehicle can reduce the crash fatality rate of about 22.4%.

## 1. Introduction

Traffic safety has always been a hot issue in people's daily life, so the safety insurance intelligent devices such as ABS and airbags are mounted in automobiles during recent years. Thanks to the greatly improved active and passive safety of automobiles, the number of deaths and death rates of car accidents in many developed countries have decreased despite the increasing number of cars. This shows that advanced safety technology can obtain obvious application effects. In order to ensure that the early warning seatbelt can maximize its protection role in car accident, a corresponding intelligent adjustment system shall be established, whose function is to control the working state and force limitation of the early warning seatbelt and then form a force buffer for the driver. Based on the background that research on the early warning

seatbelt intelligent adjustment system is in its initial stage without achieving any breakthrough results, this paper designs the corresponding early warning seatbelt intelligent adjustment system through big data analysis [1]. Characterized by large amount of data, fast speed, multiple types, high value, and strong authenticity, big data analysis refers to the analysis of large-scale data [2]. In this system design, big data analysis technology is mainly applied to the database and software functions as a method for processing and analyzing the initial collected data [3], so that the efficiency of system regulation function is enhanced. The design of intelligent adjustment system not only enables the traditional early warning seatbelt to play its role in the event of an accident but also can fully protect the driver and passengers during the entire driving process, thereby improving the safety factor of automobile [4].

## 2. Hardware of Early Warning Seatbelt Intelligent Adjustment System

The hardware of early warning seatbelt intelligent adjustment system can be divided into four parts: control module, data acquisition module, drive module, and terminal module. These four modules are connected together after modification [5]. The structural connection and distribution of the hardware system are shown in Figure 1.

**2.1. Circuit Module.** The circuit module is to connect the hardware devices in the hardware system and provide stable voltage for each module so as to ensure the normal operation of hardware system [6].

**2.1.1. Power Circuit.** The MC9S12XS128 chip, as the core component of power circuit, can realize direct power supply of 12 V automotive DC power supplies through connecting the fuse F1 to the power supply chip. By virtue of the working peak value at 750 mA, this chip can improve the power circuit better than traditional hardware system and avoid the loss of the whole hardware system circuit caused by circuit load. It is also necessary to insert a conversion chip in the power circuit to meet the voltage requirements of different hardware modules. The selected conversion chip is LM2940. In addition, a capacitor appliance, which includes two resistance devices, shall be installed in the power circuit to stabilize and decompose the voltage.

**2.1.2. Step-Down Conversion Circuit.** In general, the output voltage of automobile power circuit is 24 V, which can be adjusted via a conversion chip. Such chip can only convert the output voltage into one voltage, but the voltage required by different devices in the hardware system is not fixed, as shown in Table 1.

The step-down conversion circuit can be realized in two steps. First, the 24 V output by the power supply circuit is converted into a 5 V, and then it is converted into 3.3 V. The process of realizing step-down conversion is shown as follows:

$$V_{\text{Out}} = V_{\text{Ref}} \times \left(1 + \frac{R_1}{R_2}\right), \quad (1)$$

where parameter  $V_{\text{Ref}}$  represents the reference voltage of the step-down conversion circuit and  $R_1$  and  $R_2$ , respectively, represent the resistance of the step-down conversion circuit.

**2.2. Data Acquisition Module.** The core of data acquisition module is the sensor [7]. In this design, 5 types of sensors are installed, including vehicle speed sensor, webbing pull-out sensor, seating detection sensor, buckle sensor, and seat position sensors. A total of 10 data acquisition sensors are installed in the designated position with two in each type [8].

**2.3. Mathematical Model of the Drive Module.** The permanent magnet DC motor coming with a one-stage worm gear reduction mechanism is selected to provide power for the

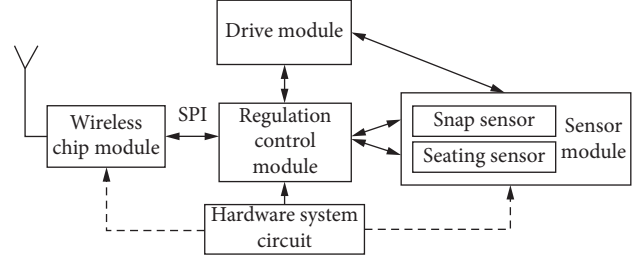


FIGURE 1: System hardware composition structure diagram.

TABLE 1: Hardware system demand voltage analysis.

Chip name	Model selection	Demand voltage
Seat node main chip	STM8S208R8	+5.0 V
Wireless transmission chip	SI4432	+1.8 V~+3.6 V
Seat inspection chip	MC33794	+5.0 V
Controller main core	STM32F103	+5.0 V

hardware system. After deceleration, the output speed of this motor is about 50 r/min, the rated power is 40 W, and the output torque can reach Nm. According to Newton's second law, the basic working principle of the drive module can be concluded as follows:

$$\begin{cases} u_a(t) = R_a i_a(t) + L_a \frac{di_a(t)}{dt} + E_a, \\ T_e = J \frac{d\omega}{dt} + B\omega, \\ E_a = K_e \omega, \\ T_e = K_T I_a, \end{cases} \quad (2)$$

where  $u_a$  is the input voltage of the drive motor,  $R_a$  is the resistance value inside the drive motor,  $L_a$  is the armature inductance,  $E_a$  is the induced electromotive force,  $T_e$  is the electromagnetic torque of the drive DC motor,  $J$  is the moment of inertia,  $B$  is the damping coefficient,  $i$  is the DC current in the motor,  $\omega$  is the output angle of the motor, and the parameters  $K_T$  and  $K_e$ , respectively, represent the torque creation and the induced electromotive force constant of the driving motor. Through the transformation and solution of formula (2), the transfer function of the DC drive motor can be obtained as

$$G_U(s) = \frac{K_T}{L_a J s^2 + (L_a B + R_a J)s + K_e K_T + R_a B}. \quad (3)$$

The devices in the entire hardware system can work coordinately by taking the motor transfer function shown in formula (3) as the working principle of the drive module.

**2.4. Adjusting Module.** The AT89C52 chip is applied as single-chip microcomputer in the adjusting module, which has an 8 bit data width processor and a total of 256 RAM units. In the adjusting module, the adjustment control

realization devices, such as pretightening unit, force limiter, and control node cluster coordinator, are connected to the single-chip microcomputer to realize the coordinated work between the hardware systems. For the sake of effectively realizing the pretightening function, a steel ball-type pretightening period is selected. Since the steel ball in the pretightening unit will move at a high speed under the action of explosive force, a sealing baffle shall be added and fixed to the steel ball collector with screws to prevent steel ball from bouncing out of the pretightening unit.

**2.5. Connection and Debugging of Hardware Devices.** In accordance with the method shown in Figure 1, all the components in the hardware system are connected for debugging. At first, the debugging process of hardware equipment needs to test the bare board, that is, carefully check the connection of printed circuit board according to the circuit principle. In this process, special attention should be paid to the power system to prevent the breakage of power supply and wrong polarity. Secondly, the break test function of the digital multimeter is used to test whether all the circuits on the bare board have circuit phenomena. After confirming the normal state of bare board, each component module unit in the hardware system is connected to the hardware motherboard. During this process, it is necessary to observe whether there is a missing device, whether the interface is firmly connected, whether the reverse polarity capacitor is distinguished, and whether the diode device is connected in the wrong direction. In case of an abnormal phenomenon, it is needed to analyze the reason, eliminate the fault, and test repeatedly until meeting the test requirements of the hardware system. The final test is a power-on operation test, that is, the power port is started in the hardware system to observe whether each hardware device can operate normally under the power-on environment. Before the power-on test, it is necessary to not only check whether the assignment of the external power supply voltage is at the required DIYYUAN voltage value but also repeatedly determine whether the polarity direction is correct, so as to avoid damage to device components caused by the short circuit at the power terminal.

### 3. Big Data Algorithm Design of Early Warning Seatbelt Intelligent Adjustment System

The prewarning seatbelt intelligent adjustment system database is set to adjust the system's mathematical algorithm [9]. The data stored in database include the vehicle parameter data, the basic parameter data of the seatbelt, the real-time seatbelt usage data, and the historical accident data [10]. The basic structure of database can be divided into the physical data layer, conceptual data layer, and user data layer [11]. In order to provide sufficient data support for the realization of system algorithm function, the real-time data collected by the sensor is processed in a unified manner and stored in the database in a fixed format [12]. Since the database table is the main unit of database, a database table firstly needs to be built through the sensor data, and then the

fixed database relationship rules shall be specified [13]. In this way, the database of early warning seatbelt intelligent adjustment system can be established.

All the information about seatbelt wearing status, such as the vehicle operating speed and the warning parameters of the seatbelt in different collision accidents, are saved in the database of automobile seatbelt intelligent adjustment system. In order to give more convenience to system users when querying the operating status of car seatbelt instantly, it is needed to not only strictly define the name of each field and the corresponding data type during the database design but also ensure that the parameters in the database are not defined in duplication. The information data  $\alpha_n$  that can better explain the parameter definition is selected in the specified database, and a weight calculation is performed on the data type formed by its secondary data information [14]:

$$C_T = \sum_{t=1}^T an(C_t - D_{st}). \quad (4)$$

In the formula,  $C_t$  represents the data level and  $D_{st}$  represents the information flow direction threshold. According to this, the effective definition under the parameter definition data constraint mechanism is as follows:

$$P_{PV}(t) = \max\{C_T, \eta_{PV}\}. \quad (5)$$

In the formula,  $\eta_{PV}$  represents the extreme value of constraint to maximize the output result of constraint effect and realize the definition of database data information. Based on the basic design principles of database tables, version SQL2016 is used for establishing independent database tables. The database table of automobile seatbelt information is shown in Table 2.

In the same way, database tables of vehicle history information and seatbelt automatic adjustment parameters can be built up.

## 4. Design of Mathematical Algorithm Model for Early Warning Intelligent Seatbelt Adjustment

**4.1. Real-Time Monitoring of Seatbelt Working Status.** The real-time monitoring of seatbelt working status can be divided into three steps. First, the real-time data of early warning seatbelt is collected; second, the big data analysis is adopted to analyze the collected parameter data; and finally, the work of seatbelt is judged based on the big data analysis results. Depending on the sensor installed at the seatbelt position, the collected real-time data of early warning seatbelt mainly include the pull-out length of the webbing and the limiting force. Big data analysis contains six basic aspects, namely, visual analysis, data mining, predictive analysis, semantic engine, data quality and data management, data storage, and data warehouse. Based on the data stored in the database, the working mode of seatbelt can be determined by following the steps of big data analysis. The working mode of early warning seatbelt can be roughly divided into nonworking, safe and dangerous states, which is

TABLE 2: Vehicle seat belt information.

Data parameter definition	Field meaning	Type of data
ID	Primary key	Number
Car_Route	Route	Text
Car_License	License plate	Text
Car_Driver	Driver	Text
Car_Seat	Number of seats	Number
Car_Seat_Number	Seat number	Number
..	..	..
Car_Seat_Number	Seat number	Number

judged according to the basic physical data of the driver and the vehicle speed.

**4.2. Design and Realization of Intelligent Adjustment Function.** The intelligent adjustment function is realized to ensure that the early warning seatbelt is always in a safe working status. However, during the vehicle driving, the safety mode parameters of seatbelt vary as the change of vehicle speed, so the early warning seatbelt needs to adjust its pretightening strength according to the actual operating conditions of the vehicle. In addition, when a vehicle crashes, the smart seat belt system will automatically adjust different safety modes, so as to maximize the personal safety of the driver and other passengers in the vehicle. In order to effectively realize the system intelligent adjustment function, seatbelt working modes are adjusted intelligently directing at the initial state of vehicle, the normal driving state, the state of a collision accident, and the state after collision accident.

**4.2.1. Design of Early Warning Reminder.** After the car engine is ignited, the control module in the hardware system switches to the working state and reminds the device to turn on. In this process, the sensor device detects the gravity of the seat, the speed signal, and the seatbelt signal step by step. If the gravity on the seat is greater than 0 and the seatbelt signal is 0, the early warning reminder program will be started immediately. If it is detected that the driver does not buckle up, the warning light flashes immediately, and the buzzer is activated to give a 1 min low-frequency and gentle buzzer, and then a 2 min high-frequency buzzer. If the driver still has not buckled up, the buzzer stops working, the indicator light continues, and the car speed is adjusted below 10 m/s by the controller. Once buckling up, the sensor status is adjusted, and the path between the power supply and the early warning device is disconnected to stop the early warning.

**4.2.2. Design of Initial Pretightening Adjustment.** The degree of seatbelt usage at the initial speed is determined according to people's body shape and weight, and then the initial pretightening force is set based on manual pull-out of the seatbelt. The pull-out amount of seatbelt is determined by the corresponding sensor and calculated as follows:

$$S_{\min} = T_{\min} \times V_{\max}, \quad (6)$$

where parameter  $T_{\min}$  represents the time that takes for the passenger to pull out the seatbelt and  $V_{\max}$  is the speed when pulling out the seatbelt. In this situation, the pretightening force of the seatbelt can be calculated as follows:

$$F_0 = PS_{\min}, \quad (7)$$

where  $F_0$  is the initial pretightening force and  $P$  is the initial preload coefficient.

**4.2.3. Design of Loading Intensity Adjustment.** The restriction force of seatbelt is related to the driving speed. Assuming that the real-time vehicle speed detected by the vehicle speed sensor is  $v$ , the pretightening restriction force of the prewarning seatbelt at the speed of  $v_x$  should be

$$F = F_0 + \eta(v_x - v), \quad (8)$$

where parameter  $\eta$  represents the coefficient of loading intensity, whose value is a constant determined by the vehicle model and scale. Next, it is needed to check whether the current seatbelt pretightening force reaches  $F$ . If it reaches  $F$ , keep the current status and continue to run; otherwise, the relevant controller shall be used to adjust the current pretightening force to  $F$ .

**4.2.4. Design of Intelligent Locking Adjustment.** Intelligent locking adjustment is to immediately adjust the seatbelt to the maximum pretightening state in the event of a traffic collision. If the starting angular acceleration of the intelligent lock adjustment controller is set at  $\varepsilon_0$  and the linear acceleration is set at  $a_0$ , then

$$M_0 = F \cdot L. \quad (9)$$

The initial working torque for controlling it can be calculated by formula (9). The parameter in the above formula represents the controlled length; then, the starting angular velocity can be calculated as follows:

$$\varepsilon_0 = \frac{M_0}{I}, \quad (10)$$

where  $I$  is the output current of controller. By substituting the result of formula (10) into formula (11), the pretightening acceleration of the seatbelt in the event of a traffic collision can be obtained.

$$a_0 = \varepsilon_0 \times S. \quad (11)$$

The calculation finds that, when a collision accident occurs, the early warning seatbelt is immediately adjusted to the maximum pretension state at acceleration  $a_0$ .

## 5. System Big Data Test

This system test is to verify the relevant performance and application value of the early warning seatbelt intelligent adjustment system designed based on big data analysis [15]. To this end, a car with intact seatbelt facilities is selected as the experimental environment [16]. The experiment is performed by dividing into two parts. One is the realization

TABLE 3: System regulation function test results.

	Security level A (mm)	Security level B (mm)	Security level C (mm)	Average value (mm)
Small passenger seatbelt tightening amount	38.6	38.4	37.9	38.3
Moderate occupant seatbelt tightening	48.1	47.2	45.6	46.9
Tightening of seat belts for tall passengers	57.2	55.9	53.8	55.7



effect of the system function is verified, which requires recruiting volunteers with large differences in body size. The other is to test whether the applied system can improve the survival rate of passengers in traffic collision accidents. In consideration of the high risk during test process, dummy with the same parameters is made as the research object. With the purpose to ensure the theoretical property of the experimental results, the traditional early warning seatbelt intelligent adjustment system is selected as a reference, so that these two systems are tested under the condition of same application environment and participants [17].

**5.1. Adjusting Function Test.** The volunteers are divided into three categories based on their heights, which are short passengers, medium passengers, and tall passengers. The short passengers indicate volunteers at height between 160 cm and 165 cm and weight about 50 kg; the medium passengers include volunteers at height between 170 cm and 175 cm and weight about 65 kg; the tall passengers are those at height between 180 and 190 cm and weight about 80 kg. In addition, three safety levels are set in the experiment, including safety level A (the vehicle speed is around 15 m/s), safety level B (the vehicle speed at 25 m/s), and safety level C (the vehicle speed at 35 m/s). At the same time, the key positions of seatbelt are marked to help observe the adjustment degree of early warning seatbelt under different security levels. In this way, the adjustment effect of seatbelt in one volunteer under three safety levels is obtained from a single experiment.

By means of the same method, the adjustment effect in three volunteers is tested, and the measurement equipment is used to quantify the test results. The results of adjustment test obtained through statistics and calculations are shown in Table 3.

Statistics obtains that the average tightening amount of the seatbelt in the designed system and traditional system separately is 46.9 mm and 38.4 mm, which proves that the designed system has a higher adjustment effect than the traditional system.

## 6. Conclusions

Nowadays, traffic safety has become a topic of concern around the world due to the increasing number of motor vehicles. However, since existing passive safety technology cannot adapt to the current traffic safety situation, it is quite important to study the early warning seatbelt intelligent adjustment system based on big data analysis in the field of active safety. Limited by the time and related conditions, there are still some shortcomings in this research, which remain to be further improved in the future.

## Data Availability

All data included in this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## Acknowledgments

This work was supported by the School-Level Key Project of Nantong Shipping Vocational and Technical College (HYKY/2019A02) and 2020 Jiangsu Province University “Blue Project” Outstanding Young Backbone Teacher Training Project.

## References

- [1] L.-H. Wang, X.-P. Zhao, J.-X. Wu, Y.-Y. Xie, and Y.-H. Zhang, “Motor fault diagnosis based on short-time fourier transform and convolutional neural network,” *Chinese Journal of Mechanical Engineering*, vol. 30, no. 6, pp. 1357–1368, 2017.
- [2] D. Lechevalier, S. Hudak, R. Ak, Y. T. Lee, and S. Fofou, “A neural network meta-model and its application for manufacturing,” in *Proceedings of the 2015 IEEE international conference on big data (big data)*, pp. 1428–1435, IEEE, Santa Clara, CA, USA, November 2015.
- [3] M. Ambrosin, P. Braca, M. Conti, and R. Lazeretti, “Odin,” *ACM Transactions on Internet Technology*, vol. 18, no. 1, pp. 1–22, 2017.
- [4] M. Belenki, “Service-to-device mapping for smart items using a genetic algorithm,” U.S. Patent 7,890,568, 2011.
- [5] H.-S. Kim and P. N. Roschke, “Design of fuzzy logic controller for smart base isolation system using genetic algorithm,” *Engineering Structures*, vol. 28, no. 1, pp. 84–96, 2006.
- [6] A. Fonseca and B. Cabral, “Prototyping a GPGPU neural network for deep-learning big data analysis,” *Big Data Research*, vol. 8, pp. 50–56, 2017.
- [7] J. F. Cao, H. Y. Cui, H. Shi, and L. J. Jiao, “Big data: a parallel particle swarm optimization-back-propagation neural network algorithm based on MapReduce,” *PLoS One*, vol. 11, no. 6, Article ID e0157551, 2016.
- [8] S.-H. Jung, J.-C. Kim, and C.-B. Sim, “Prediction data processing scheme using an artificial neural network and data clustering for big data,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, no. 1, p. 330, 2016.
- [9] L. Yu, S. Wang, and K. K. Lai, “An integrated data preparation scheme for neural network data analysis,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 2, pp. 217–230, 2006.
- [10] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [11] A. Botalb, M. Moinuddin, U. M. Al-Saggaf, and S. A. Ali, “Contrasting convolutional neural network (CNN) with multi-layer perceptron (MLP) for big data analysis,” in *Proceedings of the 2018 International Conference on Intelligent and Advanced System (ICIAS)*, pp. 1–5, IEEE, Kuala Lumpur, Malaysia, August 2018.
- [12] W. Aziguli, Y. Y. Zhang, Y. Xie, and H. Zhang, “A robust text classifier based on denoising deep neural network in the analysis of big data,” *Scientific Programming*, vol. 2017, Article ID 3610378, 10 pages, 2017.
- [13] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, “An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3137–3147, 2016.
- [14] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, “Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2106–2116, 2017.



- [15] A. Gupta, H. K. Thakur, R. Shrivastava, P. Kumar, and S. Nag, "A big data analysis framework using Apache spark and deep learning," in *Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 9–16, IEEE, New Orleans, LA, USA, November 2017.
- [16] Y. Liu, "Incomplete big data imputation mining algorithm based on BP neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 4, pp. 4457–4466, 2019.
- [17] P. Li, Z. K. Chen, L. T. Yang, Q. C. Zhang, and M. J. Deen, "Deep convolutional computation model for feature learning on big data in internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 790–798, 2018.

## Retraction

# Retracted: Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining

### Mathematical Problems in Engineering

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their

agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y.-H. Xu, L.-F. Huang, R.-R. Guo, X.-Y. Zhang, and J.-M. Zhu, "Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining," *Mathematical Problems in Engineering*, vol. 2020, Article ID 9345901, 10 pages, 2020.

## Research Article

# Metrological Analysis of Online Consumption Evaluation Influence Commodity Marketing Decision Based on Data Mining

Yun-Hua Xu,<sup>1</sup> Lin-Fang Huang ,<sup>2</sup> Rong-Rong Guo ,<sup>3</sup> Xu-Yang Zhang,<sup>4</sup> and Jia-Ming Zhu <sup>4</sup>

<sup>1</sup>School of Business Administration, Anhui University of Finance and Economics, Bengbu 233030, China

<sup>2</sup>School of Economics, Anhui University of Finance and Economics, Bengbu 233030, China

<sup>3</sup>School of Accounting, Anhui University of Finance and Economics, Bengbu 233030, China

<sup>4</sup>School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China

Correspondence should be addressed to Jia-Ming Zhu; zhujm1973@163.com

Received 13 August 2020; Revised 12 September 2020; Accepted 6 November 2020; Published 19 November 2020

Academic Editor: Shaohui Wang

Copyright © 2020 Yun-Hua Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The data of reviews and ratings in the online market can provide guidance for company's production and business activities. In this paper, firstly, we build a BP neural network model to help identify "useful consumer reviews." Then, we use the fuzzy comprehensive evaluation method to identify the most successful and failing goods. Next, we achieve the time series prediction of product reputation by making use of ARIMA model. Finally, we use word segmentation and K-means clustering algorithm to determine whether stars and comments have radiation effects.

## 1. Introduction

Reviews system refers to the information feedback of consumers by rating and scoring the purchased goods. With the increasingly fierce competition among e-commerce platforms, it has been widely used as a tool to extract consumer emotions [1, 2]. According to the data of 2017, more than 75% of companies analyzed consumer reviews to conduct market research. It can be seen that a large amount of online consumer generated data (hereinafter referred to as "reviews") has gradually become an important tool for producers to identify consumer demand and make sales and production decisions [3].

The most direct role of consumer reviews is to assist other consumers in their purchase decisions [4], while producers analyze consumer reviews to discover demand [5]. It is of great significance to identify the new demand created by consumers due to the fact that the earlier the manufacturer discovers the consumer demand and enters the market, the greater and more lasting market share benefit can be obtained. With the development of natural language processing technology, producers

can capture the new possible general interests of each user through semantic space and semantic network analysis and then discover the potential needs of consumers [6, 7].

This paper is based on the star rating and text evaluation data of three products (hair dryer, microwave oven, and baby pacifier) of Sunshine Company. After data cleaning, we first establish a BP neural network prediction model to help identify effective reviews. By using this model, the company can obtain useful information on the rating and comment data of any goods on sale. After obtaining the effective reviews, we further build a fuzzy comprehensive evaluation model to help the company determine the most successful and the most failed products and formulate the corresponding online sales strategy.

So far, we have successfully mined out the potential information in consumer reviews, but how this information affects consumer behavior such as purchase intention and then fluctuates the company's profits is unknown, which hinders the decision-making of the company's operators. So, we creatively use Python to segment valid reviews. We associate the word segmentation again to build a semantic

network to analyze consumer emotions and then judge whether `star_rating` and reviews have a radiating effect. In addition, we construct ARIMA model to analyze the temporal relationship between reviews and corporate reputation. Finally, we draw the conclusions, hoping that our results can help companies identify potential consumer demand and make production and sales decisions.

## 2. Data Sources, Data Processing, and Assumptions

**2.1. Data Sources.** As a global e-commerce giant, Amazon Mall has long introduced a comment mechanism, in which buyers can express their satisfaction, recommendation, and opinions on their products through “star ratings” and “reviews.” The data in this paper come from question C of the Mathematical Contest in Modeling in 2020. They are the scores and comments provided by customers of microwave ovens, baby pacifiers, and hair dryers sold in Amazon Market for a period of time [2, 8, 9]. The main concepts and definitions are shown in Table 1.

**2.2. Data Processing.** As we know, the amount of raw data is large and they are not processed, so we should do data processing according to the integrity and usefulness of the information by using four steps as follows:

- (i) *Step 1. Data screening:* The original data provided have the problem of redundancy and loss. Therefore, first of all, we should filter them by removing null values and removing unreadable data.
- (ii) *Step 2. Data modification:* There may be some unreasonable data in the raw data, which may have a significant impact on the conclusions drawn from the data, so we can modify part of the data through data processing. We calculate the ratio of helpful\_votes to total\_votes (the ratio of comments with zero total votes is zero) to modify the number of helpful votes, so eventually we can get the modified helpful\_votes.
- (iii) *Step 3. Data normalization:* Data normalization can provide an approach for comparison of various data and reflect the combined results of different factors. We should do the data normalization to facilitate the further analysis of the data. Firstly, we extract the length of review\_headline and review\_body and sum them and we get the review\_length sum. Secondly, we transform both vine and verify\_purchase into virtual variables by transforming  $Y$  to 1 and  $N$  to 0.
- (iv) *Step 4. Finding the indicators of data:* After the data processing above, we finally get 5 data indicators: `star_rating`, `vine`, `verify_purchase`, `review_length` sum, and modified helpful\_votes.

**2.3. Assumption.** In order to simplify our model, we make some assumption in our paper. The details are as follows: (1) We assume that the explanatory variables are deterministic variables, not random variables, and they are not correlated

with each other. The conditional mean value of the residual in the regression model, which is equal to 0, is independent of the explanatory variables, and each error variance is the same. (2) The average monthly evaluation data of time fragmentation of ARIMA model are not set as the break point of time series data, and seasonal and periodic changes are not considered. The expectation of irregular changes is 0. (3) In data processing, most data we use are standardized data. (4) The data we do dimension processing with or intuitively discard do not retain too much information about the whole. (5) We do not assume that these words expressing affection to products such as “great” and “like” must appear in the high level of star\_reviews. (6) The analysis of different brands of the same product and different products is consistent. We only analyze the situation of one product and do not process the rest again.

## 3. Recognition of Helpful Reviews Based on BP Neural Network

**3.1. Analysis Approach.** In order to select out the valuable part of all the ratings and reviews data submitted by consumers, we set up an intelligent algorithm BP neural network which has high efficiency to fit these data. Through this model, the company can obtain any information of the products on sale based on the data on ratings and reviews.

### 3.2. BP Neural Model

**3.2.1. Model Preparation.** BP neural network, which refers to terror backpropagation neural network, is the most widely used neural network Model. BP neural network generally consists of three layers: input layer, hidden layer, and output layer. The input signal acts on the output node through the hidden layer, and the transformation of the training function produces errors. When the output signal is sent, the error is returned to the network and sent to the input layer through the network. When the network propagates in the opposite direction, the network distributes errors to the neurons in each layer, adjusts the error signals obtained from each layer, and determines the weight of each cell [10].

**3.2.2. Model Establishment.** There are two steps in the process of establishing BP neural network. Firstly, we should set the network parameter  $k$  that represents the numbers of neurons. Secondly, we use existing data of input and output to triangle the network. We now drive the core formulas of the BP neutral network.

We set  $d_k$  as the expected output of the input layer's  $k$  output variable,  $o_k$  as the actual output of the input layer's  $k$  output variable, and  $y_k$  as the output of the hidden layer. Besides,  $w_{jk}$  and  $\theta_j$  are the weight and threshold of the hidden layer to the output layer, and  $V_{ij}$  and  $\theta_i$  are the sum of the weight and threshold of the hidden layer to the input layer [11].

Thus, when  $net_k = \sum_{j=0}^m (w_{jk} - \theta_j)$  the output of the input layer is as follows:

TABLE 1: The main concepts and definitions of this paper.

Serial	Variable	Variable description
1	Star_rating (integer)	A score used to rate a product with a number of stars between 1 and 5
2	Helpful_votes (integer)	The number of the votes of whether a product review is valuable or not
3	Total_votes (integer)	The total number of the votes of whether a product review is valuable or not
4	Review_headline (string)	The headline of a review
5	Review_body (string)	The content of a review
6	Vine (string)	The member of Amazon Vine Voices
7	Verify_purchase (string)	People who are defined to purchase a product

$$o_k = f(\text{net}_k), \quad k = 1, 2, \dots, l. \quad (1)$$

Then, when  $\text{net}_j = \sum_{i=0}^n (v_{ij} - \theta_i)$  the output of the hidden layer is as follows:

$$Y_j = f(\text{net}_j), \quad j = 1, 2, \dots, m. \quad (2)$$

The error between the networks output and the expected output is as follows:

$$E = \frac{1}{2} \sum_{k=1}^j (d_k - o_k)^2. \quad (3)$$

It is further expanded to the input layer as follows:

$$E = \sum_{k=1}^l \left\{ d_k - f \left[ \sum_{j=1}^m w_{jk} f(\text{net}_j) \right] \right\}^2. \quad (4)$$

**3.2.3. Error Analysis.** First, according to the formula above, function of each layer has error input in the network. Thus, by adjusting the weight and threshold, the error  $E$  can be changed. The adjustment of the weight should be proportional to the gradient of the error. When  $\alpha$  and  $\beta$  are the learning rates of the network, the formula is as follows:

$$\Delta w_{jk} = -\alpha \frac{\partial E}{\partial w_{ij}}, \quad j = 0, 1, 2, \dots, m; k = 1, 2, \dots, l, \quad (5)$$

$$\Delta v_{ij} = -\beta \frac{\partial E}{\partial w_{ij}}, \quad i = 0, 1, 2, \dots, m; j = 1, 2, \dots, m.$$

Learning rate determines the change of each weight during every training period. If the learning rate is high, the system will be unstable, and if the learning rate is low, the training time will be long and the convergence rate will be slow. Therefore, we should select a smaller positive number as the learning rate. In order to ensure the stability of the system, the selected learning rate should range from 0.1 to 0.8.

**3.3. Results Analysis.** Next, we will import the data into MATLAB, select the star\_rating and review\_length sum as input, and use the modified helpful\_votes as output to establish a neural network. Then we select 30 pieces of data randomly as the training set and 35 pieces of data as the test set and set the number of training steps as 10000, the target error as  $1 \times 10^{-5}$ , and the learning rate as 0.8. Besides, we

express the mean square error and the coefficient of determination in the picture. We can, respectively, get the comparison charts of the actual value and the predicted value of the training set, which are shown in Figure 1. The red solid line represents the true value and the blue-dotted line represents  $P$ .

From Figure 1, we can see that the fitting effect of neural network at each point is good, and RMS (root mean square) is above 0.9 and MSE (mean square error) is controlled at about 0.01 for both series.

Therefore, Sunshine Company is able to use our model to predict the sales of the three products in online marketplace and get useful information based on ratings and reviews so as to guide future sales and production activities.

## 4. Fuzzy Comprehensive Evaluation of Good and Bad Products Based on Principal Component Analysis

**4.1. Analysis Approach.** Having obtained the valuable information in the reviews, we are concerned about whether we can use this information to analyze which products are potentially successful or failing [12]. In order to determine the potentially most successful or most failing products, we establish the fuzzy comprehensive evaluation model based on principal component analysis.

### 4.2. Fuzzy Comprehensive Evaluation Model

**4.2.1. Model Preparation.** Generally, the selection of weight of the comprehensive evaluation model is influenced by some subjective factors. Using principal component analysis can help us calculate the weight of each index objectively, avoid the randomness of weight selection, and reduce the artificial errors in the process of calculation. With advantage of fuzzy comprehensive evaluation where processing line is not clear or subordinate relations is not specific, we can get more reliable results. The principles of the model are shown as follows: (1) Set the evaluation object as set  $P$ . The various indicators related to  $P$  are determined and then the evaluation factor set, namely, the indicator set  $U$ , is established. Select the set of possible evaluation results to establish the evaluation level set  $V$ . (2) Establish membership degree relation matrix  $R$ . According to the grades in the evaluation grade set  $V$ , the membership degree of each index in  $U$  is evaluated. Suppose that the subset of membership degree of corresponding subindexes in  $U$  is  $R_i = (r_{i1}, r_{i2}, \dots, r_{im})$ ,

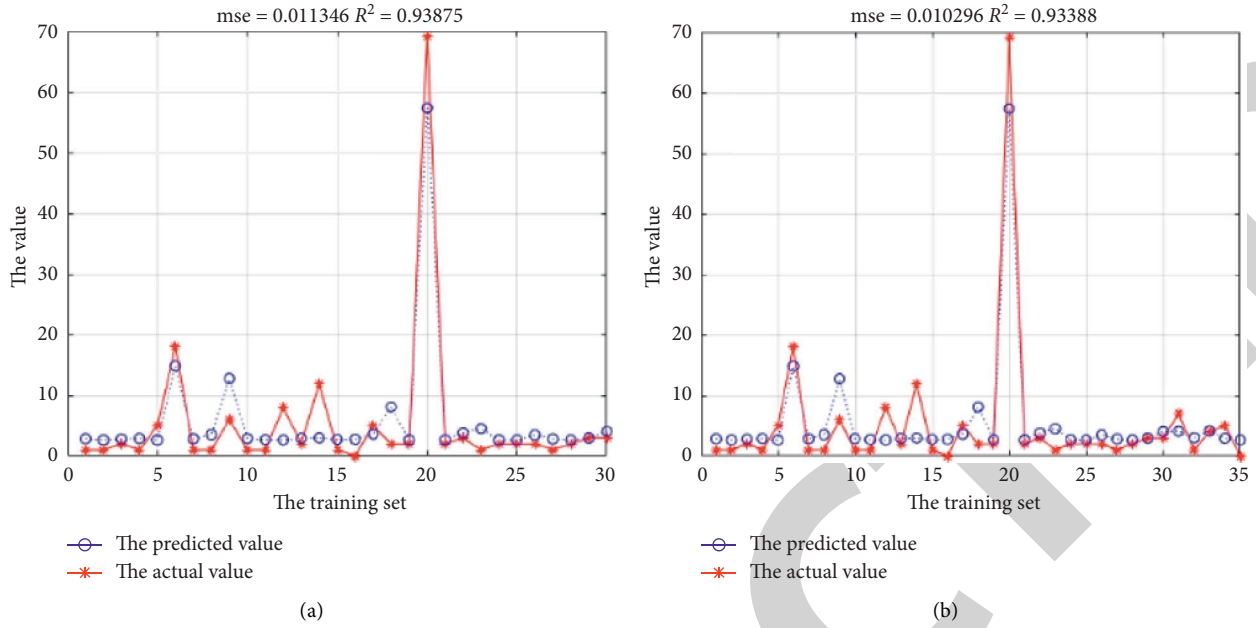


FIGURE 1: The actual value and the predicted value of the training set.

where  $r_{ij}$  represents the membership degree of  $u_j$  to the evaluation level in set  $V$  and its corresponding value  $v_j$ , and  $r_{i1} + r_{i2} + \dots + r_{im} = 1$ . The subsets of membership degree of  $m$  evaluation grades constitute the membership degree relation matrix  $R$ . The membership value is calculated by linear distribution function. (3) Establish the weight vector  $W$ . In order to reflect the importance of each subindex  $u_i$  to  $P$ , we set the weight vector as  $W_i = (w_1, w_2, \dots, w_m)$ . (4) Establish the evaluation result vector  $B$ . The evaluation result vector  $B$ , that is,  $B = WR$ , is obtained by applying the synthesis operation that can fully reflect the function of membership relation matrix  $R$  and weight vector  $W$ . How to establish our model is shown in the five steps [13, 14].

- (i) Step 1. Select the original data of the text and ratings. According to the actual situation, we select the data of five variables coming from 3 products: star\_rating, helpful\_votes, vine, verified\_purchase, and review\_length sum.
- (ii) Step 2. Conduct principal component analysis according to the selected data. We mainly determine the principal component with the standard of characteristic value greater than 1 or slightly less than 1 and the cumulative variance contribution rate between 70% and 85% to determine the weight of each product.
- (iii) Step 3. Establish the fuzzy comprehensive evaluation model. We divide the evaluation level into three levels, which correspond to the relative quality of the product. Then, we calculate the membership matrix according to the membership degree, use MATLAB to calculate the result vector, determine the level of products through weighted processing, and make corresponding analysis.

**4.2.2. Principal Component Analysis.** According to the actual situation, we select the data of five variables coming from 3 products: star\_rating, helpful\_votes, vine, verified\_purchase, and review\_length sum.

Taking the data of hair dryers as an example, we use STATA to process the principal component analysis. The results are shown in Tables 2 and 3.

According to the principle of principal component selection, we select the first three principal components and substitute the original variables with  $x_1, \dots, x_5$ . Thus, we can get a principal component expression, and the rest of expressions can be inferred like it. The formula is as follows:

$$\text{Comp1} = -0.2143x_1 + 0.3464x_2 + 0.4442x_3 - 0.5536x_4 + 0.5747x_5. \quad (6)$$

Next, we use the number of loads in factor loading matrix table to divide the variance and the square root value of the principal component characteristic in the cumulative contribution rate table; then we can get the coefficients of all the indicators in each of the main component linear combinations. Then we multiply them to the variance contribution rates of the principal component and divide them to the sum of variance contribution rates, and we can obtain the corresponding weight vector  $W = (0.1987, 0.0944, 0.3048, 0.2082, 0.1940)$ , which can reflect the different indicators  $U$  for evaluation objects to the importance of the hair dryer.

**4.2.3. The Fuzzy Comprehensive Evaluation.** We divide the degree of the evaluation object into three levels and establish the evaluation level set  $V = \{1, 2, 3\}$ . Then we establish the membership degree relation matrix  $R$  (membership degree function in the matrix is calculated by linear distribution function) to evaluate the membership degree of each index



TABLE 2: The results of the principal component analysis.

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	1.64055	0.562677	0.3281	0.3281
Comp2	1.07787	0.094071	0.2156	0.5437
Comp3	0.983801	0.327314	0.1968	0.7404
Comp4	0.656486	0.015192	0.1313	0.8717
Comp5	0.641294	0	0.1283	1

TABLE 3: The results of the principal component analysis.

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Unexplained
star_rating	-0.2143	0.4416	0.8142	-0.2013	-0.2358	0
helpful_votes	0.3464	-0.5847	0.51	-0.1724	0.4982	0
vine	0.4442	0.6018	0.0428	0.4312	0.5028	0
verified_purchase	-0.5536	-0.2532	0.2149	0.7528	0.1282	0
length_sum	0.5747	-0.1919	0.1702	0.4208	-0.6533	0

in set  $U$  according to the very level in the evaluation level set  $V$  [15].

$$R = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.8 & 0 & 0.2 \\ 0.1 & 0.54 & 0.36 \\ 0.22 & 0.14 & 0.64 \\ 0.3 & 0.6 & 0.1 \end{pmatrix}. \tag{7}$$

4.3. *Result Analysis.* The fuzzy comprehensive evaluation model we select is of outstanding-principle component type. We use MATLAB to calculate the result vector  $B = WR = (0.0755, 0.1646, 0.1332)$  and use the weighted average method to get that the final score is 2.1546. Similarly, we can get that the scores of microwave ovens and pacifiers are 1.6026 and 1.9255. It is obvious that the best potential products are hair dryers and the worst are microwave ovens.

### 5. Time Series Prediction of Products' Reputation Based on ARIMA Model

5.1. *Analysis Approach.* In order to discuss the increasing or decreasing trend of products' reputation in the online marketplace, we use the ARIMA model to identify metrics and patterns based on timeline for each piece of data [16, 17].

#### 5.2. ARIMA Model

5.2.1. *Model Preparation.* Firstly, we analyze and process the data through scatter diagram. We take a hair dryer as an example. We first extract all the data of star\_ratings and

review\_dates. Due to the large amount of data, we average the data of star\_ratings on a monthly basis to obtain the average star\_ratings data of 140 months from March 2002 to August 2015 (there are no data for some months in the middle). The scatter diagram is shown in Figure 2. The horizontal axis represents the month and the vertical axis represents the average star level of each month.

Therefore, we can find the following conclusions from Figure 2: Firstly, we can see that the data fluctuate greatly and the data generated are relatively regular. It is indicated that the number of people who buy the product at the beginning is small and there are no reviews in some months. Secondly, in 2003 and 2004, the data once reached the lowest level. It is indicated that the company may have taken a series of measures to save its reputation. Thirdly, from the end of 2006 to the beginning of 2007, the review data of the products began to stabilize with small fluctuation. It is indicated that the number of buyers began to increase. Finally, in the later period, the average star\_rating is stable at around 4.1, and it is also the average total star\_rating of the product. On the other hand, the flattening trend also reflects the fact that the correlation coefficient between the star\_rating and the other two variables is nearly zero.

Next, we import the data into STATA to process time series analysis. There are mainly four different factors that play an important role in the formation and development of time series, including long-term trend factor, seasonal change factor, cyclic change factor, and irregular change factor. Besides, as the time series formed in this topic is relatively long, we cannot ignore the influence of long-term trend. The main method for determining and analyzing long-term trend is smoothing the time series. Thus, we choose the MA (moving average) method and the formula of smoothing is as follows:

$$x'(t) = \frac{x(t-2) + x(t-1) + x(t) + x(t+1) + x(t+2) + x(t+3)}{6}. \tag{8}$$

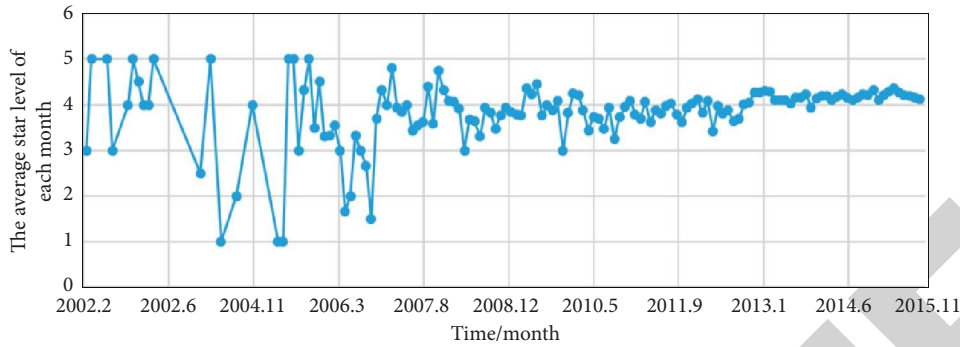


FIGURE 2: Average star\_rating per month.

**5.2.2. Model Establishment.** ARIMA model is a common and effective method in time series prediction. ARIMA method can find a model suitable for data investigation under the condition that the data pattern is unknown, so it has been widely used in financial and economic prediction. Its basic principle is as follows: Firstly, smooth the original time series by using MA (moving average) method. Secondly, determine the type of model, the order of model, and undetermined mined parameters by analyzing the characteristics of ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) of stationary sequence. Finally, analyze and predict the future times series [18, 19].

ARIMA model can be known as ARIMA ( $p, d, q$ ), where  $P$  is the autoregressive order number,  $d$  is the difference order, and  $q$  is the sliding average order number. The  $P$  value of the corresponding  $Q$  statistic is almost zero at the significance level of 0.05, so we reject the original hypothesis that the time sequence is irrelevant. Meanwhile, the  $P$  value passing the white noise test also rejects the original hypothesis that the sequence is white noise. Then, through the full test, we can get the value of  $P = 0.163 > 0.05$ , which means that it also passes the stationarity test. Thus, the ARIMA model could be used and the model we selected is ARIMA (1, 0, 1) [20].

**5.3. Results Analysis.** Through the above process, we have determined our model as ARIMA (1, 0, 1). Then we can draw the prediction results by calculation, as shown in Table 4. It can be seen from the table that each coefficient of the model is significant and the standard deviation of the residual is very small, and the  $P$  value of each term is greater than the significance level; that is, the residual passes the autocorrelation test.

Above all, we can say that the model does not have autocorrelation, and the ARIMA model can be used to predict the future changes of product reputation in the market. The formula of ARIMA model can be restored to the following:

$$\begin{aligned} \text{star}_t &= 3.859 + 0.870\text{star}_{t-1} + \varepsilon_t + 0.372\varepsilon_t, \\ \hat{\varepsilon} &= 0.154. \end{aligned} \quad (9)$$

According to the ARIMA model, the average product stars in the next five months are predicted to be 4.124739,

TABLE 4: The correlation between the star\_rating and reviews\_length sum.

Variable	Coef.	Std. err.	The value of $t$	The value of $P$
Star_rating	-0.1876	-0.1163	-1.61	0.1070
Reviews_length	0.0079	-0.0004	18.87	0.0001
_Cons	-3.8989	0.0598	7.65	0.0001

4.090306, 4.060333, 4.034242, and 4.011531, indicating that the product reputation will decline in the future.

Taking the hair dryer data as an example, we will process the sensitivity analysis of the ARIMA model. ARIMA model is the core part of the three-parameter input: the autoregressive order  $p$ , difference number  $d$ , and the change of the moving average order  $q$ . So we change the parameters and smoothing data of the initial model ARIMA (1, 1). At the same time, we make the average star\_ratings prediction for the next five months. Here, we do not restore the ARIMA model expression and do not list the results of model parameters. The selected model involves constant terms. We change parameters  $p$  and  $q$ , respectively; the range of change is the integers from 0 to 3. The sensitivity analysis is as follows in Figure 3.

Firstly, we find in Figure 3 that  $p$  is changing and the red line  $p = 0$  has the largest change range and the largest error, while other  $p$  and  $q$  are in good fitting with the original data. Secondly, the prediction of the red line  $p = 0$  circulates from low to high and finally is lower than the level of the original data. Finally, in the forecast of the last five years, no matter the change of  $p$  or  $q$ , the ARIMA model shows a downward trend, which should be paid attention to. Therefore, we can find that the conclusion of sensitivity analysis is consistent with the conclusion above, which shows that our model has good stability and strong adaptability.

## 6. Exploration of Radiation Effects Based on K-Means Clustering Algorithm

**6.1. Analysis Approach.** In order to explore whether specific star\_ratings and reviews will cause more reviews, firstly we use word segmentation to find the relationships between specific vocabularies and star\_ratings; then we use  $K$ -means clustering algorithm to further explore the radiation effects.

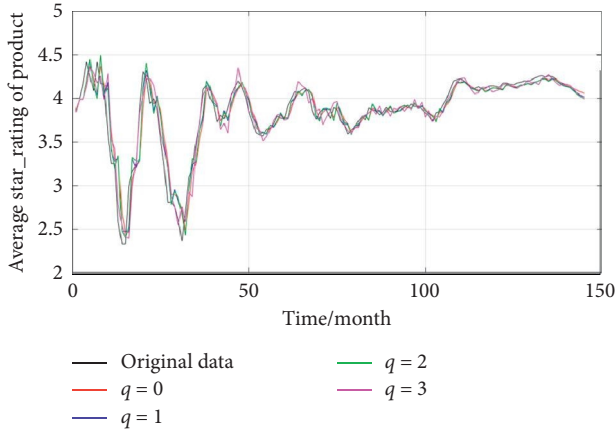


FIGURE 3: The picture of sensitivity analysis.

## 6.2. Word Segmentation

**6.2.1. Model Preparation.** Word segmentation is an important concept in NLP (Natural Language Processing) [21]. It decomposes long texts such as sentences, paragraphs, and articles into data structures, which regard words as their units. Such processing helps us to express complex problems in mathematical language [22–24]. The basic unit of word segmentation in English is a word. The basic steps of the model are as follows [25].

- (i) *Step 1.* Separating the sentences into individual words according to the space, punctuation, and so forth.
- (ii) *Step 2.* Removing the stopping words: Remove the words that appear frequently but have no practical meaning, such as “it,” “that,” “to,” “a,” and “the.”
- (iii) *Step 3.* Lemmatization (restoring the part of speech) and stemming (extracting the stem): There are some special forms of the words in English, such as the singular and plural nouns and the -ing forms and -ed forms of verbs; we need to restore these morphed words and extract the main components from them.
- (iv) *Step 4.* Eigenvector extraction [12, 26].

### 6.2.2. Model Establishment

- (1) Lemmatization and stemming. Taking the hair dryer data as an example, we extract 20 words such as “recommend” and “light” from high to low frequency through Python and convert them into word vectors. If the word “recommend” which has the highest frequency appears in the ratings while the other words do not,  $a$  is a vector with 20 rows of columns and we will mark the eigenvector as

$$a = (1 \ 0 \ 0 \ \dots \ 0)^T. \quad (10)$$

- (2) Removing stop words. After removing these words which cannot express the actual meaning in the

product ratings such as “buy” and “purchase,” we get the word frequency table, which is shown as Table 5. The number to the right of the word represents the number of occurrences of the word in the rating title and rating content.

- (3) Eigenvector conversion. Finally, we convert the star\_ratings into eigenvectors and calculate the Pearson correlation coefficient between these eigenvectors of words and eigenvectors of star\_ratings. The calculation formula is as follows [27]:

$$r(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \quad (11)$$

**6.2.3. Results Analysis.** Finally, the correlation coefficient matrix we get is shown as in Table 6. The rows of the matrix are the star\_ratings from low to high, and the columns of the matrix are the high-frequency words we choose.

As can be seen from Table 6, the correlation coefficient between each word and the star\_rating is very slight, most of which are less than 0.1 and they are almost irrelevant. However, considering the uncertainty of big data and the processing of taking things out of context (we removed negative auxiliary verbs), some of the correlations are still fairly high. On the other hand, based on some intuitive situations, the correlation coefficient between the word “great” and the five-star rating is 0.1718, and the correlation coefficient between “love” and the five-star rating is 0.2581. We can conclude that these specific descriptions are obviously closely related to the level of star\_rating and some words with negative meaning of the products did not show high correlation coefficient.

Firstly, we can easily find that only “low” and “weight” in the high-frequency words are likely to be the words with negative reviews. Secondly, the average star\_rating of hair dryers is as high as 4.1, whose number of negative reviews is relatively less. Finally, we cannot regard the negative meanings as negative meaning without the context directly, which makes us choose to abandon it when we are dealing with negative words.

## 6.3. K-Means Clustering Algorithm

**6.3.1. Model Preparation.** Clustering is the process of looking for similarities between a collection of figurative or abstract objects and dividing them into categories. The basic steps of  $K$ -means clustering algorithm are as follows:

- (i) Step 1. Select  $k$  objects from the data as the initial clustering center.
- (ii) Step 2. Calculate the distance between each cluster object and the cluster center and divide them into parts based on the distance.
- (iii) Step 3. Calculate each cluster center again.

**6.3.2. Model Establishment.** We identify this problem as a clustering problem; that is to say, we do not assume that

TABLE 5: The word frequency.

Words	Frequency	Words	Frequency	Words	Frequency	Words	Frequency
Recommend	955	Thick	728	Fine	630	Pretty	486
Light	860	Fast	693	Low	624	Retractable	465
Powerful	856	Cool	680	Quiet	522	Worth	450
Easy	845	Heavy	664	Weight	521	Perfect	445
Nice	808	Great	660	Love	517	Work	395

TABLE 6: Pearson correlation coefficient between eigenvectors of words and eigenvectors of star\_ratings.

Word	Star				
	1 star	2 stars	3 stars	4 stars	5 stars
Recommend	-0.0334	-0.0024	-0.057	-0.0298	0.0766
Light	-0.0627	-0.0241	-0.0099	0.0406	0.0215
Powerful	-0.0712	-0.0231	0.0154	0.0052	0.0391
Easy	-0.0708	-0.0446	-0.0415	0.0492	0.047
Nice	-0.0747	-0.0329	-0.0002	0.0507	0.0191
Thick	-0.0492	-0.0295	0.0086	0.0095	0.0299
Fast	-0.0716	-0.0385	-0.0337	0.0077	0.0728
Cool	-0.0162	0.0097	0.0248	0.0221	-0.0266
Heavy	-0.015	0.014	0.0655	0.0567	-0.0798
Great	-0.1512	-0.0787	-0.0996	0.0122	0.1718
Fine	-0.0131	0.0529	0.0676	0.04	-0.0872
Low	0.0181	0.0367	0.0239	0.0118	-0.0505
Quiet	-0.0475	-0.0249	-0.004	0.0078	0.0353
Weight	-0.0494	-0.0254	-0.01	0.0217	0.0292
Love	-0.1251	-0.1016	-0.1135	-0.0935	0.2581
Pretty	-0.0248	-0.0073	0.0385	0.0752	-0.0632
Retractable	-0.0093	0.0225	0.0273	0.0192	-0.0358
Worth	0.0394	-0.0039	-0.0265	-0.0286	0.0165
Perfect	-0.0676	-0.0423	-0.0574	-0.0186	0.1064
Work	-0.1004	-0.0473	-0.0056	0.093	0.0105

these words expressing affection to products such as “great” “like” must appear in the high level of star\_reviews. Instead, we first find out the similarities of all kinds of reviews by means of clustering and divide them into categories. Then, according to these categories, we in turn look for corresponding star\_ratings to determine whether customers have a tendency to publish such reviews after a series of certain stars comments [28].

Firstly, we are given the reviews\_text data set, where  $d$  represents the dimension of the data set, and  $n$  represents the amount of data in the data set and the number  $K$  of partition sets we want to generate. Through the  $K$ -means algorithm, we can divide the object into  $K$  different clusters  $C_k$ ,  $k = 1, 2, \dots, k$ . Now we select a central  $k$  in each cluster. Similar to the general optimization problem, the clustering problem is to minimize the sum of squares between the midpoint of each cluster and each center. That is,  $L = \sum_{k=1}^K \sum_{x_i^d \in C_k} \|x_i^d - \mu_k\|^2$  to the minimum. The distance here is referring to the Euclidean metric.

$K$ -means clustering algorithm is a classical algorithm to solve clustering problems, which is simple and fast. For processing large data sets, the algorithm keeps contractibility and high efficiency. The effect is best when the cluster is close to the normal distribution.

From Figure 4, we can see that most of the points are loosely arranged, but some of them have the potential to become the center of the cluster. Then we process on these points with  $K$ -means clustering and five cluster centers are obtained corresponding to each star.

**6.3.3. Results Analysis.** We do separation with five kinds of color to mark the various elements of the clusters represented with a point and the center represented with an  $x$ . We restore the principal component through the serial number of the five clusters to find these star\_ratings of the reviews initially and we get the corresponding relationship of the reviews, clustering, and star\_ratings (we will not duplicate recording the number of reviews). Thus, we can find 3 significant results. Firstly, blue cluster has five-star reviews with the maximum number of 212, which appears in the corresponding term vectors “how many” 132 times and “love” 89 times. Secondly, the green cluster has most one-star reviews with the number of 38 where the corresponding word vector appeared: “like” 25 times, which may be the reason customers wanted to express “do not like” or “dislike.” This is also shown in the picture. Thirdly, it can be concluded from the model that customers are most likely to

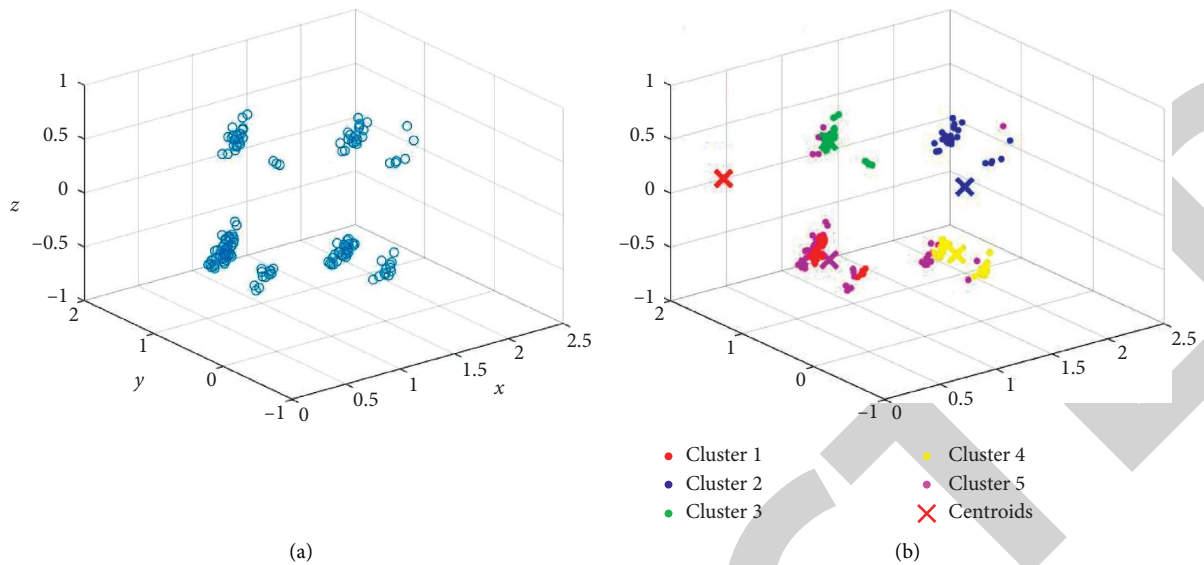


FIGURE 4: The result of K-means clustering algorithm. (a) Clustering results. (b) Cluster assignments and centroid.

directly express their love for the product or recommend others to buy it when they see five-star reviews; however, customers are also likely to directly express their dissatisfaction when they see one-star negative review [29]. So, finally, Sunshine Company should adjust its sales strategy in time according to the star\_ratings of the review to increase the product production when seeing the continuous good reviews and to save the reputation of own product in time when seeing bad reviews increase.

## 7. Conclusion

Starting from a case analysis, this paper establishes a model suitable for any platform and any commodity to mine the potential information behind the ratings and reviews of online product, which provides certain help for companies to make online sales strategy and identify the demand of consumers. However, in order to simplify the analysis process, some assumptions made by us make the model deviate from the reality. This article uses BP neural network and other methods to establish multiple models, adopts the method of combining qualitative description and quantitative calculation, makes assumptions and demonstrations, uses clear logic to explain the principles and practical applications of various models, and uses MATLAB and other methods. Therefore, we strongly believe that our conclusions can help companies use consumer review data to explore the market deeply and make appropriate production plans and sales strategies.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the General Project of Philosophy and Social Science Planning of Anhui Province: Research on Government R&D Subsidies Promoting Green Innovation Efficiency of Anhui High-Tech Industries (no: AHSKY2019D085); Anhui University of Finance and Economics School-Level Teaching and Research Fund Project (acjyyb2020011); Key Projects of Support Program for Outstanding Young Talents in Anhui Province Colleges and Universities (no. gxyq2018119); Key Project Funds from Anhui Education Ministry (no. 2019rcsfjd089); and Student Scientific Research Fund Project of School of Economics, Anhui University of Finance and Economics (ACJXYZD2002).

## References

- [1] E.-J. Lee and S. Y. Shin, "When do consumers buy online product reviews? effects of review quality, product type, and reviewer's photo," *Computers in Human Behavior*, vol. 31, pp. 356–366, 2014.
- [2] H. Hong, D. Xu, G. A. Wang, and W. Fan, "Understanding the determinants of online review helpfulness: a meta-analytic investigation," *Decision Support Systems*, vol. 102, pp. 1–11, 2017.
- [3] Z.-P. Fan, G.-M. Li, and Y. Liu, "Processes and methods of information fusion for ranking products based on online reviews: an overview," *Information Fusion*, vol. 60, pp. 87–97, 2020.
- [4] K. Floyd, R. Freling, S. Alhoqail, H. Y. Cho, and T. Freling, "How online product reviews affect retail sales: a meta-analysis," *Journal of Retailing*, vol. 90, no. 2, pp. 217–232, 2014.
- [5] R. Decker and M. Trusov, "Estimating aggregate consumer preferences from online product reviews," *International Journal of Research in Marketing*, vol. 27, no. 4, pp. 293–307, 2010.
- [6] X. Xu, X. Wang, Y. Li, and M. Haghghi, "Business intelligence in online customer textual reviews: understanding consumer



## Research Article

# Construction of the Evaluation System of Sustainable Utilization of Large Stadiums Based on the AHP Method

Li-Cai Zhu,<sup>1,2</sup> Zhi Gao ,<sup>1</sup> Jia-Ming Zhu ,<sup>3</sup> and Di Zhang<sup>2</sup>

<sup>1</sup>*Institute of Sports Economics, Anhui University of Finance and Economics, Bengbu 233030, China*

<sup>2</sup>*School of International Economics and Trade, Anhui University of Finance and Economics, Bengbu 233030, China*

<sup>3</sup>*School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China*

Correspondence should be addressed to Zhi Gao; gicy0415@163.com

Received 11 July 2020; Revised 19 September 2020; Accepted 20 October 2020; Published 9 November 2020

Academic Editor: M. Javaid

Copyright © 2020 Li-Cai Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Large-scale sports stadiums are an important part of China's sports stadium system and an important material basis and prerequisite for China's basic public sports service system. This paper aims to establish the evaluation index system by exploring the utilization of large stadiums and gymnasiums under the background of sustainable development and to provide important reference for the construction and utilization of stadiums and gymnasiums in the future. Using literature methods, quantitative statistics, analytic hierarchy process (AHP), and other methods, focusing on sustainable use, this study develops a postgame utilization evaluation system for large stadiums, considering the design, construction, operation, and management of the stadium. This paper identifies 5 first-level indicators, 13 second-level indicators, and 38 third-level indicators, which mainly included five aspects of comprehensive performance, service level, environmental performance, economic performance, and venue maintenance. The AHP method was used to determine the index weight coefficient, thus forming a sustainable use evaluation index system for the large stadiums. Finally, taking Bengbu Sports Center in Anhui province as an example, this paper makes an empirical analysis.

## 1. Introduction

Large-scale sports stadiums, which are an important part of China's sports stadium system, are an important material basis and prerequisite for China's basic public sports service system. Since the 21st century, China has hosted a series of large-scale international sports events such as the Beijing Olympic Games, the Guangzhou Asian Games, and the World University Games. Moreover, various provinces and cities have hosted a series of large-scale national sports events such as the National Games and the Provincial Games. China has thus accumulated a large number of large-scale sports venues and stadiums. The sixth national sports census showed that as of the end of 2013, China had built 1,093 large stadiums [1]. With the Healthy China and the National Fitness Program becoming a national strategy, people's growing sports demand and the supply of public sports services have produced serious contradictions [2]. There are a series of problems behind the current situation of

shortage and waste in the supply of stadiums and gymnasiums [3]. The "13th Five-Year Plan for Sports Development" indicated that it is necessary to actively promote the reform of the stadium management system and the innovation of operational mechanisms; it is also of great importance to improve the sports stadium operation evaluation system and to implement the integrated mode of venue design, construction, operation, and management in order to combine the demand for the event and a comprehensive postgame use of the stadiums in an organic way, strictly adhering to the scientific outlook on development [4]. Therefore, it is of great guiding significance and practical importance to construct a postgame utilization evaluation system for large stadiums, considering the design, construction, operation, and management of the stadium and focusing on sustainable utilization. Based on the comprehensive performance of stadiums and gymnasiums, service level, environmental performance, economic performance, and stadium maintenance, this paper constructs an



evaluation system for the sustainable utilization of large-scale stadiums and gymnasiums after competition, including the whole life cycle of stadiums and gymnasiums design, construction, operation, and management, and constructs a fuzzy comprehensive evaluation model. The article focuses on whether the use of large stadiums and gymnasiums after the games is in line with the concept of sustainable development.

## 2. Literature Review and Analysis

*2.1. Literature Review.* Sports venues are an important part of public buildings and the basis for sports activities. They have the characteristics of large initial investment, high operating costs, and high resource consumption. In the context of the era of green development, stadiums should meet not only the development needs but also the requirements of resource conservation. The economic sustainability of stadiums was assessed also considering Emission Trading System, energy saving, and energy production [5]; in terms of site selection and distribution, Johnson indicated that in the early 20th century, stadiums were usually located in the city center, very close to the train station or in the public pedestrian area, because most people lived in and around the city center [6]. Fabris points out that sports stadiums and arenas have sustainability features [7]. Most studies on venue selection are focused on the city center or the city [8]. Compared with urban suburbs, city centers are more conducive to the development of the venue. Roger and Kemp believe that the extent of the development of public transportation systems around stadiums is an important factor in testing the pressure on stadium parking lots [9]. Therefore, it is a wise choice to locate stadiums in residential areas. The California Public Policy Research Institute indicated that there is an unequal distribution of public sports facilities. More than half of Californians agree that poor communities do not enjoy their fair share, and it is difficult to use well-preserved parks and sports facilities [10]. A study at the University of Southern California also found an imbalance in the Los Angeles park, where communities with low-income and colored populations find it more difficult to access public sports facilities than white and high-income communities [11]. Some scholars have suggested that the existing analysis model of individual responsibility and personal behavior choices for physical exercise should be transformed into social responsibility and an overall macro research model, focusing more on the role of environment and public policy in shaping lifestyle and behavioral patterns [12]. Norman et al. assessed the carbon emissions of the community from the three levels of community construction, construction operation, and transportation mode and concluded that the transportation mode has the greatest impact on carbon emissions [13]. Winkelman et al. reported that site selection and construction methods have a significant impact on the emissions of transportation infrastructure construction [14]. Burke and Shang indicated that whether large stadiums can be managed after the game is an important issue for event organizers in the initial stage of construction [15]. If the stadium is not used continuously after the game, the maintenance cost will be high. In the case of the London Olympics, it was decided that the main Olympic venues will be reduced in size and put into use after the games, and the temporary venues will be replanned

for future urban construction [16]. In terms of venue operations, Santo analyzed a large amount of data in the field of stadium operation mode research, concluding that large stadiums have a facilitating effect on some local economies as well as some negative effects [17]. China's sports stadiums are mainly responsible for the training and competition of professional teams. They are not oriented to the masses and have a narrow business scope [18]. The usage rate of Chinese sports venues is also low. Many professional sports venues charge the public excessively, making few people choose professional sports venues as a place for fitness exercise. A significant number of stadiums are difficult to use, except for some big games [19]. Aljehani concludes that stadium scalability is impacted by urban form factors such as land uses density, accessibility, and connectivity. These factors impact the decision to scale down or up a stadium for better urban interactions [20]. Crompton et al. indicated that the current US stadium facilities mainly use public-private joint financing model—public private partnership—that is, both government and private investment in the construction of stadiums [21]. Rebeggiani drew on the financing model and tax policy of American stadiums and studied the investment operations of German stadiums [22]. Westerbeeck et al. analyzed the success of the operation of a large-scale comprehensive sports stadium and the reasons for its success [23]. Based on a summary of these previous studies, the present study develops a sustainable postgame utilization evaluation system for large stadiums focusing on the relationship between venues, environment, and people.

*2.2. Challenges and Gap Analysis.* Some literatures promote the development of stadiums and gymnasiums from the perspectives of external environment, such as the site selection of stadiums and gymnasiums, the perfection degree of surrounding transportation system, and the distribution of public sports facilities. From the perspective of subjective factors, individual behavior patterns and personal responsibilities will also affect the sustainable use of venues. From the perspective of venue operation, investment and financing mode and government tax will also have a far-reaching impact on venue development [24]. It can be concluded that most of the previous literature separated people, venues, and environment. The challenge of this paper is to integrate both subjective and objective factors into the evaluation system to make it complete and comprehensive.

## 3. Value Basis and Principles of the Evaluation Index System for Sustainable Postgame Use of Large-Scale Stadiums

The value bases of the evaluation system for sustainable utilization of large stadiums and gymnasiums are Revalue, Renew, Reuse, Reduce, and Recycle. Principles include comprehensiveness, scientificity, and operability.

*3.1. Value Base.* The concept of sustainable development can be summarized as follows: coordinating “the development of a natural-economic-social composite system without

destroying the resources and environmental carrying capacity” [25]. The sustainable use of sports venues can be expressed as follows: the implementation of green development throughout the life cycle of a stadium, including site selection, planning, design, construction, use, postmatch operations, and construction materials. Sustainable development minimizes the damage to the natural and social environment and maximizes the efficient use of stadiums.

The value base of sustainable postmatch use of large-scale sports stadiums in China is mainly reflected in the 5R principles of sustainable development: Revalue, Renew, Reuse, Reduce, and Recycle. Revalue refers to “re-evaluation and rethinking.” China is facing a contradiction between people’s growing sports demand and the supply of public sports services. People’s sports demand and the status quo of sports supply must be reconsidered. A balanced and sound public sports service system that meets people’s diverse sports needs must be built. Renew refers to “revival and transformation.” An improper postgame use and the backwardness of the operation management mechanism have led to a sharp increase in the loss and damage of the venue. It is necessary to strengthen the focus on the maintenance and renewal of stadiums. Reuse refers to “rerun and reutilization.” The supply of sports venues presents a situation of shortage and waste, so it is necessary to reuse the stadiums after the competition, expand the open area and time, and improve the utilization rate of stadiums. Reduce refers to “reduction.” During the construction and operation of sports venues, stadiums are designed with reasonable structure, and green materials are used to reduce the damage to nature. Recycle refers to “recycling,” mainly in terms of the operation of sports venues, attempting to save energy, using recyclable resource materials, and improving the service life-span of stadiums.

Sustainable postmatch use of stadiums in China is based on the relationship between venues, environment, and people, and it can help solve the contradiction between the growing sports demand of the Chinese people and the supply of public sports services. In the whole life cycle of the venue, the principle of “improving the utilization rate of the stadium, improving the service life of the stadium, and reducing the impact on the environment” is the priority principle, and the 5R principles of sustainable development are implemented throughout the life cycle of the stadium. Based on the concept of sustainable development and considering the 5R principles, the life cycle diagram of large stadiums is constructed, as shown in Figure 1.

**3.2. Principles.** The process of postgame use of large-scale sports stadiums involves various levels of government and related functional departments, sports business organizations, the public, and other entities, and its impact is broad, diverse, and complex. Therefore, in establishing a postmatch evaluation system of large-scale sports stadiums, it is necessary to consider not only the current situation of the construction and operation of various sports venues in China but also the economic, social, and natural impacts of stadiums. The goal of the system is to achieve a people-oriented, harmonious coexistence between man and nature and sustainable development. The system strives to reflect not only its application adaptability but also the

guiding significance of policy, based on the status quo as well as future directions.

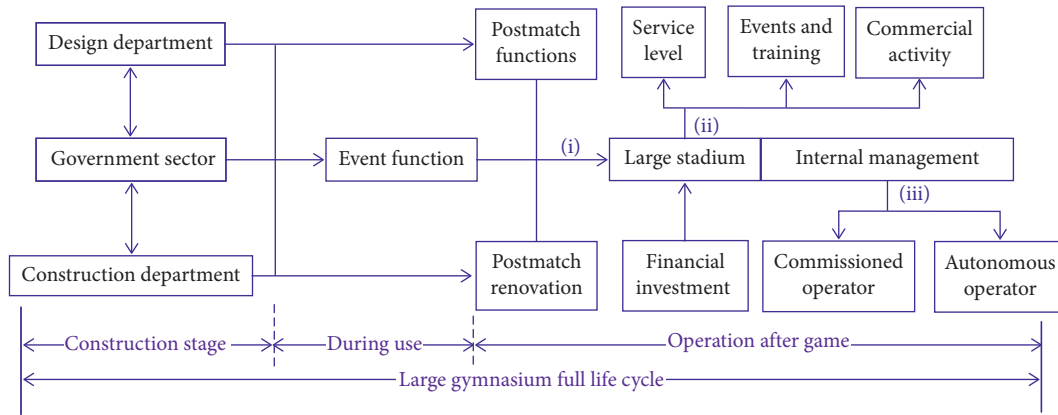
**3.2.1. Comprehensiveness.** The postgame use of large-scale sports stadiums involves multiagent, far-reaching, multi-objective system engineering, not only related to the postgame operation stage of the stadium but also affected by the construction stage and the stage of the use for the event, and it covers the entire life cycle of large stadiums. At the same time, due to the interaction of social, economic, and natural factors, the evaluation system needs to be completed by venue management department, government departments, and the public. In the process of establishing such an evaluation system, it is necessary to emphasize the theme of sustainable use of large-scale sports stadiums based on inheriting the social service and economic performance of general venue evaluation and to closely link society, people, and nature in order to achieve the goal of coordinated development of these three aspects. To make the whole evaluation system more systematic and comprehensive, it is necessary that the system reflects not only the current operational status of the stadium but also its development potential, sustainable development, and the attributes of resource optimization.

**3.2.2. Scientific Nature.** The scientific nature of the indicator system is the basis for ensuring accurate and reasonable evaluation results. First, the selected indicators should reflect the essential characteristics of the postgame use of sports stadiums, the status quo of stadium supply and demand, and the theme of sustainable use. Second, there are many factors influencing the use of gymnasiums after holding sports matches, and there is a certain correlation between them. It is necessary to ensure the independence of each index and to avoid an overlap with attributes and connotations. Finally, the indicator system should focus on the goal of sustainable use and comprehensively reflect the performance of the postgame use of sports stadiums, including both positive and negative performance evaluation, and important indicators cannot be omitted.

**3.2.3. Operability.** The selection of indicators should consider theoretical rationality and the actual situation of the postgame operation of gymnasiums in China to ensure that the indicators can be collected more accurately. At the same time, it is necessary to comprehensively use absolute and relative value indicators, quantitative and qualitative indicators, and short-term and long-term indicators to embody specific evaluation indicators and make the evaluation operable.

## **4. Construction of Sustainable Postmatch Utilization Evaluation Index System for Large-Scale Stadiums**

Starting from China’s actual national conditions, this study designs a preliminary indicator system based on the theme of sustainable use, drawing on the rich research results of its predecessors, and then constructing a comprehensive indicator system of five first-level indicators, including



- (i) The postgame functional transformation;
- (ii) After the stadium is open to use;
- (iii) The stadium operation management model.

FIGURE 1: Full life cycle diagram of a large stadium.

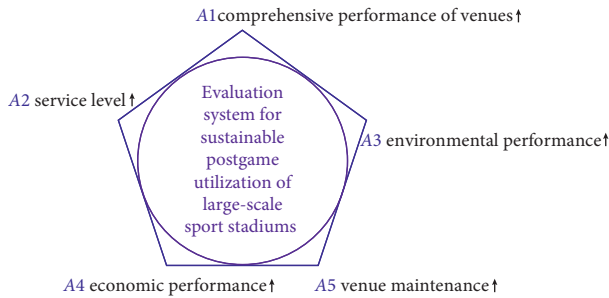


FIGURE 2: Evaluation index system for the sustainable use of large stadiums.

comprehensive performance, service level, environmental performance, and economic performance of the venue, through expert interviews and questionnaires (Figure 2).

**4.1. Comprehensive Performance of the Venue (A1).** A stadium is an arena for sports competitions, sports training, and public fitness needs. The comprehensive performance of the stadium is the key planning content before stadium construction, and it is the basis for the stadium to be fully utilized after a competition. Infrastructure is the material basis for the comprehensive performance of the venue, which can be measured in terms of facility scale, supporting facilities, safety facilities, and sanitation. Social environment is an external guarantee for the comprehensive performance of the venue, which includes traffic environment and population density. The management system is the core of the comprehensive performance of the venue, mainly reflected in the operation management mechanism, organizational structure, and talent support.

**4.2. Service Level (A2).** The service level of the stadium is the embodiment of the core value of the stadium, which can be evaluated from the aspects of utilization level and activity bearing.

Utilization level mainly reflects the utilization rate and openness of stadiums in sports events, training, and other activities. Therefore, it includes open area, opening hours, and number of receptions. Open area is measured by the proportion of the actual use of the stadium to the building area of the stadium. Open time conditions are measured using the weekly average opening hours of the stadium for service levels. The number of receptions is measured by the number of people in the stadium.

Activity bearing capacity mainly reflects the ability of sports venues to host large-scale events such as sports events, exhibitions, and cultural performances. It can be measured from the following four categories: sports events, sports training, public fitness activities, and commercial activities.

**4.3. Environmental Performance (A3).** Environmental performance is used to measure the quality of the natural environment of the stadium and the pollution and damage to the natural environment, including energy consumption, environmental load, environmental pollution, and indoor environmental quality. The energy consumption of stadiums mainly refers to the electricity and water consumed in the daily operation of stadiums. Environmental load indicates the influence of the building’s architectural body on the natural environment around the stadium, which mainly includes wind environment and groundwater system. Environmental pollution mainly refers to air pollution, water pollution, and solid waste generated by the daily operation of stadiums. Indoor environmental quality is a collection of various environmental factors in the interior of the stadium when people perform sports activities, including sound, heat, air, and light.

**4.4. Economic Performance (A4).** Economic performance is the most basic evaluation content for the postmatch use of large-scale sports stadiums. It refers to the economic benefits of sports and expenditures of various public or business activities. It can be measured by economic income and expenditure. The

economic income of stadiums and gymnasiums mainly comes from fiscal appropriation, fiscal subsidy, and commercial income, while the economic expenditure includes basic and project expenditure.

4.5. *Venue Maintenance (A5)*. Venue maintenance is the material basis for the sustainable use of stadiums. It is measured by two types of indicators: stadium maintenance funds and venues' fixed assets loss. The sixth national sports census showed that, as of the end of 2013, 90% of large stadiums in China adopted an independent business model. The short-term behavior of venues brought by the independent business model caused the loss and damage of some venues to increase sharply, and the maintenance costs also rose sharply. As a result, the fixed assets loss of a venue is a serious issue, so the use of depreciation and the proportion of new fixed assets is measured. In the management of public sports venues, China generally adopts the state financial allocation method to address the maintenance of houses, venues, and large-scale equipment in public sports venues [26]. However, the significant operating costs will inevitably result in a financial burden on relevant government departments. Therefore, it is necessary to operate stadiums through multiple channels [27]. Venue maintenance funds should be measured by financial maintenance funds input and other maintenance funds.

### 5. Weight Determinations of Sustainable Postmatch Use Evaluation Indicators of Large-Scale Stadiums

The Analytic Hierarchy Process (AHP) was introduced by T. L. Saaty, a professor at the University of Pittsburgh in the mid-1970s. Its basic idea is to decompose a complex problem into various components and then group these factors into dominance relationships to form an ordered hierarchical structure. The relative importance of the factors in the hierarchy is determined by means of a pairwise comparison, and then the judgment of the person is integrated to determine the overall order of the relative importance of the factors of the decision. AHP has made it convenient for decision makers to solve decision-making problems that are difficult to quantify, so its application covers almost any scientific field. In this study, eight field management experts were invited to create two points for the evaluation indicators. The steps for determining the weight of the first-level indicators are as follows:

- (1) Ask the eight experts to refer to the judgment matrix scale definition table (Table 1) to assign scores to the five first-level indicators and then statistically organize the assigned scores to obtain the judgment matrix A (Table 2). The scores of each indicator in the judgment matrix are obtained by the average of the scores assigned by the eight experts.

- (2) The largest eigenvalue  $\lambda_{\max}$  of the judgment matrix A and its corresponding normalized eigendirection are calculated as  $W = [W_1, W_2, W_3, W_4, W_5]^T$ , and eigenvector  $W = [W_1, W_2, W_3, W_4, W_5]^T$  is the weight vector of the evaluation unit, calculated by the geometric average method (square root method)  $\lambda_{\max}$  with W.
  - (i) From the formula  $M_i = \prod_{j=1}^n b_{ij}$  ( $i = 1, 2, \dots, n$ ), we can get

$$\begin{aligned}
 M_1 &= 1 \times \frac{13}{24} \times \frac{7}{12} \times \frac{5}{4} \times \frac{5}{4} = \frac{39}{79} = 0.49371, \\
 M_2 &= \frac{5}{2} \times 1 \times \frac{3}{2} \times \frac{5}{2} \times \frac{3}{2} = \frac{225}{16} = 14.0625, \\
 M_3 &= \frac{9}{4} \times \frac{3}{4} \times 1 \times \frac{9}{4} \times \frac{7}{8} = \frac{299}{90} = 3.32227, \\
 M_4 &= \frac{7}{8} \times \frac{5}{12} \times \frac{11}{24} \times 1 \times \frac{11}{24} = \frac{1}{13} = 0.07659, \\
 M_5 &= \frac{9}{4} \times \frac{7}{8} \times \frac{5}{4} \times \frac{9}{4} \times 1 = \frac{299}{54} = 5.53711.
 \end{aligned} \tag{1}$$

- (ii) Compute the  $n$ th root of  $M_i$ ,  $\bar{W}_i = \sqrt[n]{M_i}$ :

$$\begin{aligned}
 \bar{W}_1 &= \sqrt[5]{M_1} = \sqrt[5]{0.49371} = 0.86835, \\
 \bar{W}_2 &= \sqrt[5]{M_2} = \sqrt[5]{14.0625} = 1.69673, \\
 \bar{W}_3 &= \sqrt[5]{M_3} = \sqrt[5]{3.32227} = 1.27414, \\
 \bar{W}_4 &= \sqrt[5]{M_4} = \sqrt[5]{0.07659} = 0.59818, \\
 \bar{W}_5 &= \sqrt[5]{M_5} = \sqrt[5]{5.53711} = 1.40817.
 \end{aligned} \tag{2}$$

- (iii) Use the formula  $W_i = \bar{W}_i / \sum_{j=1}^n \bar{W}_j$  to normalize the vector  $\bar{W} = [\bar{W}_1, \bar{W}_2, \bar{W}_3, \bar{W}_4, \bar{W}_5]^T$ , calculate  $W_i$  separately, and get the eigenvector  $W = [W_1, W_2, W_3, W_4, W_5]^T$ :

$$\begin{aligned}
 W_1 &= \frac{0.86835}{5.84285} = 0.14862, \\
 W_2 &= \frac{1.69673}{5.84285} = 0.29039, \\
 W_3 &= \frac{1.27414}{5.84285} = 0.2176, \\
 W_4 &= \frac{0.59818}{5.84285} = 0.10238, \\
 W_5 &= \frac{1.40817}{5.84285} = 0.24101.
 \end{aligned} \tag{3}$$

So, the feature vector is  $W = [0.14862, 0.29039, 0.2176, 0.10238, 0.24102]^T$ .

- (iv) The largest eigenvalue  $\lambda_{\max}$  of the judgment matrix  $A$  is calculated by the formula,  $(AW)_i$  is the factor of the vector  $AW$ :

$$\begin{aligned}
 (AW)_1 &= 1 \times 0.14862 + \frac{13}{24} \times 0.29039 + \frac{7}{12} \times 0.2176 \\
 &\quad + \frac{5}{4} \times 0.10238 + \frac{5}{4} \times 0.24101 = 0.86208, \\
 (AW)_2 &= \frac{5}{2} \times 0.14862 + 1 \times 0.29039 + \frac{3}{2} \times 0.2176 \\
 &\quad + \frac{5}{2} \times 0.10238 + \frac{3}{2} \times 0.24101 = 1.6058, \\
 (AW)_3 &= \frac{9}{4} \times 0.14862 + \frac{3}{4} \times 0.29039 + 1 \times 0.2176 \\
 &\quad + \frac{9}{4} \times 0.10238 + \frac{7}{8} \times 0.24101 = 1.21102, \\
 (AW)_4 &= \frac{7}{8} \times 0.14862 + \frac{5}{12} \times 0.29039 + \frac{11}{24} \times 0.2176 \\
 &\quad + 1 \times 0.10238 + \frac{11}{24} \times 0.24101 = 0.56361, \\
 (AW)_5 &= \frac{9}{4} \times 0.14862 + \frac{7}{8} \times 0.29039 + \frac{5}{4} \times 0.2176 \\
 &\quad + \frac{9}{4} \times 0.10238 + 1 \times 0.24101 = 1.33185, \\
 \lambda_{\max} &= \sum_{i=1}^n \frac{(AW)_i}{5W_i} = \frac{(AW)_1}{5W_1} + \frac{(AW)_2}{5W_2} + \frac{(AW)_3}{5W_3} \\
 &\quad + \frac{(AW)_4}{5W_4} + \frac{(AW)_5}{5W_5} \\
 &= \frac{0.86208}{5 \times 0.14862} + \frac{1.6058}{5 \times 0.29039} + \frac{1.21102}{5 \times 0.2176} \\
 &\quad + \frac{0.56361}{5 \times 0.10238} + \frac{1.33285}{5 \times 0.24101} = 5.28541.
 \end{aligned} \tag{4}$$

- (3) The consistency test of the judgment matrix  $A$  is carried out by the formula  $C.R. = C.I./R.I.$ :

Consistency index  $C.I. = (\lambda_{\max} - n)/(n - 1) = (5.28541 - 5)/(5 - 1) = 0.07135$ , Average random consistency index  $R.I. = 1.12$ .

$C.R. = (C.I./R.I.) = (0.07135/1.12) = 0.06371 < 0.10$ . It can be considered that the judgment matrix  $A$  passes the consistency test.

Therefore, the weights of  $A_1, A_2, A_3, A_4, A_5$  are, respectively, 0.14862, 0.29039, 0.2176, 0.10238, and 0.24101.

Similarly, according to the above five steps, the weights of second and third level indicators at a single level can be calculated, and finally, the combined weights at each level can be calculated according to the weights of the indicators at each single level (Table 3).

## 6. Research on Fuzzy Comprehensive Evaluation Model for Sustainable Postmatch Use of Large-Scale Sports Venues

Large-scale sports venues in China are not only used for large-scale events but are also included in the Chinese sports stadium system for postgame utilization and to continue to serve the public. Large-scale sports stadiums serve the general public in competition, training, fitness, and watching, so the design and construction of sports venues, postgame operating modes, and other factors are fundamental factors for the sustainable postgame use of large stadiums. In the sustainable postgame use evaluation index system of large-scale sports stadiums, there are both quantitative and qualitative indicators. Quantitative indicators can be obtained by collecting relevant data, while qualitative indicators cannot be accurately measured, as they reflect the intuitive feelings of the service objects. Therefore, the membership degree of quantitative indicators can be determined by collecting relevant data, and the membership degree of qualitative indicators can be determined by the multiple levels of fuzzy comprehensive evaluation method.

### 6.1. Indicator Evaluation Corpus

**6.1.1. Quantitative Indicator Evaluation Corpus.** Among the first-level indicators of the sustainable postgame use evaluation index system for large-scale sports stadiums, only the  $A_2$  service level can be determined by collecting relevant data. The evaluation corpus adopts the "Comprehensive Evaluation System for Large-Scale Sports Facilities Operation Management" promulgated by the State Sports General Administration of China (Table 4).

**6.1.2. Qualitative Indicator Evaluation Corpus.** The qualitative evaluation index of the evaluation index system for the sustainable use of large stadiums is {excellent, good, pass, poor, very poor}; the corresponding evaluation value set  $V$  is {1.0, 0.8, 0.6, 0.4, 0.2}.

**6.2. Evaluation Object and Research Object Selection.** A case study was conducted using the Bengbu Sports Center in Anhui province, China. The construction of the stadium was completed in 2018, and it hosted the 14th Anhui Provincial Games in October of the same year. It has been a year since the hosting of this large-scale event. Therefore, an evaluation of the continuous use of Bengbu

TABLE 1: Judgment matrix scale definition table.

Scale	Definition
1	Indicates that $A_i$ is of equal importance to $A_j$
3	Indicates that $A_i$ is slightly more important than $A_j$
5	Indicates that the indicator $A_i$ is more important than the indicator $A_j$
7	Indicates that the indicator $A_i$ is significantly more important than the latter $A_j$
9	Indicates that the indicator $A_i$ is extremely important compared to the indicator $A_j$
2, 4, 6, 8	Represent the intermediate values of the above adjacent judgment
Else	If the ratio of the importance of indicators $A_i$ to indicators $A_j$ is $a_{ij}$ , then the ratio of the importance of indicators to indicators is $a_{ji} = 1/a_{ij}$

TABLE 2: Level 1 indicator judgment matrix  $A$ .

$A$	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$
$A_1$	1	13/24	7/12	5/4	5/4
$A_2$	5/2	1	3/2	5/2	3/2
$A_3$	9/4	3/4	1	9/4	7/8
$A_4$	7/8	5/12	11/24	1	11/24
$A_5$	9/4	7/8	5/4	9/4	1

Sports Center can play a very important reference role. The subjective evaluation of qualitative indicators was carried out through on-site stratified random distribution and questionnaire recovery. Sports trainers, sports staff, venue managers, audiences, general fitness people, and other people were selected for the questionnaire survey. A total of 200 questionnaires were distributed, and 192 copies were collected, of which, 185 were valid questionnaires. Respondents' basic information is shown in Table 5.

### 6.3. Indicator Score Calculation

6.3.1. *Quantitative Index Score Calculation.* By examining the "Bengbu Sports Center 2018 Opening Work Plan," the "2018 City Sports Bureau Government Information Disclosure Annual Work Report," and other related documents, the membership degree of the three-level evaluation index of the stadium's service level was determined, and it was combined with the weights of the evaluation indicators at each level, calculating the scores of the indicators at each level.

6.3.2. *Calculation of Qualitative Indicator Score.* Due to the professionalism of index evaluation, some indicators are evaluated by specific survey targets in the indicator system. Therefore, the  $A_3$  environmental performance

evaluation index was selected as an example for analysis. This index was investigated in the 185 valid questionnaires.

- (i) Determining the degree of membership of the three levels of environmental performance indicators

The 185 valid questionnaires were counted according to the grades of qualitative indicators and normalized, and the proportion of people selected for each evaluation level was obtained, which is the membership degree of environmental performance of the three-level indicators (Table 6).

- (ii) Constructing the fuzzy membership relation matrix of the environmental performance of the three-level index:

$$\begin{aligned}
 R_{31} &= \begin{bmatrix} 0.141 & 0.173 & 0.373 & 0.238 & 0.076 \\ 0.092 & 0.249 & 0.427 & 0.173 & 0.065 \end{bmatrix}, \\
 R_{32} &= \begin{bmatrix} 0.103 & 0.205 & 0.265 & 0.205 & 0.222 \\ 0.189 & 0.184 & 0.405 & 0.141 & 0.081 \end{bmatrix}, \\
 R_{33} &= \begin{bmatrix} 0.157 & 0.222 & 0.319 & 0.200 & 0.103 \\ 0.135 & 0.173 & 0.378 & 0.157 & 0.157 \\ 0.146 & 0.222 & 0.254 & 0.184 & 0.195 \end{bmatrix}, \\
 R_{34} &= \begin{bmatrix} 0.168 & 0.265 & 0.222 & 0.195 & 0.151 \\ 0.157 & 0.227 & 0.265 & 0.168 & 0.184 \\ 0.151 & 0.211 & 0.330 & 0.195 & 0.114 \\ 0.168 & 0.157 & 0.389 & 0.184 & 0.103 \end{bmatrix}.
 \end{aligned} \tag{5}$$

- (iii) Establishing the three-level fuzzy comprehensive evaluation model of environmental performance:



TABLE 3: Evaluation index system for sustainable postgame utilization of large-scale sports stadiums.

First class indicator <i>A</i>		Second class indicator <i>B</i>		Third class indicator <i>C</i>	
Index	Weight	Index	Weight	Index	Weight
A1 comprehensive performance of venues	0.14862	B11 infrastructure	0.05945	C111 facilities size	0.01783
				C112 supporting facilities	0.01783
				C113 safety facilities	0.01189
				C114 health facility	0.01189
		B12 social environment	0.04459	C121 traffic environment	0.02229
				C122 density of population	0.02229
				C131 the mechanism of operation and management	0.01783
		B13 management system	0.04459	C132 organization structure	0.01338
				C133 talent support	0.01338
				A2 service level	0.29039
C212 opening hours	0.04356				
C213 the number of reception	0.05808				
C221 sports event	0.04356				
B22 organized activities	0.14520	C222 exercise training	0.04356		
		C223 public fitness program	0.03630		
		C224 commercial activity	0.02178		
		A3 environmental performance	0.2176		
C312 water	0.02176				
B32 environmental load	0.04352			C321 wind environment	0.02176
				C322 underground water	0.02176
B33 environmental pollution	0.06528			C331 atmosphere	0.01632
				C332 water pollution	0.02285
				C333 solid waste	0.02611
				C341 sound	0.01278
B34 indoor environmental quality	0.06528	C342 temperature	0.01331		
		C343 air	0.01472		
		C344 light	0.01464		
		A4 economic performance	0.10238	B41 economic income	0.05119
C412 revenue from fiscal subsidies	0.01536				
B42 economic expenditure	0.05119			C413 business revenue	0.02559
				C421 basic expenditure	0.02559
				C422 project expenditure	0.02559
				A5 venue maintenance	0.24101
C512 other maintenance expenses	0.07230				
B52 loss of stadium fixed assets	0.12051	C521 fixed assets depreciation and new proportion	0.12051		

TABLE 4: Service level indicator evaluation.

Index	Excellent/1.0	Good/0.8	Pass/0.6	Poor/0.4	Very poor/0.2
C211	80%–100%	70%–79%	60%–69%	50%–59%	40%–49%
C212	50 hours per week or above	46–50 h	41–45 h	36–40 h	Less than 35 h
C213	Over 100,000 times	80,000 to 90,000 times	60,000 to 70,000 times	40,000 to 50,000 times	Less than 40,000 times
C221	More than 10 times	8 to 9 times	6 to 7 times	3 to 5 times	Less than 3 times
C222	More than 10 times	8 to 9 times	6 to 7 times	3 to 5 times	1 to 2 times
C223	More than 10 times	8 to 9 times	6 to 7 times	3 to 5 times	1 to 2 times
C224	More than 10 times	8 to 9 times	6 to 7 times	3 to 5 times	1 to 2 times

TABLE 5: Basic information of the survey object.

Property	Proportion (%)	Age	Proportion (%)	Schooling	Proportion (%)
Sports trainers	14.05	Under 18	11.35	Below senior high school	7.03
Sports staff	23.24	18–30	28.11	Senior high school	16.76
Venue managers	8.64	30–45	30.27	Technical secondary school	33.51
Audiences	22.16	45–60	17.3	Bachelor	35.68
General fitness people	31.91	At least 60	12.97	Graduate or above	7.02

TABLE 6: Degree of subordination of environmental performance of large stadiums.

Index	Evaluation									
	Excellent		Good		Pass		Poor		Very poor	
	Number	Proportion	Number	Proportion	Number	Proportion	Number	Proportion	Number	Proportion
C311	26	0.141	32	0.173	69	0.373	44	0.238	14	0.076
C312	17	0.092	46	0.249	79	0.427	32	0.173	12	0.065
C321	19	0.103	38	0.205	49	0.265	38	0.205	41	0.222
C322	35	0.189	34	0.184	75	0.405	26	0.141	15	0.081
C331	29	0.157	41	0.222	59	0.319	37	0.200	19	0.103
C332	25	0.135	32	0.173	70	0.378	29	0.157	29	0.157
C333	27	0.146	41	0.222	47	0.254	34	0.184	36	0.195
C341	31	0.168	49	0.265	41	0.222	36	0.195	28	0.151
C342	29	0.157	42	0.227	49	0.265	31	0.168	34	0.184
C343	28	0.151	39	0.211	61	0.330	36	0.195	21	0.114
C344	31	0.168	29	0.157	72	0.389	34	0.184	19	0.103

$$\begin{aligned}
 B_{31} &= (0.02176 \ 0.02176) \times \begin{bmatrix} 0.141 & 0.173 & 0.373 & 0.238 & 0.076 \\ 0.092 & 0.249 & 0.427 & 0.173 & 0.065 \end{bmatrix} = (0.00507 \ 0.00918 \ 0.01741 \ 0.00894 \ 0.00307), \\
 B_{32} &= (0.02176 \ 0.02176) \times \begin{bmatrix} 0.103 & 0.205 & 0.265 & 0.205 & 0.222 \\ 0.189 & 0.184 & 0.405 & 0.141 & 0.081 \end{bmatrix} = (0.00635 \ 0.00846 \ 0.01458 \ 0.00753 \ 0.00659), \\
 B_{33} &= (0.01632 \ 0.02285 \ 0.02611) \times \begin{bmatrix} 0.157 & 0.222 & 0.319 & 0.200 & 0.103 \\ 0.135 & 0.173 & 0.378 & 0.157 & 0.157 \\ 0.146 & 0.222 & 0.254 & 0.184 & 0.195 \end{bmatrix} = (0.00946 \ 0.01337 \ 0.02048 \ 0.01166 \ 0.01036), \\
 B_{34} &= (0.0128 \ 0.0133 \ 0.0147 \ 0.0146) \times \begin{bmatrix} 0.168 & 0.265 & 0.222 & 0.195 & 0.151 \\ 0.157 & 0.227 & 0.265 & 0.168 & 0.184 \\ 0.151 & 0.211 & 0.330 & 0.195 & 0.114 \\ 0.168 & 0.157 & 0.389 & 0.184 & 0.103 \end{bmatrix} = (0.0089 \ 0.0118 \ 0.0169 \ 0.0103 \ 0.0076).
 \end{aligned}
 \tag{6}$$

(iv) Establishing a fuzzy comprehensive evaluation model of environmental performance

Based on the fuzzy comprehensive evaluation results of the three-level indicators of environmental performance, the fuzzy membership relation matrix of two-level indexes of environmental performance was constructed; the fuzzy comprehensive

evaluation model of two-level indexes of environmental performance was established by combining the weight vector of two-level indexes of environmental performance.

$$\begin{aligned}
 A_3 &= (0.04352 \ 0.04352 \ 0.06528 \ 0.06528) \times \begin{bmatrix} 0.00507008 & 0.00918272 & 0.01740800 & 0.00894336 & 0.00306816 \\ 0.00635392 & 0.00846464 & 0.01457920 & 0.00752896 & 0.00659328 \\ 0.00945905 & 0.01337251 & 0.02047532 & 0.01165569 & 0.01035986 \\ 0.00891895 & 0.01181247 & 0.01691687 & 0.01029234 & 0.00756484 \end{bmatrix} \\
 &= (0.001696888 \ 0.002412089 \ 0.003833045 \ 0.002149643 \ 0.001590589).
 \end{aligned}
 \tag{7}$$

TABLE 7: Evaluation scores of sustainable postmatch use of Bengbu Sports Center.

First class indicator <i>A</i>		Second class indicator <i>B</i>		Third class indicator <i>C</i>	
Index	Score	Index	Score	Index	Score
<i>A1</i> comprehensive performance of venues	0.6609	<i>B11</i> infrastructure	0.6976	C111 facilities size	0.6973
				C112 supporting facilities	0.6962
				C113 safety facilities	0.6811
		<i>B12</i> social environment	0.6611	C114 health facility	0.7168
				C121 traffic environment	0.6627
				C122 density of population	0.6595
		<i>B13</i> management system	0.5958	C131 the mechanism of operation and management	0.6001
				C132 organization structure	0.5978
				C133 talent support	0.5881
<i>A2</i> service level	0.725	<i>B21</i> utilization	0.78	C211 open area	0.8
				C212 opening hours	1.0
		<i>B22</i> organized activities	0.67	C213 the number of reception	0.6
				C221 sports event	0.8
				C222 exercise training	1.0
				C223 public fitness program	0.4
				C224 commercial activity	0.2
<i>A3</i> environmental performance	0.6081	<i>B31</i> energy consumption	0.6194	C311 electricity	0.6136
				C312 water	0.6296
		<i>B32</i> environmental load	0.6021	C321 wind environment	0.5524
				C322 underground water	0.6518
		<i>B33</i> environmental pollution	0.5997	C331 atmosphere	0.6266
				C332 water pollution	0.5944
				C333 solid waste	0.5886
				C341 sound	0.6214
		<i>B34</i> indoor environmental quality	0.6152	C342 temperature	0.6016
				C343 air	0.6186
C344 light	0.6212				
<i>A4</i> economic performance	0.5681	<i>B41</i> economic income	0.5244	C411 revenue from appropriation	0.6136
				C412 revenue from fiscal subsidies	0.6949
		<i>B42</i> economic expenditure	0.6118	C413 business revenue	0.3864
				C421 basic expenditure	0.6102
				C422 project expenditure	0.6136
<i>A5</i> venue maintenance	0.6902	<i>B51</i> expenditure for venue maintenance	0.6583	C511 fiscal investment	0.6441
				C512 other maintenance expenses	0.6678
		<i>B52</i> loss of stadium fixed assets	0.7220	C521 fixed assets depreciation and new proportion	0.7220
Total value: 0.6656					

(v) Calculating the fuzzy comprehensive evaluation results of environmental performance *A3* was normalized:

$$A3' = (0.145253476 \ 0.206474624 \ 0.328108343 \ 0.184009267 \ 0.13615429). \tag{8}$$

The environmental performance evaluation score is

$$T3 = A3' \times V = (0.1453 \ 0.2065 \ 0.3281 \ 0.1840 \ 0.1362) \times \begin{bmatrix} 1.0 \\ 0.8 \\ 0.6 \\ 0.4 \\ 0.2 \end{bmatrix} = 0.608132746. \tag{9}$$

Similarly, according to the above five steps, we conducted fuzzy comprehensive evaluation on A1 venue comprehensive performance, A4 economic performance, A5 venue maintenance, and other first-level indicators. The evaluation scores are shown in Table 7.

6.3.3. *Fuzzy Comprehensive Evaluation of Sustainable Postmatch Utilization of Large Stadiums.* Based on the quantitative and qualitative index evaluation scores calculated in the above steps, the first-level index evaluation score vector was constructed, and combined with the first-level index weight, the final overall comprehensive evaluation score was obtained.

$$A = (0.6609 \ 0.725 \ 0.6081 \ 0.5681 \ 0.6902) \times \begin{bmatrix} 0.14862 \\ 0.29039 \\ 0.2176 \\ 0.10238 \\ 0.24101 \end{bmatrix} = 0.6656. \tag{10}$$

According to the evaluation results, the evaluation score of the sustainable use of Bengbu Sports Center in Anhui province, China, is 0.6656, which is between 0.8 and 0.6; thus, it can be judged that the sustainable postmatch use of the stadium is “good.” Among the first-level index evaluation scores, service level has the highest evaluation score, reaching 0.725, mainly due to the success of the 14th Anhui Provincial Games, bringing a series of small- and medium-sized competitions to make use of the stadium. The activity bearing capacity of the stadium has been effective, but there is room for improvement, and the development of the stadium in terms of commercial activities needs to be further strengthened. The economic performance evaluation score is the lowest, which belongs to the “pass” level. The main reason for this is that business income is small. The stadium has not been market-oriented after the game, and the operation management mode is lagging. It is thus necessary to introduce market mechanism and increase business income. The above evaluation score results are basically consistent with those of other researchers’ qualitative analysis, which is

in line with the current situation of the use of large-scale sports stadiums in China.

### 7. Conclusions

- (1) The evaluation system for the sustainable utilization of large-scale stadiums and gymnasiums in China is to measure the operation and performance of stadiums and gymnasiums, which is constructed by five level indicators, namely, comprehensive performance, service level, environmental performance, economic performance, and maintenance of stadiums and gymnasiums
- (2) Based on the fuzzy comprehensive evaluation model, the paper evaluates Bengbu Sports Center in Anhui province, China, and concludes that the sustainable utilization of the stadium is a “good” grade
- (3) At present, the postcompetition utilization of large stadiums and gymnasiums in China has some problems, such as outdated system and mechanism, insufficient development power, single industrial chain, and no mature systematic operation

### Data Availability

Data are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This work was supported by the Anhui Province Philosophy and Social Sciences Planning Project (no. AHSKY2016D63), Anhui Education Department Teaching and Research Fund Project (no. 2018jyxm1305), Key Teaching and Research Project of Anhui University of Finance and Economics “The Construction of Teaching Competition” Double Core Teaching Mode of Public Physical Education in Colleges and Universities (acjyzd2019012), and Anhui University of Finance and Economics School-Level Teaching and Research Fund Project (acxkjsjy201803zd and acjyyb2018006).

## References

- [1] China Sport General Administration, "The sixth China sports census data bulletin," 2014, <http://www.sport.gov.cn/n16/n1077/n1467/n3895927/n4119307/7153937.html>.
- [2] J. Dai, S. Zhang, Y. Tang, X. J. Guo, and Z. H. Ma, "Value orientation and path selection of public sports service system innovation in governance context," *China Sports Science*, vol. 35, no. 11, pp. 3–12, 2015.
- [3] C. Yuanxin and W. Jian, "Current status, constraints and countermeasures of the operation of large-scale sports venues in China," *Journal of Shanghai Sport University*, vol. 34, no. 5, pp. 17–21, 2010.
- [4] China Sport General Administration, "The 13th five-year plan for China sports development," 2016, <http://www.sport.gov.cn/n316/n340/c723004/content.html>.
- [5] M. Manni, A. Petrozzi, V. Coccia, A. Nicolini, and C. Franco, "Investigating alternative development strategies for sport arenas based on active and passive systems," *Journal of Building Engineering*, vol. 31, Article ID 101340, 2020.
- [6] G. Johnson, "The economic impact of new stadiums and arenas on cities," *University of Denver Sports & Entertainment Law Journal*, vol. 230, pp. 3–40, 2011.
- [7] P. Fabris, "Sports stadiums and arenas showcase sustainability features," *Building Design & Construction*, [https://schlr.cnki.net/Detail/index/SPQD\\_01/SPQDCAF0D3A8DC369B73A27EE0ED25B3A95B](https://schlr.cnki.net/Detail/index/SPQD_01/SPQDCAF0D3A8DC369B73A27EE0ED25B3A95B), 2018.
- [8] A. Thornley, "Urban regeneration and sports stadia," *European Planning Studies*, vol. 10, no. 7, pp. 813–818, 2002.
- [9] L. Roger and R. L. Kemp, *Cities and Sports Stadium:A Planning Handbook*, Mc Farland Press, Jefferson, NC, USA, 2009.
- [10] B. A. Griffin, C. Eibner, C. E. Bird et al., "The relationship between urban sprawl and coronary heart disease in women," *Health & Place*, vol. 20, no. 3, pp. 51–61, 2013.
- [11] I. Jewell and A. G. Leblanc, "Systematic review of the health benefits of physical activity and fitness in school-aged children and youth," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 7, no. 1, pp. 1–16, 2010.
- [12] M. V. D. Hurk and K. Verhoest, "On the fast track? Using standard contracts in public-private partnerships for sports facilities: a case study," *Sport Management Review*, vol. 20, pp. 226–239, 2017.
- [13] J. Norman, H. L. Maclean, and C. A. Kennedy, "Comparing high and low residential density: life-cycle analysis of energy use and greenhouse gas emissions," *Journal of Urban Planning and Development*, vol. 132, no. 1, pp. 10–21, 2006.
- [14] S. Winkelman, A. Bishins, and C. Kooshian, "Planning for economic and environmental resilience," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 8, pp. 575–586, 2010.
- [15] J. Burke and J. Shang, "Post event sustainability: London 2012 Olympics," *World Archit*, vol. 8, pp. 92–95, 2015.
- [16] B. Hanway and S. Sun, "London 2012 olympic park masterplan and legacy," *Landscape Architecture*, vol. 3, pp. 102–110, 2012.
- [17] C. Santo, "The economic impact of sports stadiums: recasting the analysis in context," *Journal of Urban Affairs*, vol. 27, no. 2, pp. 177–192, 2010.
- [18] B. R. Muijsson, "Increasing profitability of multifunctional sport stadiums: a research on using a decision support model to increase profitability and feasibility of multifunctional sport stadiums," 2015, <http://resolver.tudelft.nl/uuid:145856b5-487b-4a66-8cca-9871228f01fe>.
- [19] Y. X. Ding and L. Zhang, "Application research on TOT model used to public sports stadium under the background of ownership and operation separation," *Journal of Beijing Sport University*, vol. 40, no. 2, pp. 24–29, 2017.
- [20] A. Amna and F. Salim, "Sports arena development: scalability impact on urban fabric integration," *Architecture Research*, vol. 6, no. 6, 2016.
- [21] J. L. Crompton, D. R. Howard, and T. Var, "Financing major league facilities: status, evolution and conflicting forces," *Journal of Sport Management*, vol. 17, no. 2, pp. 156–184, 2010.
- [22] L. Rebeggiani, "Public vs. private spending for sports facilities—the case of Germany," *Journal of Public Finance and Management*, vol. 6, no. 3, pp. 395–435, 2006.
- [23] H. Westerbeek, P. Emery, and P. Turner, *Managing Sport Facilities and Major Events*, Routledge, New York, NY, USA, 2006.
- [24] Y. Hongyan and Z. Xu, "Research on sustainable building innovation strategy based on life cycle analysis," *Architecture Science*, vol. 8, pp. 9–13, 2007.
- [25] R. Apanavičienė, D. Ala, T. Baltramaitis, and V. Maliene, "Sustainability aspects of real estate development: Lithuanian case study of sports and entertainment arenas," *Sustainability*, vol. 7, no. 6, 2015.
- [26] Y. I. Li, "Research on sustainable building innovation strategy for sports venues in China," *Journal of Beijing Sport University*, vol. 36, no. 03, pp. 39–43, 2013.
- [27] L. Lei, S. H. Xiao, Q. Fu, and L. Chen, "Operation management of large-scale sports venues in China: model selection and path arrangement," *Journal of Beijing Sport University*, vol. 36, no. 10, pp. 10–15, 2013.

## Research Article

# Several Asymptotic Bounds on the Balaban Indices of Trees

Bo Deng <sup>1,2,3,4</sup>, Chengfu Ye,<sup>1</sup> Weilin Liang,<sup>1</sup> Yalan Li,<sup>5</sup> and Xueli Su<sup>1</sup>

<sup>1</sup>School of Mathematics and Statistics, Qinghai Normal University, Xining 810001, China

<sup>2</sup>Academy of Plateau, Science and Sustainability, Xining, Qinghai 810008, China

<sup>3</sup>Key Laboratory of Tibetan Information Processing, Ministry of Education, Xining, Qinghai Province, China

<sup>4</sup>Tibetan Intelligent Information Processing and Machine Translation Key Laboratory, Qinghai 810008, China

<sup>5</sup>School of Computer, Qinghai Normal University, Xining 810001, China

Correspondence should be addressed to Bo Deng; dengbo450@163.com

Received 16 August 2020; Revised 30 September 2020; Accepted 21 October 2020; Published 6 November 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Bo Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Balaban index (also called the  $J$  index) of a connected graph  $G$  is a distance-based topological index, which has been successfully used in various QSAR and QSPR modeling. Although the index was introduced 30 years ago, there are few results on the asymptotic relations. In this paper, several asymptotic bounds on the Balaban indices of trees with diameters 3 and 4 are shown, respectively.

## 1. Introduction

All graphs considered in this paper are simple and undirected. Let  $G$  be a graph with its edge set  $E(G)$  and vertex set  $V(G)$ . We set  $|V(G)| = n$  and  $|E(G)| = m$ . The star of order  $n$  is denoted by  $S_n$ . The distance between vertices  $u$  and  $v$  in  $G$  is denoted by  $d_G(u, v)$ , and the sum of the distance between vertex  $u$  and each vertex of  $G$  is denoted by  $\sigma_G(u)$ , that is,  $\sigma_G(u) = \sum_{w \in V(G)} d_G(u, w)$ .

The Balaban index [1] of a connected graph  $G$  (or the  $J$  index for short) is defined as

$$J(G) = \frac{m}{\mu + 1} \sum_{uv \in E(G)} \frac{1}{\sqrt{\sigma_G(u)\sigma_G(v)}}, \quad (1)$$

where  $\mu$  is the cyclomatic number and  $\mu = m - n + 1$ .

The Balaban index (also called the  $J$  index) of a connected graph  $G$  is a distance-based topological index, which has been successfully used in various QSAR and QSPR modeling [2, 3]. Many applications in chemistry can be found in [4–6]. By comparing with the Wiener index regarding alkanes in [7], it was found that the Balaban index reduces the degeneracy of the latter index and provides much higher discriminating ability.

So, the Balaban index is also called the sharpened Wiener index. Some results on the maximal and minimal Balaban index [8–10] have been presented. In [11–14], the asymptotic behaviors of the Balaban indices for various infinite families of graphs are observed.

Until now, there are few results on the asymptotic relations on the Balaban index. In this paper, several asymptotic bounds on the Balaban indices of trees with diameters 3 and 4 are shown, respectively. The two kinds of trees are depicted as follows.

If a tree is with diameter 3, then this tree can be obtained by attaching some pendent edges to the two end-vertices of one edge. Then, this tree is denoted by  $T_n(3, a, b)$ , see Figure 1, which has  $n$  vertices, diameter 3, and satisfies that there are  $a$  pendent edges attached at one end-vertex of one edge and  $b$  pendent edges attached at the other end-vertex of the edge, where  $a \geq 1$ ,  $b \geq 1$ , and  $a + b = n - 2$ . The set of this kind of trees is denoted by  $\mathcal{T}_n(3, a, b)$ .

The tree with order  $n$  and diameter 4 denoted by  $T_n^l(4, a_1, a_2, \dots, a_l)$ , see Figure 2, is obtained from a star  $S_{l+1}$  by attaching  $a_1, a_2, \dots, a_l$  pendent edges to the  $l$  pendent vertices of the star, respectively, where  $a_i \geq 0$ ,  $1 \leq i \leq l$ . The set of this kind of trees is denoted by  $\mathcal{T}_n^l(4, a_1, a_2, \dots, a_l)$ .



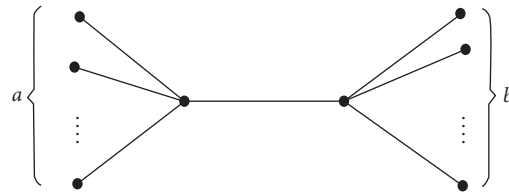


FIGURE 1: The graph  $T_n(3, a, b)$ .

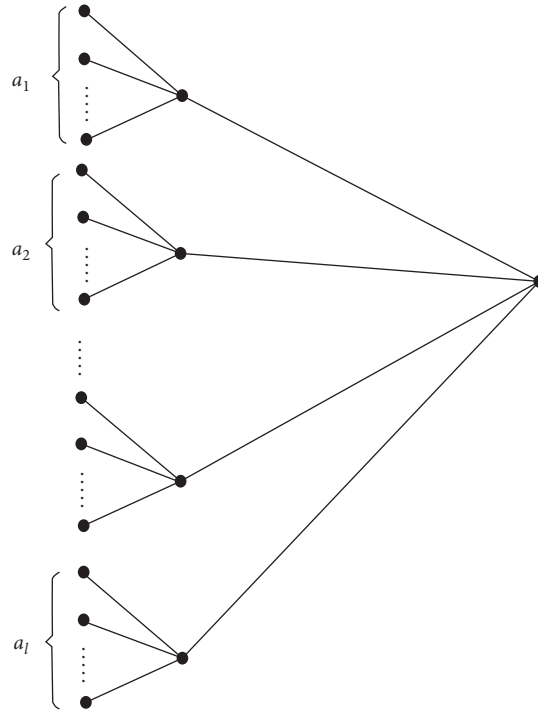


FIGURE 2: The graph  $T_n^l(4, a_1, a_2, \dots, a_l)$ .

### 2. The Balaban Indices of Trees with Diameter 3

In this section, the asymptotic bounds on the Balaban indices of trees with diameter 3 will be given.

**Theorem 1.** For any tree  $T_n(3, a, b) \in \mathcal{T}_n(3, a, b)$ , we have

$$O\left(\frac{2n}{\sqrt{15}}\right) \leq J(T_n(3, a, b)) \leq O\left(\frac{n}{\sqrt{2}}\right). \quad (2)$$

*Proof.* Suppose  $a \leq b$ . Then, by  $a + b = n - 2$ , we have  $a \leq ((n - 2)/2)$ . By direct calculation, the Balaban index of  $T_n(3, a, b) = T_n((3, a, n - a - 2))$  is as follows:

---


$$J(T_n(3, a, n - a - 2)) = (n - 1) \cdot \frac{a}{\sqrt{(3n - a - 5)(2n - a - 3)}} + \frac{1}{\sqrt{(2n - a - 3)(n + a - 1)}} + \frac{n - a - 2}{\sqrt{(2n + a - 3)(n + a - 1)}} \quad (3)$$

Through computation, the derivative of  $J(T_n(3, a, n - a - 2))$  related to  $a$  is

$$\frac{\partial J(T_n(3, a, n - a - 2))}{\partial a} = \frac{n - 1}{2} \cdot \left[ \frac{a(5n - 2a - 8)}{((3n - a - 5)(2n - a - 3))^{3/2}} + \frac{2}{\sqrt{(3n - a - 5)(2n - a - 3)}} + \frac{2 + 2a - n}{((2n - a - 3)(n + a - 1))^{3/2}} - \frac{2}{\sqrt{(n + a - 1)(2n + a - 3)}} + \frac{(2 + a - n)(3n + 2a - 4)}{((2n + a - 3)(n + a - 1))^{3/2}} \right]. \tag{4}$$

Next, the sign of  $(\partial J(T_n(3, a, n - a - 2)))/\partial a$  will be determined. For the above equation, we see that the first two terms are positive, and the last three terms are nonpositive.

Then, we use the sum of the first two terms to divide the sum of the absolute values of the last three terms. So, we get that

$$\frac{(a(5n - 2a - 8)/((3n - a - 5)(2n - a - 3))^{3/2}) + (2/(\sqrt{(3n - a - 5)(2n - a - 3)}))}{n - 2a - 2/((2n - a - 3)(n + a - 1))^{3/2} + 2/\sqrt{(n + a - 1)(2n + a - 3)} + (n - a - 2)(3n + 2a - 4)/((2n + a - 3)(n + a - 1))^{3/2}} = \frac{12n^2 - 38n + a(8 - 5n) + 30}{((3n - a - 5)(2n - a - 3))^{3/2}((n - 2a - 2)/((n + a - 1)(2n - a - 3))^{3/2}) + (2/\sqrt{(n + a - 1)(2n + a - 3)}) + ((n - a - 2)(3n + 2a - 4)/((2n + a - 3)(n + a - 1))^{3/2})}. \tag{5}$$

Denote the above fraction by  $f(n, a)$ . Observing this quotient, its value is equal to 1 if  $n = 2k + 2$  and  $a = k$  for  $k = 2, 3, \dots, \lfloor n - 2/2 \rfloor$ . In this case, it means that  $T_n(3, a, n - a - 2)$  attains the minimum value. Otherwise, its value is less than 1 as  $n$  tends to infinity, that is,

$$\lim_{n \rightarrow \infty} f(n, a) = \frac{4}{7\sqrt{3}} < 1. \tag{6}$$

Thus, the derivative of  $J(T_n(3, a, n - a - 2))$  related to  $a$  is negative as  $n$  is big enough. In this case,  $J(T_n(3, a, n - a - 2))$  increases along with parameter  $a$  decreasing. Since  $1 \leq a \leq \lfloor (n - 2)/2 \rfloor$ , we see that

$$J(T_n(3, a, n - a - 2)) \leq J(T_n(3, 1, n - 3)),$$

$$J(T_n(3, a, n - a - 2)) \geq J(T_n(3, \lfloor (n - 2)/2 \rfloor, \lceil (n - 2)/2 \rceil)). \tag{7}$$

By calculation, we get

$$J(T_n(3, 1, n - 3)) = O\left(\frac{n}{\sqrt{2}}\right), \tag{8}$$

$$J\left(T_n\left(3, \lfloor (n - 2)/2 \rfloor, \lceil (n - 2)/2 \rceil\right)\right) = O\left(\frac{2n}{\sqrt{15}}\right).$$

Hence, we obtain

$$O\left(\frac{2n}{\sqrt{15}}\right) \leq J(T_n(3, a, b)) \leq O\left(\frac{n}{\sqrt{2}}\right). \tag{9}$$

□

### 3. The Balaban Indices of Trees with Diameter 4

In this section, we present some asymptotic bounds for the Balaban indices of trees with diameter 4.

**Theorem 2.** For any tree  $T_n^l(4, a_1, a_2, \dots, a_l) \in \mathcal{T}_n^l(4, a_1, a_2, \dots, a_l)$ , we have

$$J(4, a_1, a_2, \dots, a_l) \leq O\left(\frac{n}{\sqrt{2}}\right). \tag{10}$$

*Proof.* The Balaban index of a tree  $T_n^l(4, a_1, a_2, \dots, a_l)$  is as follows:

$$J(T_n^l(4, a_1, a_2, \dots, a_l)) = (n - 1) \cdot \left[ \sum_{i=1}^l \frac{a_i}{\sqrt{(4n - 2a_i - l - 6)(3n - 2a_i - l - 4)}} + \sum_{i=1}^l \frac{1}{\sqrt{(2n - l - 2)(3n - 2a_i - l - 4)}} \right]. \tag{11}$$

Suppose  $1 \leq a_1 \leq a_2$ , and let  $a_1 = a$  and  $c = n - a - a_2$ . Then,

$$J(T_n^l(4, a_1, a_2, \dots, a_l)) = (n-1) \cdot \left[ \begin{aligned} & \frac{a}{\sqrt{(4n-2a-l-6)(3n-2a-l-4)}} \\ & + \frac{n-a-c}{\sqrt{(2n+2a+2c-l-6)(n+2a+2c-l-4)}} \\ & + \sum_{i=3}^l \frac{a_i}{\sqrt{(4n-2a_i-l-6)(3n-2a_i-l-4)}} \\ & + \frac{1}{\sqrt{(2n-l-2)(3n-2a-l-4)}} + \frac{1}{\sqrt{(2n-l-2)(n+2a+2c-l-4)}} \\ & + \sum_{i=3}^l \frac{1}{\sqrt{(2n-l-2)(3n-2a_i-l-4)}} \end{aligned} \right]. \quad (12)$$

And we get the derivative of  $J(T_n^l(4, a, a_2, \dots, a_l))$  related to  $a$  as follows:

$$\frac{\partial J(T_n^l(4, a, a_2, \dots, a_l))}{\partial a} = (n-1) \cdot \left[ \begin{aligned} & \frac{-2-l+2n}{((-2-l+2n)(-4-2a-l+3n))^{3/2}} \\ & + \frac{1}{((-4-2a-l+3n)(-6-2a-l+4n))^{1/2}} \\ & - \frac{-2-l+2n}{((-2-l+2n)(-4-l+n+2a+2c))^{3/2}} \\ & - \frac{1}{((-4-l+n+2a+2c)(-6-l+2n+2a+2c))^{1/2}} \\ & + \frac{a(-20-8a-4l+14n)}{2((-4-2a-l+3n)(-6-2a-l+4n))^{3/2}} \\ & - \frac{(-a-c+n)(-20-4l+6n+8a+8c)}{2((-4-l+n+2a+2c)(-6-l+2n+2a+2c))^{3/2}} \end{aligned} \right]. \quad (13)$$

The sum of positive terms above is denoted by  $S^+(n, a, l)$ , i.e.,

$$S^+(n, a, l) = (n-1) \cdot \left[ \begin{aligned} & \frac{-2-l+2n}{((-2-l+2n)(-4-2a-l+3n))^{3/2}} \\ & + \frac{1}{((-4-2a-l+3n)(-6-2a-l+4n))^{1/2}} \\ & + \frac{a(-20-8a-4l+14n)}{2((-4-2a-l+3n)(-6-2a-l+4n))^{3/2}} \end{aligned} \right]. \quad (14)$$

And the sum of absolute values of negative terms above is denoted by  $S^-(n, a, l)$ , i.e.,

$$S^-(n, a, l) = (n - 1) \cdot \left[ \begin{aligned} & \frac{-2 - l + 2n}{((-2 - l + 2n)(-4 - l + n + 2a + 2c))^{3/2}} \\ & + \frac{1}{((-4 - l + n + 2a + 2c)(-6 - l + 2n + 2a + 2c))^{1/2}} \\ & + \frac{(-a - c + n)(-20 - 4l + 6n + 8a + 8c)}{2((-4 - l + n + 2a + 2c)(-6 - l + 2n + 2a + 2c))^{3/2}} \end{aligned} \right]. \tag{15}$$

Then, we use  $S^+(n, a, l)$  to divide  $S^-(n, a, l)$ , and we get

$$\lim_{n \rightarrow \infty} \frac{S^-(n, a, l)}{S^+(n, a, l)} = \frac{5}{2^{3/2}} > 1. \tag{16}$$

So, the derivative of  $J(T_n^l(4, a_1, a_2, \dots, a_l))$  related to  $a_1$  is less than 0 as  $n$  is big enough. It means that the corresponding Balaban index increases along with the number of

pendent edges  $a_1$  decreasing and the number of pendent edges  $a_2$  increasing, i.e.,

$$J(T_n^l(4, a_1, a_2, \dots, a_l)) \leq J(T_n^l(4, a_1 - 1, a_2 + 1, \dots, a_l)). \tag{17}$$

Analogously, we obtain

$$\begin{aligned} J(T_n^l(4, a_1, a_2, \dots, a_l)) &\leq J(T_n^l(4, 1, a_1 + a_2 + a_3 + \dots + a_l - 1, 0, \dots, 0)) \\ &= J(T_n^l(4, 1, n - l - 2, 0, \dots, 0)). \end{aligned} \tag{18}$$

For the above tree  $T_n^l(4, 1, n - l - 2, 0, \dots, 0)$ , it can be seen that there are only two numbers, i.e., 1 and  $n - l - 2$ , as the numbers of pendent edges, respectively, attach to two pendent vertices of a star  $S_{l+1}$ . Thus, it is easy to check that

$$J(T_n^l(4, 1, n - l - 2, 0, \dots, 0)) = O\left(\frac{n}{\sqrt{2}}\right). \tag{19}$$

Hence,

$$J(T_n^l(4, a_1, a_2, \dots, a_l)) \leq O\left(\frac{n}{\sqrt{2}}\right). \tag{20}$$

On the contrary, the asymptotically tight lower bound of such a tree is not easy to be given due to the determination of parameter  $l$ , but we find that, in  $\mathcal{T}_n^l(4, a_1, a_2, \dots, a_l)$ , the tree attained the asymptotically tight lower bound which possesses a property satisfying  $|a_i - a_j| \leq 1$  for  $1 \leq i \neq j \leq l$ . In case of  $l = 2$ , we obtain the following result.  $\square$

**Theorem 3.** For any tree  $T_n^2(4, a_1, a_2) \in \mathcal{T}_n^2(4, a_1, a_2)$ , where  $1 \leq a_1 \leq a_2$ , we have

$$J(T_n^2(4, a_1, a_2)) \geq O\left(\frac{n}{\sqrt{6}}\right). \tag{21}$$

*Proof.* From the proof of Theorem 2, we see that the derivative of  $J(T_n^2(4, a_1, a_2))$  related to  $a_1$  is less than 0 as  $n$  is big enough. Thus, if  $a_1 = \lfloor n - 3/2 \rfloor, a_2 = \lceil n - 3/2 \rceil$ , then

$$J(T_n^2(4, a_1, a_2)) \geq J\left(T_n^2\left(4, \lfloor \frac{n-3}{2} \rfloor, \lceil \frac{n-3}{2} \rceil\right)\right). \tag{22}$$

Note that

$$J\left(T_n^2\left(4, \lfloor \frac{n-3}{2} \rfloor, \lceil \frac{n-3}{2} \rceil\right)\right) = O\left(\frac{n}{\sqrt{6}}\right). \tag{23}$$

Thus,

$$J(T_n^2(4, a_1, a_2)) \geq O\left(\frac{n}{\sqrt{6}}\right). \tag{24}$$

$\square$

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research was supported by the NSFQH (no. 2018-ZJ-925Q) and NSFC (no. 11701311).

### References

- [1] A. T. Balaban, "Highly discriminating distance-based topological index," *Chemical Physics Letters*, vol. 89, no. 5, pp. 399-404, 1982.
- [2] J. Devillers and A. T. Balaban, *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, The Netherlands, 1999.
- [3] R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley, Weinheim, Germany, 2000.
- [4] A. T. Balaban, P. V. Khadikar, C. T. Supuran, A. Thakur, and M. Thakur, "Study on supramolecular complexing ability vis-à-vis estimation of pKa of substituted sulfonamides:

- dominating role of Balaban index ( $j$ ),” *Bioorganic & Medicinal Chemistry Letters*, vol. 15, no. 17, pp. 3966–3973, 2005.
- [5] G. Grassy, B. Calas, A. Yasri et al., “Computer-assisted rational design of immunosuppressive compounds,” *Nature Biotechnology*, vol. 16, no. 8, pp. 748–752, 1998.
- [6] P. V. Khadikar, C. T. Supuran, A. Thakur, and M. Thakur, “QSAR study on benzene-sulphonamide carbonic anhydrase inhibitors: topological approach using balaban index,” *Bioorganic & Medicinal Chemistry*, vol. 12, pp. 789–793, 2004.
- [7] A. T. Balaban, “A comparison between various topological indices, particularly between the index  $J$  and Wiener index  $W$ ,” in *Topology in Chemistry Discrete Mathematics of Molecules*, D. H. Rouvray and R. B. King, Eds., Horwood, Chichester, England, 2002.
- [8] H. Dong and X. Guo, “Character of graphs with extremal balaban index,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 63, pp. 799–812, 2010.
- [9] L. Sun, “Bounds on the balaban index of trees,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 63, pp. 813–818, 2010.
- [10] B. Zhou and N. Trinajstić, “Bounds on the balaban index,” *Croatica Chemica Acta*, vol. 81, pp. 319–323, 2008.
- [11] A. T. Balaban, N. Ionescu-Pallas, and T. S. Balaban, “Asymptotic values of topological indices  $j$  and  $j'$  (average distance sum connectivities) for infinite cyclic and acyclic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 17, pp. 121–146, 1985.
- [12] M. Knor, J. Kranjc, R. Škrekovski, and A. Tepeh, “A search for the minimum value of balaban index,” *Applied Mathematics and Computation*, vol. 286, pp. 301–310, 2016.
- [13] M. Knor, J. Kranjc, R. Škrekovski, and A. Tepeh, “A note on accumulation points of balaban index,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 78, pp. 163–168, 2017.
- [14] M. Knor, J. Kranjc, R. Škrekovski, and A. Tepeh, “On the minimum value of sum-balaban index,” *Applied Mathematics and Computation*, vol. 303, pp. 203–210, 2017.

## Retraction

# Retracted: Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries

### Mathematical Problems in Engineering

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Z.-J. Zhou, Y. Wang, M.-M. Lu, and J.-M. Zhu, "Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8723062, 10 pages, 2020.



## Research Article

# Government Intervention, Financial Support, and Comprehensive Efficiency of Enterprise Independent Innovation: Empirical Analysis Based on the Data of Chinese Strategic Emerging Industries

Ze-Jiong Zhou <sup>1</sup>, Yao Wang <sup>2</sup>, Miao-Miao Lu <sup>1</sup>, and Jia-Ming Zhu <sup>3</sup>

<sup>1</sup>School of Economics, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

<sup>2</sup>School of Finance, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

<sup>3</sup>School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

Correspondence should be addressed to Jia-Ming Zhu; zhujm1973@163.com

Received 22 August 2020; Revised 3 October 2020; Accepted 19 October 2020; Published 3 November 2020

Academic Editor: Shaohui Wang

Copyright © 2020 Ze-Jiong Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Government intervention and financial support are two major means to promote the independent innovation performance of enterprises in strategic emerging industries, and government intervention has induced crowding-out effects on financial support, which leads to the uncertainty of the dual incentive effect of government intervention and financial support on enterprises' independent innovation. The research object of this paper is 657 strategic emerging enterprises listed in Shanghai and Shenzhen. We empirically studied the impact of government intervention and financial support on the comprehensive efficiency of independent innovation of strategic emerging enterprises. The empirical study draws the following conclusions. Firstly, the comprehensive efficiency of independent innovation of enterprises is in the trend of continuous improvement and technical efficiency and scale efficiency are also increasing, but the technical efficiency is lower than the scale efficiency, which shows that the improvement of independent innovation efficiency mainly depends on the expansion of innovation scale. Secondly, both government intervention and financial support promote the comprehensive efficiency of independent innovation of strategic emerging industry enterprises, but the incentive effect of government intervention is more obvious. Thirdly, there is an inverted U-shaped relationship between government intervention and the comprehensive efficiency of independent innovation. Fourthly, the regression coefficient of the interaction between government intervention and financial support and the comprehensive efficiency of enterprise independent innovation is negative, which indicates that government intervention has an inhibitory effect on the effect of financial support on the overall efficiency of enterprise independent innovation. Finally, we put forward countermeasures and suggestions.

## 1. Introduction

In recent years, the world economy has seriously slumped, international trade and investment have shrunk, and China's domestic consumption, investment, and exports have declined significantly. At present and in the future, China's economic development is facing unprecedented risks and challenges, and it is urgent to find a new driving force for economic growth. Strategic emerging industries are the key industries for China to cultivate new driving

forces for economic development, realize economic transformation, and gain new advantages in global competition. With government intervention and financial support, China's strategic emerging industries continue to invest in independent innovation, but the overall efficiency of enterprise independent innovation is not high. The comprehensive efficiency of enterprises' independent innovation represents the input-output ratio of enterprises in the process of innovation. It is a comprehensive summary of enterprises' independent

innovation input and output and is a key indicator to measure the sustainable development of enterprises. It reflects the ability of enterprises' independent innovation resource allocation, technology transformation, and new product output [1, 2]. Government intervention and financial support have dual incentive effects on the improvement of the comprehensive efficiency of enterprise independent innovation. The government uses financial subsidies and tax incentives to encourage enterprises to increase investment in independent innovation and accelerate the transformation of independent innovation achievements and the production of new products [3]. Financial institutions guide private and social capital into enterprises' independent innovation activities through bank credit and other financial instruments and allocate funds to strategic emerging industry enterprises with investment value, so as to ease the financial constraints faced by enterprises [4]. When the government invests too much money into the independent innovation activities of enterprises, the financing behavior of enterprises will deviate from the principle of market efficiency, and the incentive effect of financial support on enterprises' independent innovation will be weakened [5, 6].

Scholars have done a lot of research on the influence of government intervention on enterprises' independent innovation. The existing researches mainly focus on the ideal role of government intervention in the process of enterprise independent innovation. Scholars usually divide the role of government intervention in the process of enterprise independent innovation into three situations: the hand of aid, the hand of plunder, and the hand of inaction. One view is that the government provides financial funds for enterprises' independent innovation activities, which plays a "helping hand" to ensure the smooth development of enterprise innovation activities and improve the efficiency of independent innovation [7–13]. Another view is that government intervention can easily lead to government rent-seeking behavior in enterprise independent innovation activities, which plays a role of "predatory hand"; that is to say, government intervention increases the cost of enterprise independent innovation and hinders the improvement of comprehensive efficiency of enterprise independent innovation [14–22]. In addition, some scholars believe that the initial purpose of government intervention is to encourage enterprises to carry out independent innovation, but in special industries and under specific conditions, government intervention will play the role of "hands of inaction"; that is to say, government intervention has no effect [23–27].

In the study of the relationship between government intervention and enterprise independent innovation, we need to consider the financial support. In the presence of government intervention, financial support can not only ease the financing constraints of enterprises' independent innovation but also lead to insufficient investment in innovation. One view is that the government guides financial institutions to provide financial support for enterprises' independent innovation, so as to solve the

problem of financial constraints of enterprises' independent innovation [28–32]. Another view is that government intervention distorts the allocation function of financial resources, resulting in insufficient investment in independent innovation and difficulty in improving the efficiency of independent innovation [33–35].

There are two deficiencies in the existing literature: firstly, when studying the effect of government intervention and financial support on enterprise independent innovation, most scholars are interested in the impact of government intervention and financial support on the input and output of independent innovation, and few pieces of literature study the influence of government intervention and financial support on the final output (i.e., comprehensive efficiency) of enterprise independent innovation. The independent innovation of enterprises includes three stages: innovation input, innovation output, and innovation final output. If we ignore the research on the final output of independent innovation (i.e., comprehensive efficiency), we cannot completely open the "black box of independent innovation." Secondly, when studying the effect of government intervention and financial support, the existing literature ignores the induced effect and crowding-out effect of government intervention on financial support, so the dual incentive effect of government intervention and financial support on enterprises' independent innovation cannot be correctly evaluated. Thirdly, in the analysis of the impact of government intervention on enterprise independent innovation, the indicators selected in the existing literature are relatively single, which cannot correctly measure the effect of government intervention.

Based on the data of 657 strategic emerging enterprises listed in Shenzhen and Shanghai from 2012 to 2018, we empirically analyze the impact of government intervention and financial support on the comprehensive efficiency of independent innovation, so as to open the black box of independent innovation of strategic emerging industry enterprises. When selecting the sample companies, the following principles should be followed: select the strategic emerging industry listed companies in the wind database; the initial starting time of the sample is set to 2012; the selected listed companies are in the growth period; exclude the listed companies with ST and \*ST types; exclude the listed companies with tax rate less than 0 or greater than 1.

The innovation of this paper is as follows. Firstly, we divide the process of enterprise independent innovation into three stages, innovation input, innovation output, and innovation final output, and conduct empirical research on the impact of government intervention and financial support on the final output of enterprise independent innovation, so as to open the last stage of enterprise independent innovation black box. Secondly, it analyzes the dual incentive effect of government intervention and financial support on enterprise independent innovation and the influence of government intervention on financial support effect. Thirdly, we establish the index system of government intervention to make the measurement of the effect of government intervention more accurate.

## 2. Measurement of the Efficiency of Government Intervention, Financial Support, and Independent Innovation

*2.1. Measurement of Government Intervention.* By means of financial subsidies and tax incentives, the government intervenes in the comprehensive efficiency of enterprise independent innovation. According to the latest accounting standards, this paper selects the government subsidy index under other income accounting subjects in enterprise financial statements as the measurement index. Tax incentives include the deduction of R&D expenses, value-added tax, and income tax. The selection methods of these three tax preference indicators are as follows. Referring to the research of Han Renyue, this paper takes the actual amount of enterprise R&D expenses in 2012–2016 and the actual amount of enterprise R&D expenses in 2017 and 2018 multiplied by 50% and 75%, respectively, as the indicators to measure the additional deduction of R&D expenses. Referring to the practice of Wang Chunyuan, this paper uses the VAT deduction of new fixed assets of enterprises in the year as the measurement index of VAT preference. Referring to the practice of Li Xiangju, we take corporate income tax as the reverse indicator of tax preference.

Tax incentives include an additional deduction of R&D expenses, preferential treatment of value-added tax, and preferential treatment of income tax. When we study the input and output stages of enterprise independent innovation, we need to use a number of indicators to measure the effect of government intervention, and in the stage of comprehensive efficiency of enterprise independent innovation, we pay more attention to the comprehensiveness of multiple indicators. Referring to the existing literature, this paper uses the panel factor analysis method and uses the final comprehensive index as the measurement index of government intervention [36, 37].

*2.1.1. Measurement Model of Government Intervention.* In order to standardize the government intervention indicators and eliminate the differences in the number and unit of each variable, we establish the following factor analysis model:

$$\begin{aligned} X &= (x_1, x_2, x_3, \dots, x_p), \\ F &= (F_1, F_2, F_3, \dots, F_m), \\ e &= (e_1, e_2, e_3, \dots, e_p), \end{aligned} \quad (1)$$

where  $X$  is the standardized government intervention matrix,  $x_p$  is the components of the government intervention matrix,  $F$  is the common factor of  $X$ ,  $F_m$  is the components of the common factor matrix,  $e$  is the special factor of  $X$ , and  $e_p$  is the components of the special factor  $e$ .

The expressions of formula (1) are as follows:

$$\begin{cases} x_1 = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + \dots + a_{1m}F_m + e_1, \\ x_2 = a_{21}F_1 + a_{22}F_2 + a_{23}F_3 + \dots + a_{2m}F_m + e_2, \\ x_3 = a_{31}F_1 + a_{32}F_2 + a_{33}F_3 + \dots + a_{3m}F_m + e_3, \\ \dots \\ x_p = a_{p1}F_1 + a_{p2}F_2 + a_{p3}F_3 + \dots + a_{pm}F_m + e_p. \end{cases} \quad (2)$$

The above formula can be simplified into the matrix form:

$$\begin{aligned} X &= AF + e, \\ A &= (a_{ij}). \end{aligned} \quad (3)$$

According to the requirements of the factor analysis model,  $\text{Cov}(F, e) = 0$ ,  $\text{Cov}(F) = 1$ , and the covariance of  $e$  is a diagonal matrix.  $X$  is the standardized government intervention matrix,  $x_p$  is the component of the government intervention matrix,  $F$  is the common factor of  $X$ ,  $F_m$  is the component of the common factor matrix,  $a_{ij}$  is the factor load matrix,  $A$  is the factor load matrix, and  $e$  is the special factor of  $X$ .

### 2.1.2. Panel Factor Analysis of Government Intervention Indicators

(1) *Test of Factor Analysis.* Whether the sample data is suitable for factor analysis needs to be tested for applicability. There are two commonly used fitness test methods, namely, KMO measure and Bartlett's spherical test. The applicability of factor analysis on government intervention indicators was tested. The test results are as follows. From 2012 to 2018, the KMO statistical values of government intervention are in the range of 0.6–0.8, and the probability values corresponding to LR test statistics are all less than 0.05. It can be seen that many indicators of government intervention are suitable for factor analysis. The test results are listed in Table 1.

(2) *Analysis of Panel Factor.* This paper makes factor analysis on the cross-sectional data of government intervention from 2012 to 2018 and obtains the final government intervention panel data. The normalized data in 2012 were rotated to obtain the eigenvalue and contribution rate. The eigenvalues of factor 1 and factor 2 are greater than 1, which are 1.3424 and 1.2332, respectively. The cumulative variance contribution rate of factor 1 and factor 2 after the maximum orthogonal rotation is 90.63%, which indicates that factor 1 and factor 2 can be extracted as public factors of subsidies and taxes. See Table 2 for details. Referring to the factor analysis process of 2012 government intervention indicators, the four indicators of government intervention from 2013 to 2018 were reduced in order (the calculation results are not listed). The results of factor 1 and factor 2 are named as subsidy factor G1 and tax factor G2, and the comprehensive

TABLE 1: Applicability test of government intervention factor analysis from 2012 to 2018.

Particular year	2012	2013	2014	2015	2016	2017	2018
Statistical value of KMO	0.6886	0.7637	0.7248	0.7087	0.7339	0.7096	0.7078
LR test: $\chi^2(6)$ : Prob > $\chi^2$	43.59	44.75	47.97	87.91	126.61	135.98	127.52
	0.0018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 2: Eigenvalues and variance contribution rate of raw data and maximum orthogonal rotation data of government intervention in 2012.

Factors	Original data			After maximum orthogonal rotation		
	Eigenvalues	Variance contribution	Cumulative contribution rate of variance	Eigenvalues	Variance contribution	Cumulative contribution rate of variance
Factor 1	1.47861	0.3697	0.3697	1.3424	0.4570	0.4570
Factor 2	1.07837	0.2696	0.6393	1.2332	0.4493	0.9063
Factor 3	0.6983	0.1749	0.8142	—	—	—
Factor 4	0.5824	0.1398	0.954	—	—	—

score of government subsidy and tax factor is taken as the measurement index  $g$  of government intervention. See Table 3 for details.

(3) *Comprehensive Score of Government Intervention and Its Descriptive Statistical Analysis.* Using the method of panel factor analysis, this paper calculates the score of government subsidy factor and tax factor and then obtains the comprehensive score of government intervention. The results of descriptive statistical analysis are shown in Table 4. It can be seen from Table 4 that the average value of subsidy factor is greater than that of tax factor, indicating that the promotion effect of government subsidy on independent innovation effect of strategic emerging industry enterprises is greater than that of tax preference. The standard deviation of subsidy factor is larger than that of tax factor, indicating that the data stability of subsidy factor is inferior to that of tax factor. The minimum and maximum values of  $G_1$ ,  $G_2$ , and  $G_3$  are close to each other, indicating that the data of the three are relatively stable as a whole.

*2.2. Measurement of Financial Support.* Financial support means include bank loans and foreign direct investment, and bank loans include short-term loans and long-term loans.

Referring to the existing literature, this paper selects short-term bank loans as one of the measurement indicators of financial support, long-term bank loans as one of the indicators to measure financial support, and the overseas borrowed funds in the total project investment approved by the relevant government departments as the index to measure the foreign direct investment. In the same way as the government intervention indicators, we will reduce the dimension of financial support indicators to get the comprehensive score of financial support.

*2.2.1. Panel Factor Analysis of Financial Support.* The applicability of factor analysis is tested for financial support indicators, and the test results are shown in Table 5. It can be seen from Table 5 that the KMO values of financial support

indicators are greater than 0.7 from 2012 to 2018, and the probability corresponding to LR test statistical values is less than 0.05, indicating that the financial support indicators of each year can be factor analyzed.

The eigenvalues of factor 1 and factor 2 are greater than 1, and the eigenvalues of factor 1 and factor 2 are 1.69389 and 1.30467, respectively. The cumulative variance contribution rate of factor 1 and factor 2 is 99.95%. Therefore, factors 1 and 2 are extracted as common factors of domestic investment  $F_1$  and foreign investment  $F_2$ . See Tables 6 and 7 for details.

*2.2.2. Comprehensive Score of Financial Support and Its Descriptive Statistical Analysis.* Using the panel factor analysis method, we calculate the comprehensive score of financial support. The results of descriptive statistical analysis are shown in Table 8. It can be seen from Table 8 that the average and maximum values of domestic investment factor and foreign investment factor are close, and the minimum value of comprehensive factor is far less than that of domestic investment factor and foreign investment factor. The standard error of domestic investment factors is much larger than that of foreign investment factors and the labeling difference of comprehensive factors, which indicates that the data stability of domestic investment factors is not high.

### 2.3. Measurement of the Comprehensive Efficiency of Enterprise Independent Innovation

*2.3.1. Input and Output Variables of Enterprise Independent Innovation.* Referring to the existing literature, we take capital input in R&D and personnel input in R&D as the independent innovation input variables of strategic emerging industry enterprises and the number of patent applications and main business income as the output variables of independent innovation [38]. The input and output variables of enterprise independent innovation are shown in Table 9.



TABLE 3: Naming of government intervention factors.

Sign of factor	High load index	Factor naming
$G_1$	Government grants	Subsidy factor
	Additional deduction of R&D expenses	
	Preferential treatment of value-added	
$G_2$	tax	Tax factor
	Income tax	

TABLE 4: Descriptive statistics of comprehensive scores of government intervention.

Variable	Mean	Std. dev.	Min	Max
G1	0.0949	1.0536	-4.3724	4.9740
G2	0.0445	0.8175	-3.6354	4.6622
G3	0.0085	0.5489	-2.8774	4.3763

2.3.2. *Evaluation Model of Enterprise Independent Innovation Efficiency.* DEA-CCR model and DEA-BBC model are commonly used to evaluate efficiency. BBC model focuses on pure technical efficiency and scale efficiency, while this paper focuses on measuring the comprehensive efficiency of independent innovation, so DEA-CCR model is selected to measure the comprehensive efficiency of enterprise independent innovation [39, 40].

Suppose the number of DMUs is  $m$ . For any DMU<sub>0</sub>, it is assumed that there are  $n$  types of inputs and  $s$  types of outputs,  $x_{ij}$  represents the type  $i$  input of the  $j$ th DMU, and  $y_{rj}$  represents the type  $r$  output of the  $j$ th DMU. DEA-CCR model can be expressed as follows:

$$\begin{aligned} \min & \left[ \theta - \varepsilon \left( \sum_{i=1}^n s_i^- + \sum_{r=1}^s s_r^+ \right) \right], \\ \text{s.t.} & \sum_{j(\neq k)=1}^m x_{ij} \lambda_j + s_i^- \leq \theta X_{i0}, \\ & \sum_{j(\neq k)=1}^m y_{rj} \lambda_j + s_r^+ \leq Y_{r0}, \quad \lambda \geq 0, s_r^+ \geq 0, s_i^- \geq 0. \end{aligned} \tag{4}$$

where  $i = 1, 2, \dots, n; j = 1, 2, \dots, m; r = 1, 2, \dots, s; m$  represents the number of decision-making units,  $n$  represents the number of input variables,  $s$  represents the number of output variables.  $s_i^+$  represents the relaxation variable of the  $r$ th output,  $s_i^-$  represents the relaxation variable of the  $i$ th input, and  $\theta$  represents the comprehensive efficiency.

2.3.3. *Cross-Sectional Data and Descriptive Statistical Analysis of the Comprehensive Efficiency of Independent Innovation.* Using DEA-CCR model, we get the cross-sectional data of the comprehensive efficiency of independent innovation of strategic emerging industry enterprises. Descriptive statistical analysis of the cross-sectional data of each year is carried out, and the results are shown in Table 10.

As can be seen from the regression analysis results in Table 10, the overall efficiency of enterprise independent innovation was generally in the trend of continuous improvement; its value increased from 0.264 in 2012 to 0.746 in 2018. In the process of improving the comprehensive efficiency of enterprise independent innovation, its technical efficiency and scale efficiency are increasing, but the former is lower than the latter.

2.3.4. *Panel Data of Comprehensive Efficiency of Independent Innovation and Its Descriptive Statistical Analysis.* Based on the annual cross-sectional data of the comprehensive efficiency of enterprise independent innovation, we get the panel data of the comprehensive efficiency of enterprise independent innovation. We make a descriptive statistical analysis on the panel data of the comprehensive efficiency of enterprise independent innovation. The results are shown in Table 11.

It can be seen from Table 11 that the average value of scale efficiency is greater than that of pure technical efficiency, indicating that the expansion of innovation scale has a more obvious effect on promoting the efficiency of independent innovation of enterprises. The standard errors of comprehensive efficiency, pure technical efficiency, and scale efficiency are very close, which indicates that the efficiency of independent innovation is relatively stable. The maximum values of comprehensive efficiency, pure technical efficiency, and scale efficiency all reach 1, but their minimum values are obviously different.

### 3. Empirical Research on the Influence of Government Intervention and Financial Support on the Comprehensive Efficiency of Independent Innovation

3.1. *Index Selection.* The explained variable is the comprehensive efficiency of independent innovation (crste), and its value is the comprehensive efficiency value calculated by DEA-CCR. The explanatory variables are government intervention ( $G$ ) and regulatory variable financial support ( $F$ ). The values of these two variables have been obtained by factor analysis.

The controlling variables are enterprise scale (size), the profitability of enterprises (ep), the age of enterprises (age), the capital structure level (lev), and the types of enterprise ownership (own). The enterprise scale is the logarithm of the total assets of the enterprise, the profitability is the proportion of net profit and operating income, the age of an enterprise is the number of years from the time the company was listed to the research deadline, and the capital structure level is the asset-liability ratio. The values of the type of enterprise ownership are as follows: the value of state-owned enterprise holding is 0, and the value of others is 1.

3.2. *Tobit Model.* The range of the explained variable (comprehensive efficiency of independent innovation) in this paper is (0, 1], while some explanatory variables are less

TABLE 5: Applicability of financial support factor analysis in 2012–2018.

Particular year	2012	2013	2014	2015	2016	2017	2018
Statistical value of KMO	0.711	0.7503	0.7253	0.7390	0.7459	0.7384	0.7402
LR test: chi2(6): Prob > chi2	1186.93	1719.61	2298.65	2158.06	3568.46	3194.84	3063.46
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 6: Eigenvalues and variance contribution rate of raw data and maximum orthogonal rotation data of financial support in 2012.

Factors	Original data			After maximum orthogonal rotation		
	Eigenvalues	Variance contribution	Cumulative contribution rate of variance	Eigenvalues	Variance contribution	Cumulative contribution rate of variance
Factor 1	1.72184	0.5739	0.5739	1.69389	0.5646	0.5646
Factor 2	1.27672	0.4256	0.9995	1.30467	0.4349	0.9995
Factor 3	0.00144	0.0005	1.0000	—	—	—

TABLE 7: Naming of financial support factors.

Sign of factor	High load index	Factor naming
$F_1$	Short-term loans from banks	Domestic investment factor
	Long-term loans from banks	
$F_2$	Foreign direct investment	Foreign investment factors

TABLE 8: Descriptive statistics of comprehensive scores of financial support.

Variable	Mean	Std. dev.	Min	Max
$F_1$	0.1216	1.0020	-4.4420	4.8377
$F_2$	0.1952	0.5644	-2.5980	5.5670
$F_3$	0.1572	0.2850	-5.0457	5.5145

than zero. If the panel OLS or panel effect model is used for empirical analysis, the regression result will be biased. Therefore, using the methods of existing literature, this paper uses Tobit model to study the comprehensive

efficiency of enterprise independent innovation [41, 42]. The specific expression of Tobit model is as follows:

$$\begin{cases} Y_i^* = \beta_0 + \beta^T X_i + \mu_i, \\ Y_i = Y_i^*, \text{ if } Y_i^* \geq 0, \\ Y_i = 0, \text{ if } Y_i^* < 0, \end{cases} \quad (i = 1, 2, \dots, n), \quad (5)$$

where  $Y_i$  is the explanatory variable,  $X_i$  is the explanatory variable,  $Y_i^*$  is the potential variable,  $\beta^T$  is the parameter vector, and  $\mu_i$  is a random error term. In this paper,  $Y_i$  is the comprehensive efficiency value of independent innovation,  $X_i$  is government intervention, and  $Y_i^*$  is financial support.

When  $Y_i^* < 0$ , the probability density function is as follows:

$$\begin{aligned} P(Y = 0) &= P(Y_i^* \leq 0) = \varphi\left(\frac{\beta^T x_i}{\sigma}\right) \\ &= 1 - \varphi\left(\frac{\beta^T x_i}{\sigma}\right). \end{aligned} \quad (6)$$

Tobit model follows the concept of maximum likelihood method. When  $Y_i = Y_i^*$ , its likelihood function is expressed as follows:

$$I(\beta^T) = \sum_{Y_i > 0} \ln \left[ \frac{1}{\sigma} \varphi\left(\frac{Y_i - \beta^T x_i}{\sigma}\right) \right] + \sum_{Y_i = 0} \ln \left[ 1 - \varphi\left(\frac{Y_i - \beta^T x_i}{\sigma}\right) \right]. \quad (7)$$

The derivative method is used to maximize  $\beta$  and  $\sigma$ , so as to obtain the maximum likelihood value.

### 3.3. Results and Analysis of Empirical Research

3.3.1. Analysis of the Lag Effect of the Comprehensive Efficiency of Independent Innovation. Since the output of independent innovation has a lag effect, this paper adds the square term of the efficiency of independent innovation in the Tobit model to test the lag effect and “inverted U”

relationship of the comprehensive efficiency of independent innovation.

The regression result of the lag effect of the comprehensive efficiency of independent innovation is shown in model (a) in Table 12. According to the regression result, the regression coefficient between the efficiency of independent innovation in the lag period and that in the current period is positive; it shows that the efficiency of independent innovation in the lag period can improve the comprehensive efficiency of independent innovation in the current period.

TABLE 9: Input and output variables of enterprise independent innovation.

Primary variables	Secondary variables	Sign of variables	Definition and calculation of variables
Input variables	Capital input in R&D	lnR&D	Take logarithm for R&D capital investment of enterprises
	Personnel input in R&D	lnL	Take logarithm for the number of R&D personnel
Output variables	Number of patent applications	lnPAT	Take logarithm for the number of patent applications
	Income from main business	lnMBI	Take logarithm for the income from main business

TABLE 10: Descriptive statistics of comprehensive efficiency of enterprise independent innovation from 2012 to 2018.

Years	Variables	Mean	Sd	Min	Max
2012	Crste	0.264	0.0622	0.0379	1
	Vrste	0.288	0.0636	0.0648	1
	Scale	0.37	0.0347	0.0429	1
2013	Crste	0.372	0.0696	0.0627	1
	Vrste	0.289	0.0663	0.0612	1
	Scale	0.581	0.0212	0.1855	1
2014	Crste	0.402	0.049	0.0802	1
	Vrste	0.552	0.0382	0.255	1
	Scale	0.548	0.035	0.114	1
2015	Crste	0.42	0.048	0.2	1
	Vrste	0.549	0.0394	0.26	1
	Scale	0.619	0.0341	0.2826	1
2016	Crste	0.662	0.0533	0.17	1
	Vrste	0.581	0.0596	0.271	1
	Scale	0.679	0.0346	0.1803	1
2017	Crste	0.649	0.041	0.281	1
	Vrste	0.601	0.0402	0.0802	1
	Scale	0.687	0.0183	0.282	1
2018	Crste	0.746	0.0379	0.3724	1
	Vrste	0.866	0.0428	0.3762	1
	Scale	0.61	243.2	0.361	769

Crste represents the comprehensive efficiency of independent innovation, Vrste represents pure technical efficiency, and Scale represents scale efficiency.

TABLE 11: Descriptive statistics of panel data of independent innovation comprehensive efficiency of strategic emerging enterprises.

Variables	Mean	Sd	Min	Max
Crste	0.562	0.0689	0.0379	1
Vrste	0.589	0.0727	0.0648	1
Scale	0.635	0.0776	0.0429	1

But there is no “inverted U” relationship between the efficiency of independent innovation in the first stage of lag and the efficiency of independent innovation in the current period.

3.3.2. *Analysis of the Effect of Government Intervention and Financial Support.* Using Tobit model, this paper analyzes the effect of government intervention and financial support on the comprehensive efficiency of enterprise independent innovation, and the regression analysis result is shown in model (a) in Table 12. In order to test the nonlinear relationship between government intervention and the

TABLE 12: Tobit panel regression results.

Variables	Model (1)	Model (2)	Model (3)
$vrste_{it-1}$	0.0636** (0.72)	0.0717** (0.89)	0.0909** (1.81)
$Vrste2_{it-1}$	-0.0013 (-0.82)	-0.0026 (-0.98)	-0.0019 (-0.91)
$G$	0.015** (2.24)	0.0151** (2.25)	0.0143** (2.12)
$F$	0.0047* (1.54)	0.0047* (1.55)	0.0038* (1.20)
$G \times F$	—	-0.0062* (-1.42)	-0.0057* (-1.23)
$G^2$	—	—	-0.0097* (-0.98)
$F^2$	—	—	-0.0009 (-1.54)
$Lnsize$	-0.0005 (-0.22)	-0.0007 (-0.32)	-0.0004 (-0.17)
$ep$	0.0004 (0.3)	0.00038 (0.25)	0.00038 (0.27)
Age	0.0001 (0.11)	0.0001 (0.12)	0.00021 (0.2)
Lev	-0.0008 (-0.77)	-0.0008 (-0.79)	-0.0007 (-0.75)
Own	-0.0015 (-0.33)	-0.0018 (-0.39)	-0.0017 (-0.37)
Constant	0.892*** (22.10)	0.8925*** (22.26)	0.8931*** (22.29)
Wald test	14.34**	14.72**	11.01*
Log likelihood	-96.73	-92.51	-80.12

Wald test statistics pass 5% significance test, indicating that the model fits well.

comprehensive efficiency of enterprise independent innovation, as well as the nonlinear relationship between financial support and the comprehensive efficiency of enterprise independent innovation, the square terms of government intervention and financial support are added to the regression analysis model. The regression results are shown in model (3) in Table 12.

According to the regression results of model (1), the regression coefficient between government intervention and the comprehensive efficiency of independent innovation is positive and has passed the significance level of 5%, which indicates that government intervention promotes the improvement of comprehensive efficiency of independent innovation of enterprises. The regression coefficient is 0.0150, which indicates that, for every 1% increase of government intervention, the comprehensive efficiency of enterprise independent innovation will increase by 0.015%. The



regression coefficient between financial support and the comprehensive efficiency of independent innovation is positive and has passed the significant level of 10%, which indicates that financial support also promotes the improvement of comprehensive efficiency of independent innovation. The regression coefficient is 0.0047, which indicates that when the financial support increases by 1%, the comprehensive efficiency of enterprise independent innovation will increase by 0.015%. Comparing the two regression coefficient values, we can see that the promotion effect of government intervention on the comprehensive efficiency of independent innovation is far greater than that of financial support, which indicates that, in recent years, the government intervention in the allocation of innovation resources of enterprises is too strong, and the allocation efficiency of financial resources is relatively low.

According to the regression analysis results of model (3), the regression coefficient between the square term of government intervention and the comprehensive efficiency of enterprise independent innovation is negative and has passed the significance level of 10%, which indicates that there is an inverted U-shaped relationship between government intervention and the comprehensive efficiency of independent innovation. The regression coefficient between the square term of financial support and the comprehensive efficiency of independent innovation is negative, which fails to pass the significance level of 10%, which indicates that there is no "inverted U" relationship between financial support and comprehensive efficiency of independent innovation.

It can be seen that the degree of government intervention in the independent innovation of strategic emerging industry enterprises is in a reasonable range, which indicates that the reasonable increase of government intervention can improve the comprehensive efficiency of enterprise independent innovation. When the government intervention exceeds a certain range, the comprehensive efficiency of independent innovation will decline.

**3.3.3. The Inhibitory Effect of Government Intervention on Financial Support.** The interaction between government intervention and financial support is added to the regression analysis to study how government intervention affects the promotion of financial support on the overall efficiency of enterprise independent innovation. The regression analysis results are shown in model (2) in Table 12.

The regression analysis results show that the regression coefficient of the interaction between government intervention and financial support and the comprehensive efficiency of independent innovation is negative and has passed the significant level of 10%, which indicates that the government intervention has an inhibitory effect on the promotion effect of financial support on the comprehensive efficiency of enterprise independent innovation. The regression coefficient is 0.0062, which indicates that when the interaction items increase by 1%, the comprehensive efficiency of enterprise independent innovation will decrease by 0.0062%. Due to the "inverted U" relationship between

government intervention and the comprehensive efficiency of enterprise independent innovation and the inhibitory effect of government intervention on financial support effect, increasing government intervention may lead to the decline of comprehensive efficiency of enterprise independent innovation. Therefore, we must make good use of both government intervention and financial support, increase the induced effect of government intervention on financial support, and reduce the crowding-out effect of government intervention on financial support.

According to the regression analysis results, the coefficients of the control variables *lnsize*, *lev*, and *own* are all negative, but they fail to pass the 10% significance level, which indicates that the enterprise scale, capital structure level, and enterprise ownership type have a negative impact on the comprehensive efficiency. The coefficients of control variables *ep* and *age* are both positive, but they fail to pass the 10% significance level proposal, which shows that the profitability and age of enterprises have a positive impact on the overall efficiency, but this effect is not obvious.

#### 4. Conclusions and Suggestions

Strategic emerging industries are the key industries to cultivate new driving forces for economic development and achieve high-quality economic development. However, enterprises in strategic emerging industries have insufficient innovation power and low innovation performance in China. Government intervention and financial support are two external means to promote the independent innovation performance of strategic emerging industry enterprises, and the government intervention has induced effect and crowding-out effect on financial support, which leads to the dual incentive effect of government intervention and financial support on independent innovation to be tested.

This paper takes 657 strategic emerging enterprises as the research object and empirically studies the influence of government intervention and financial support on the comprehensive efficiency of independent innovation. The conclusions are as follows. Firstly, the comprehensive efficiency of independent innovation is in the trend of continuous improvement and technical efficiency and scale efficiency are also increasing, but the technical efficiency is lower than the scale efficiency, which shows that the improvement of independent innovation efficiency mainly depends on the expansion of innovation scale. Secondly, both government intervention and financial support promote the comprehensive efficiency of independent innovation of strategic emerging industry enterprises, but the incentive effect of government intervention is more obvious. Thirdly, there is an inverted U-shaped relationship between government intervention and the comprehensive efficiency, while there is no inverted U-shaped relationship between financial support and the comprehensive efficiency. Fourthly, the regression coefficient of the interaction between government intervention and financial support and the comprehensive efficiency of enterprise independent innovation is negative, which indicates that government

intervention has an inhibitory effect on the effect of financial support on the overall efficiency of enterprise independent innovation. The reason is that government intervention has crowding-out effect on financial support and distorts the allocation of financial resources. In addition, excessive government intervention will lead the external financing behavior of innovation subject to deviate from the principle of market efficiency and further weaken the role of financial support.

The following measures should be taken:

- (1) Pay attention to the capital investment and R&D personnel training, and promote the sustainability of independent innovation. First of all, improve the government fund guidance and tax preferential policies to promote enterprises to increase investment in independent innovation; secondly, improve the talent training system and policies to promote the effectiveness of enterprise researchers training.
- (2) Appropriately expand government intervention, and improve the accuracy of government subsidies and tax incentives. Firstly, in the output and efficiency stage of enterprise independent innovation, the government should increase the incentive for the transformation of independent innovation achievements. Secondly, formulate accurate financial subsidies and tax preferential system for different types of strategic emerging enterprises' independent innovation activities, so as to improve the effect of government intervention.
- (3) Coordinate the government intervention and financial support to improve the dual incentive effect. We should strengthen financial support for independent innovation activities and relax the control of financial resources, strengthen the guidance of financial institutions and encourage them to innovate financial instruments, and reduce the external financing constraint threshold of enterprises' independent innovation activities and reduce their financing costs.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This work was supported by the project of Anhui University Humanities and Social Sciences Research (Grant no. SK2020A0011), Key Project of National Social Science Foundation (Grant no. 17AJY018), and the Teaching and Research Fund Project of the Anhui University of Finance and Economics (acjyyb2020011).

### References

- [1] J. Li and Y. Mu, "Research on innovation performance of Chinese listed enterprises based on DEA method," *Science of Science and Management of Science and Technology*, vol. 36, no. 2, pp. 111–121, 2015.
- [2] C. Wu, L. Huang, and C. Wen, "Research on technological innovation efficiency and its influencing factors in the Yangtze River economic belt," *China Soft Science*, no. 5, pp. 160–170, 2017.
- [3] L. Yuan and X. Zheng, "Coupling effects of environmental regulation and government subsidies on enterprise technological innovation," *Resource Science*, vol. 39, no. 5, pp. 911–923, 2017.
- [4] T. Wang, L. Chen, and L. Meng, "A comparative study on the impact of government support path on technological innovation investment of wind power enterprises in China—system GMM analysis based on dynamic panel data," *Soft Science*, vol. 33, no. 2, pp. 12–16, 2019.
- [5] X. Liu, H. Song, and L. Fan, "Signaling effects of government subsidies and enterprise innovation on investors' investment decisions," *Science and Technology Progress and Countermeasures*, vol. 37, no. 2, pp. 26–33, 2020.
- [6] X. Yan and G. Huang, "Government financial subsidies, enterprise economy, innovation performance and overcapacity: an empirical study based on strategic emerging industries," *Nankai Economic Research*, no. 1, pp. 176–198, 2020.
- [7] A. D. Paul, H. H. Bronwyn, and A. T. Andrew, "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence," *Research Policy*, vol. 29, no. 4–5, pp. 497–529, 2000.
- [8] K. Aerts and D. Czarnitzki, "Using Innovation survey data to evaluate R&D Policy: the case of Belgium," *ZEW, Centre for European Economic Research Discussion*, vol. 21, no. 4, pp. 4–55, 2004.
- [9] R. B. James, M. Gustav, and C. P. Bruce, "What promotes R&D? Comparative evidence from around the world," *Research Policy*, vol. 44, no. 2, pp. 447–462, 2017.
- [10] C. Á. A. Inmaculada, K. Chihwa, and R. J. Desiderio, "Long run effect of public grants and tax credits on R&D investment: a non-stationary panel data approach," *Economic Modelling*, vol. 75, no. 11, pp. 93–104, 2018.
- [11] X. Li and N. He, "Is tax incentive conducive to enterprise technological innovation?" *Economic Science*, no. 1, pp. 18–30, 2019.
- [12] J. Zhan, X. Shao, and M. Xu, "The impact of government subsidies on R&D investment behavior of agricultural enterprises," *Scientific Research Management*, vol. 40, no. 4, pp. 103–111, 2019.
- [13] B. Mei and J. Luo, "Financial subsidies, R&D investment and enterprise innovation performance: a moderated mediating effect model test under institutional differences," *Economic Longitude*, vol. 37, no. 1, pp. 1–19, 2020.
- [14] C. M. David and N. L. Richard, "Spinning off and spinning on?: the federal government role in the development of the US computer software industry," *Research Policy*, vol. 25, no. 6, pp. 947–966, 1996.
- [15] F. Israel and K. Yoav, "Taxes and subsidies in a polluting and politically powerful industry," *Journal of Asian Economics*, vol. 15, no. 3, pp. 481–492, 2004.
- [16] J. Edler and L. Georghiou, "Public procurement and innovation-Resurrecting the demand side," *Research Policy*, vol. 36, no. 7, pp. 949–963, 2007.

## Research Article

# The Influence of Government Intervention on the Performance of Independent Innovation under Financial Support Based on Data of Listed Companies in Strategic Emerging Industries

Ze-Jiong Zhou,<sup>1</sup> Peng-Yue Zhang,<sup>2</sup> Miao-Miao Lu,<sup>1</sup> and Zhi Gao <sup>3</sup>

<sup>1</sup>School of Economics, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

<sup>2</sup>School of Finance, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

<sup>3</sup>Department of Physical Education, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

Correspondence should be addressed to Zhi Gao; gicy0415@163.com

Received 21 August 2020; Revised 16 October 2020; Accepted 17 October 2020; Published 30 October 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Ze-Jiong Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the data of the listed companies in strategic emerging industries in China, this paper uses GMM panel estimation method to measure the impact of government intervention on the performance of independent innovation and analyzes the impact of financial support on the effect of government intervention. The conclusions are as follows: Firstly, there is a lag effect in the performance of independent innovation, and there is also a lag effect in the patent application and main business income. In addition, there is an inverted U-shaped relationship between the patent application of one period lag and the current patent application, while there is no inverted U-shaped relationship between the main business income of one period lag and the current main business income. Secondly, government subsidies, additional deduction of R&D expenses, and value-added tax incentives have a significant effect on the number of patent applications, and the reduction of income tax burden can improve the main business income. Thirdly, after adding the financial support adjustment variables, we find that the influence direction of government intervention on the independent innovation performance has not changed, but the influence degree is weaker. Fourthly, the capital investment and labor input can significantly improve the performance of enterprise independent innovation.

## 1. Introduction

With the increasing uncertain factors of economic development, China urgently needs a new driving force of economic growth to ease the downward pressure of the economy. Strategic emerging industries play an important leading and exemplary role in economic development and have become the key areas to cultivate new impetus for economic development and realize economic transformation. In order to cultivate strategic emerging industries, the Chinese government has introduced a variety of supporting policies, such as taxation and subsidies, to improve the independent innovation ability of strategic emerging industries. However, the government only plays an “auxiliary role,” which is difficult to achieve all-round support, and it cannot completely solve the problems of low enthusiasm and

insufficient innovation ability of enterprises in strategic emerging industries. Therefore, the central government and local governments should set up special support funds, issue financial support policies, and improve the guidance mechanism of government intervention, so as to encourage enterprises in strategic emerging industries to carry out independent innovation to the greatest extent.

Scholars at home and abroad have rich research on the role of government intervention in enterprise independent innovation. The main research conclusions include incentive effect, inhibition effect, and ineffective use. Firstly, government intervention has an incentive effect. According to the national innovation system theory and triple helix theory, government intervention is one of the decisive factors to improve innovation capability [1, 2]. The empirical research results show that tax credit has a significant

incentive effect on enterprise innovation in the short term [3, 4], and government subsidies and tax incentives have a significant incentive effect on enterprise independent innovation in the long run [5–7]. Secondly, government intervention has a restraining effect. Some scholars have found that government intervention inhibits the independent innovation of enterprises in a specific industry and market environment [8–11]. Other scholars have found that when the innovation input is at a low threshold, the government intervention has a significant inhibitory effect on the independent innovation of enterprises [12–14]. Thirdly, government intervention is ineffective. Some studies believe that government intervention does not play a role in improving the independent innovation ability of enterprises [15–17].

At present, the research conclusions on the impact of government intervention on the independent innovation of enterprises have not been unified, and most scholars ignore the role of financial support in the study of government intervention in the independent innovation of enterprises. This paper takes financial support as a moderating variable into the dynamic panel data model of enterprise independent innovation performance analysis and uses the System GMM method to empirically study the influence of government intervention on enterprise independent innovation. The independent innovation performance of enterprises has the lag effect and accumulation effect, and the current independent innovation performance has an impact on the next period of independent innovation performance. Due to the endogenous problems in the performance of independent innovation, if the fixed effect model is used, the regression results will be biased and inconsistent estimates. Therefore, this paper constructs a dual dynamic panel data model for the performance analysis of enterprise independent innovation, which includes the dependent variable of one period of lag and the square term of dependent variable of one lag period.

The dual dynamic panel model constructed in this paper can reveal the dynamic and nonlinear relationship of economic variables. The model is widely used in the study of long-term equilibrium relationship and short-term nonlinear dynamic behavior of economic phenomena. Compared with the traditional dynamic panel data model, the dual dynamic panel data model established in this paper has made two improvements: one is to add the square term of dependent variable to reflect the nonlinear relationship between variables; the other is to select two dependent variables to reflect innovation output performance more comprehensively.

## 2. Research Design

**2.1. Sample Selection.** The sample of this paper is the listed strategic emerging industry enterprises in Shanghai and Shenzhen, and the research period is 2012–2018. We follow the following principles when selecting research samples: Firstly, according to *the latest guidance on industry classification of listed companies*, the strategic emerging industry listed companies in wind database are selected, and the initial starting time of the sample is set as 2012. Secondly, the

selected listed companies are in the growth period and have certain business operation ability and R&D innovation ability; that is, the listed companies have certain economic strength to support enterprises to carry out R&D innovation activities. Thirdly, ST and \*ST listed companies are excluded. Fourthly, the listed companies with a tax rate less than 0 or greater than 1 are excluded.

**2.2. Data Sources.** The financial information on listed companies in strategic emerging industries comes from wind database, and the number of patent applications and R&D personnel are from the Guotai'an database. Part of the data comes from the annual reports of listed companies.

### 2.3. Index Selection

**2.3.1. The Performance of Independent Innovation.** The independent innovation ability of enterprises is reflected not only in technology but also in economic benefits. Therefore, patent application and main business income are the indicators that can directly reflect the technical and financial performance. Patent application and main business income represent technical performance and enterprise financial performance, respectively, which can directly reflect the technological and economic achievements of independent innovation of strategic emerging industries. Therefore, in this paper, patent application and main business income are taken as the indicators of the performance of independent innovation.

- (1) Number of patent applications (PAT): patent is the most direct expression of the performance of enterprises' independent innovation. Some scholars use the number of patent applications and the number of patent authorizations as indicators to measure the performance of enterprises' independent innovation. Among them, patent applications can reflect the enterprise's knowledge stock and technological innovation ability, while patent authorization is affected by policies and licensing cycle and needs to be tested and paid annual fees. There are many administrative influencing factors, so it is difficult to reflect the level of the independent innovation of enterprises in that year. Patents can be divided into three types: invention patents, practical patents, and appearance patents. However, from the concept of independent innovation and dual innovation, appearance patents are not within the scope of independent innovation. Therefore, this paper takes a number of patent applications to measure the innovation ability of enterprises. Referring to the research of relevant scholars [18, 19], this paper conducts Winsorize treatment on 1% and 99% percentiles of the number of patents and then adds 1 to take the natural logarithm.
- (2) Main business income (MBI): the output indicators of enterprise independent innovation include new product income and main business income. It is



difficult for most enterprises to count the sales revenue of new products for several consecutive years, but the main business income can reflect the continuous economic benefit data after the R&D and innovation investment activities. Therefore, this paper takes the main business income as an indicator to measure the economic benefits of independent innovation.

**2.3.2. Government Intervention.** The explanatory variable is government intervention (G), and government subsidies and tax incentives are the main means of government intervention.

- (1) Government subsidy: Government subsidy (S) includes financial discount, research, and development subsidies and policy subsidies. According to the practices of relevant scholars [20], we select the government subsidy index under other income accounting subjects in the enterprise financial statements as the measurement index.
- (2) Tax preference: tax preference includes R&D expenses plus deduction, value-added tax preference, and income tax preference.

① Additional deduction of R&D expenses (EXP): additional deduction of enterprise R&D expenses is a kind of tax method and means for government intervention in enterprise R&D innovation activities. According to the actual amount of R&D expenditure, a certain proportion is added as the deduction of tax payable. Referring to the research of some scholars [21], this paper directly multiplies the actual amount of R&D expenses of enterprises by 50% during 2012–2016, and the actual amount of enterprise R&D expenses times 75% during 2017–2018, which are used as indicators to measure the additional deduction of R&D expenses.

② Value-added tax incentives (VAT): Value-added tax is one of the most important tax burdens of enterprises. The preferential policies of value-added tax mainly include VAT return and reduction and VAT deduction. Most of them are aimed at small- and medium-sized enterprises with insufficient profits. However, strategic emerging industries have certain enterprise scale and profitability. Most listed companies have no obvious preferential policies for tax return and exemption; especially, the sales of independent innovation products of strategic emerging industries need to go through a long period of time to enjoy VAT refund and exemption. Some scholars use VAT deduction as the VAT preference in the process of R&D and innovation, which mainly refers to the VAT deduction of new fixed assets. This paper uses the VAT deduction of newly added fixed assets of enterprises as the measurement index of value-added tax incentives [22].

③ Enterprise income tax (EIT): Enterprise income tax is an important tax burden of enterprises, and it is an important means of government intervention. Referring to the existing literature [23, 24], this paper directly uses income tax as the reverse indicator of tax preference to explain the impact of government intervention on independent innovation investment.

**2.3.3. The Input of Enterprise Independent Innovation.** The input of enterprise independent innovation includes capital input of independent innovation and personnel input of independent innovation.

- (1) Capital input of independent innovation (R&D): Capital input of independent innovation is the investment of strategic emerging industry enterprises for research and development, which is mainly used in basic research, applied research, and experimental development. In this paper, the expenditure in R&D is taken as a measure of the scale of capital input of independent innovation [25].
- (2) Personnel input of independent innovation (L): The personnel input of independent innovation is one of the important resources for enterprises' independent innovation and the key source of enterprise's independent innovation vitality. Therefore, this paper takes the proportion of technical personnel in the whole company as an index to measure the personnel input of independent innovation.

**2.3.4. Financial Support.** In the process of independent innovation of strategic emerging enterprises, financial capital plays an important role. It is an important capital input for enterprises to realize financing of independent innovation, medium-term resource integration, and value-added new products. Financial support (F) belongs to microenterprise level and will be affected by government intervention. Therefore, this paper takes financial support as a moderating variable to study the impact of financial support on independent innovation investment of enterprises. Financial support here includes bank loans and foreign direct investment.

- (1) Bank loan: Bank loans include short-term loans and long-term loans. ① Short-term loans from banks (SI): The special support fund policy for strategic emerging industries clearly points out that it is necessary to expand financial support means and support enterprises to implement major industrialization projects by means of bank loans and bond financing. Bank short-term loan is an important source of funds for enterprises to invest in innovation projects and activities. Referring to the existing literature, we select short-term bank loans as one of the indicators to measure financial support [26]. ② Long-term loans from banks (LI): bank long-term loan is the key source of funds for

enterprises to invest in R&D and innovation and maintain production and operation projects. It is mainly divided into fixed assets investment loan, renewal and transformation loan, science and technology development, and new product trial production loan. Referring to the practice of existing literature [27], this paper chooses the long-term loan of banks as another indicator to measure financial support.

- (2) Foreign direct investment (FDI): FDI is one of the important capital investments for enterprise independent innovation. Referring to the existing literature, we choose the funds borrowed from overseas of enterprises in the total project investment approved by the relevant government departments as the measurement of foreign direct investment [28].

**2.3.5. Control Variables.** In order to control the influence of other factors on the empirical results, we select size of the enterprise (size), profitability of the enterprise (ep), age of the enterprise (age), capital structure of the enterprise (lev), and ownership type of the enterprise (own) as control variables [29]. Generally speaking, if an enterprise's scale is larger, its profitability is stronger, and its capital is sufficient, the enterprise will be able to provide sufficient R&D and innovation funds. According to the life cycle theory of strategic emerging industries, the age of enterprises can reflect the development stage of enterprises and judge the investment in independent innovation. The level of capital structure can reflect the investment decision-making ability and debt-paying ability of listed companies for R&D innovation projects. The type of enterprise ownership determines whether the development of an enterprise is determined by the private or the government. In order to study the holding of state-owned enterprises, we take the value of state-owned holding as 0 and that of other controllers as 1.

The variables and their calculation methods are shown in Table 1.

**2.4. Model Construction.** The independent innovation of enterprises can realize knowledge increase, technological innovation, and economic benefit growth and meet the characteristics of Romer Jones' knowledge production function. Therefore, based on Romer Jones' knowledge production function, we establish the following performance function of the independent innovation of enterprises:

$$Y_{it} = [(1 - a_K)K_{it}]^\alpha [(1 - a_L)L_{it}]^{1-\alpha}, \quad (1)$$

where  $Y$  represents the performance of independent innovation,  $K$  represents the capital input of independent innovation,  $L$  represents the personnel input of independent innovation, and  $a$  and  $a - 1$  are elasticity coefficients of  $K$  and  $L$ , respectively.

The variables of government intervention and financial support belong to the capital input of enterprise R&D innovation. Therefore, the variables of government

intervention and financial support are added to equation (1), and the innovation performance ( $Y$ ) is replaced by the number of patent applications, so as to obtain the following two innovation performance functions:

$$\text{PAT}_{it} = [(1 - a_1)G_{it}]^\alpha [(1 - a_2)K_{it}]^\beta \cdot [(1 - a_3)L_{it}]^\gamma [(1 - a_4)F_{it}]^{1-\alpha-\beta-\gamma}, \quad (2)$$

$$\text{MBI}_{it} = [(1 - a_1)G_{it}]^\alpha [(1 - a_2)K_{it}]^\beta \cdot [(1 - a_3)L_{it}]^\gamma [(1 - a_4)F_{it}]^{1-\alpha-\beta-\gamma}, \quad (3)$$

where PAT represents the number of patent applications, MBI represents the enterprise's main business income, G represents the government intervention, and F represents the financial support.

In order to eliminate the heteroscedasticity problem, we establish the following panel dynamic models based on formulas (2) and (3):

$$\begin{aligned} \ln \text{PAT}_{it} = & \beta_0 + \beta_1 \ln \text{PAT}_{it-1} + \beta_2 \ln \text{AT}_{it-1}^2 + \beta_3 \ln S_{it} \\ & + \beta_4 \ln \text{EXP}_{it} + \beta_5 \ln \text{VAT}_{it} + \beta_6 \ln \text{EIT}_{it} + \beta_7 \ln K_{it} \\ & + \beta_8 \ln L_{it} + \beta_9 \ln \text{SI}_{it} + \beta_{10} \ln \text{LI}_{it} \\ & + \beta_{11} \ln \text{FDI}_{it} + \sum_{\lambda=1} \lambda \text{Contr}_{it} + \mu_i + \varepsilon_{it}, \end{aligned} \quad (4)$$

$$\begin{aligned} \ln \text{MBI}_{it} = & \beta_0 + \beta_1 \ln \text{PAT}_{it-1} + \beta_2 \ln \text{PAT}_{it-1}^2 + \beta_3 \ln S_{it} \\ & + \beta_4 \ln \text{EXP}_{it} + \beta_5 \ln \text{VAT}_{it} + \beta_6 \ln \text{EIT}_{it} \\ & + \beta_7 \ln K_{it} + \beta_8 \ln L_{it} + \beta_9 \ln \text{SI}_{it} + \beta_{10} \ln \text{LI}_{it} \\ & + \beta_{11} \ln \text{FDI}_{it} + \sum_{\lambda=1} \lambda \text{Contr}_{it} + \mu_i + \varepsilon_{it}, \end{aligned} \quad (5)$$

where  $\beta_0$  is a constant term,  $\beta_1, \beta_2, \dots, \beta_{11}$  are the coefficient corresponding to each variable,  $\lambda$  is the coefficient corresponding to the control variable,  $\mu_i$  is the year effect, which is used to control the impact of time change, and  $\varepsilon_{it}$  is the error term.

Due to the lag term and fixed effect of dependent variables in dynamic panel model, the traditional OLS method will lead to the deviation of estimation results. GMM method can effectively control the endogeneity of dynamic panel model and solve the problem of biased estimation. Compared with the difference GMM method, System GMM method can not only realize the estimation of difference equation but also realize the estimation of horizontal equation and eliminate the bias caused by short-term panel data [30]. Therefore, this paper uses System GMM method to analyze the dynamic panel model of enterprise independent innovation performance.

In the dynamic panel data model, the combination of highly continuous time series and significant individual heterogeneity may lead to the problem of weak instrumental variables in System GMM estimation. In view of the limitations of the model, we can choose the objective function of the continuous update estimator and use more effective tool

TABLE 1: The variables and their calculation methods.

Variable properties	Primary variable	Secondary variable	Variable symbol	Calculation method of variables
Explained variable	Performance of independent innovation	Number of patent applications	PAT	Logarithm of the number of enterprise patent applications plus 1
		Main business income	MBI	Logarithm of the income from main business
		Government subsidies	S	Logarithm of government subsidies
Explanatory variable	Government intervention	Additional deduction of R&D expenses	EXP	Logarithm of R&D expenses multiplied by the specific deduction proportion
		Value-added tax incentives	VAT	Logarithm of VAT deduction for new fixed assets of enterprises
		Income tax	EIT	Logarithm of enterprise income tax
		Capital input of independent innovation	K	Logarithm of R&D expenditure
Adjustment variable	Input of enterprise independent innovation	Personnel input of independent innovation	L	Logarithm of the number of R&D personnel
		Short-term loans from banks	SI	Logarithm of bank short-term loan
		Long-term loans from banks	LI	Logarithm of long-term bank loans
		Foreign direct investment	FDI	Logarithm of capital borrowed from abroad
		Size of the enterprise	size	Logarithm of the total assets of the enterprise
Control variable	Intervention within enterprises	Profitability of the enterprise	ep	Ratio of net profit to operating income
		Age of the enterprise	age	The difference between research year and listed year
		Capital structure of the enterprise	lev	Asset liability ratio of enterprises
		Ownership type of enterprise	own	The holding value of state-owned enterprises is 0, and the value of others is 1

variables to improve and optimize the System GMM, so that it can avoid the use of weak tool variables and have a good large sample and limited sample properties.

### 3. Empirical Results and Analysis

Because of the endogenous problems in the performance of independent innovation, we first test and analyze the lag effect of independent innovation and then empirically study the influence of government intervention and financial support on the output stage of independent innovation.

*3.1. Model Test.* Through the autocorrelation test, it is found that AR (1) accepts the original hypothesis and AR (2) rejects the original hypothesis, which indicates that there is an endogenous problem. In this paper, stata12.0 software is used to conduct Hansen test and Sargan test. It is found that rejecting the original hypothesis that all variables are exogenous variables indicates that the selection of test tool variables is effective, and there is no problem of overidentification of tool variables. It shows that the tool variables selected are effective; that is, the econometric model is robust. We further tested the significance of the joint coefficient. The test results show that the dynamic panel estimation results are reliable.

*3.2. Lag Effect Analysis of Independent Innovation Output.* The empirical results are shown in Table 2. The regression results of models 1 and 3 reflect the impact of government

intervention on patent application and main business income of strategic emerging enterprises. The performance variables of the lag period are added in both models to study the lag effect of independent innovation. From the regression result of model 1, in the case of only government intervention, the number of patent applications in the previous period has a positive correlation with the number of patent applications in the current period and has passed the 1% significance level. The correlation coefficient is 0.503; that is, for every 1% increase in the number of patent applications in the previous period, the number of patent applications in the current period will increase by 0.503%, indicating that, in the process of R&D and innovation, knowledge has continuity and accumulation, and the status of R&D and innovation in the early stage has a great impact on the current period.

In order to test the nonlinear relationship between the number of patent applications in the previous period and the current patent applications, we add the square term of the number of patent applications in the previous period into the regression analysis model. Regression analysis results show that the correlation coefficient between the square term of the previous patent application and the current patent application is negative, indicating that there is an inverted U-shaped relationship between the number of patent applications in the previous period and that in the current period. The reason is that after the patent application is successful, the period of patent protection is fixed. Even within the protection period, there will still be many enterprises imitating in the market, which will damage the



TABLE 2: Dynamic panel model estimation results of independent innovation output stage.

Variables	lnPAT		Variables	lnMBI	
	Model 1	Model 2		Model 3	Model 4
$\ln PAT_{it-1}$	0.5030*** (13.87)	0.4583*** (12.27)	$\ln MBI_{it-1}$	1.1920*** (12.85)	1.1056*** (10.06)
$\ln PAT_{it-2}^2$	-0.0867* (-8.2)	-0.0676* (-6.62)	$\ln MBI_{it-2}^2$	-0.00169 (-0.10)	-0.0013 (-0.09)
lnS	0.0118* (1.91)	0.0106 (1.06)*	lnS	0.0062 (-0.57)	0.0017 (-0.06)
lnEXP	0.0968*** (4.83)	0.9300*** (4.05)	lnEXP	-0.0005 (-0.13)	-0.0056 (-0.40)
lnVAT	0.0136*** (2.37)	0.0125** (2.07)	lnVAT	-0.0032 (-0.22)	-0.0033 (-0.23)
lnEIT	-0.0001 (-0.01)	-0.0097 (-0.29)	lnEIT	-0.0246** (-2.18)	-0.0252** (-2.70)
lnK	1.0120*** (10.77)	1.1150*** (11.36)	lnK	0.6051*** (14.25)	0.5311*** (14.24)
lnL	1.0918*** (9.12)	1.1587*** (9.48)	lnL	0.0818*** (4.88)	0.0388** (1.98)
lnSI	—	0.0737(4.43)	lnSI	—	0.1813*(5.28)
lnLI	—	-0.0608(-0.4)	lnLI	—	-0.1721**(-2.24)
lnFDI	—	0.0014(0.40)	lnFDI	—	0.0004(0.10)
lnsize	0.0248 (0.17)	0.0231 (0.11)	lnsize	0.01583 (0.09)	0.01517 (0.06)
Ep	0.00489 (0.009)	0.007 (0.0038)	ep	0.0398 (0.19)	0.0324 (0.15)
age	0.00733 (0.88)	0.00639 (0.34)	age	0.0347 (1.72)	0.0299 (1.35)
lev	0.00165 (0.51)	0.00929 (1.22)	lev	0.0199 (0.76)	0.00677 (0.94)
Own	-0.1590*** (-0.64)	-0.2190*** (-1.94)	own	-0.0796*** (-0.64)	-0.262*** (-0.46)
Constant	0.6441*** (3.09)	-4.3547 (-0.6)	Constant	2.0382*** (4.72)	-0.5776 (-0.43)
AR(1)	0.0002	0.0026	AR(1)	0.000	0.0002
AR(2)	0.9182	0.4197	AR(2)	0.175	0.331
Hansen	0.5183	0.5595	Hansen	0.6288	0.6735
Sagan test	0.88	0.92	Sagan test	0.96	0.9
Wald test	0.000	0.000	Wald test	0.000	0.000

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ , and the corresponding values in brackets are  $t$  values.

economic interests of innovators. That is to say, the more the R&D and innovation achievements of innovators are, the more the economic losses they will suffer, until the loss of economic benefits exceeds the economic benefits brought by patented inventions and the number of patent applications decreases.

From the regression result of model 3, the main business income of the previous period has a positive effect on the current main business income, and the correlation coefficient is 1.1920. If the economic benefit of the enterprise is improved, the enterprise will expand production and further reduce the unit production cost. Therefore, enterprises will increase R&D innovation efforts, and the economic benefits of enterprises will continue to improve without decreasing the scale economy.

After adding the square item of the main business income of the lag phase, it failed to pass the significance level of 10%, which indicates that there is no “inverted U” relationship between the square item of the main business income of the lag phase and the current main business income, which corresponds to the “inverted U” relationship of patent application. The regression results show that there is no “inverted U” relationship in the main business.

**3.3. The Impact of Government Intervention on the Output of Independent Innovation.** From the results of model 1 and model 3, the coefficient between government subsidies and patent applications is positive at the 10% significance level; the estimation coefficient of government subsidies and main business income is positive, but it does not pass the significance level of 10%. This shows that government subsidies have a certain positive effect on the increase of patent

applications, but its effect on improving the main business income is not obvious.

The estimated coefficient of additional deduction of R&D expenses and patent application is positive and passes the significance level of 1%. The correlation coefficient is 0.0968, which indicates that for every 1% increase in additional deduction of R&D expenses, the enterprise patent application increases by 0.9680%. It can be seen that additional deduction of R&D expenses has a positive effect on patent application. The estimated coefficient of additional deduction of R&D expenses and main business income is negative and fails to pass the significance level of 10%, which indicates that additional deduction of R&D expenses has a negative impact on the main business income, but it is not obvious.

The estimated coefficient of value-added tax incentives and patent application is positive and has passed the significance level of 5%, which indicates that the value-added tax incentives promote the patent application of enterprises. The estimated coefficient is 0.0136, which means that the number of patent applications increases by 0.0136% for every 1% increase of value-added tax incentives. The correlation coefficient between the value-added tax incentives and the enterprise's main business income is negative, but it fails to pass the significance level of 10%, which indicates that the value-added tax incentives have a negative impact on the enterprise's main business income, but this effect is not obvious.

The estimated coefficient of income tax and patent application is negative, but it fails to pass the significance test, which indicates that the influence of tax preference on enterprise patent application is not obvious. The estimated

coefficient of income tax and the enterprise's main business income is negative, and the significance test shows that tax preference can promote the improvement of the enterprise's main business income.

The reason is that the government has increased government subsidies, an additional deduction of R&D expenses, and value-added tax concessions to the independent innovation of strategic emerging industries, which is equivalent to reducing the R&D innovation cost of enterprises, but it does not directly increase the economic benefits of enterprises. Therefore, the effect of government intervention measures such as government subsidies, an additional deduction of R&D expenses, and value-added tax incentives on the main business income of enterprises is not significant. The reduction of income tax can directly increase the income of enterprises. Therefore, the impact of income tax preference on the main business income is more obvious.

*3.4. The Influence of Government Intervention on the Performance of Independent Innovation under Financial Support.* On the basis of government intervention, the financial support adjustment variable is added to get the regression results of model 2 and model 4, so as to further analyze the impact of government intervention on the presence of financial support.

According to the regression result of model 2, the estimated coefficients of government subsidies, an additional deduction of R&D expenses, and value-added tax incentives are still positively correlated with enterprise patent application, and the estimated coefficient of income tax and enterprise patent application is also negatively correlated. Similar to the results of model 1, except for income tax, the correlation coefficients between the other three kinds of government intervention and enterprise patent application have passed the significance test, indicating that government intervention has a positive incentive effect on the number of enterprise patent applications.

According to the regression result of model 4, considering the financial support, the influence of government intervention on the main business income of enterprises has not changed significantly. The estimated coefficient between government subsidies and the enterprise's main business income is positive, but it fails to pass the significance test. The estimated coefficients of additional deduction of R&D expenses, value-added tax incentives, and the enterprise's main business income are negative and have not passed the significance test. The estimated coefficient of income tax and the enterprise's main business income is negative and has passed the significance test.

According to the regression results of model 2 and model 4, the estimated coefficient of short-term loans from banks and patent application is positive, but it fails to pass the significance test. The estimated coefficient of short-term loans from banks and main business income is positive, which has passed the significance level of 10%. It can be seen that short-term loans from banks have no obvious impact on patent application, but it has an obvious deposit effect on the increase of main business income. The reason is that short-

term loans from banks are the financing of enterprises through financial institutions, which are mainly used for the investment of high-yield projects. The economic benefits of these financing far exceed the patent application.

The estimated coefficient of long-term loans from banks and patent application is negative, but it fails to pass the significance test, indicating that long-term loans from banks will hinder enterprise patent application, but this effect is not obvious. The estimation coefficient of long-term loans from banks and main business income is also negative, but the significance test shows that long-term loans from banks will reduce the increase of the enterprise's main business income. The reason is that long-term loans from banks will bring great financial risks to enterprises. When the profit before interest and tax is reduced, the profit after tax will be reduced by a greater margin. The estimated coefficients of FDI, patent application, and main business income are all positive, but they have not passed the significance test.

From the perspective of control variables, the estimation coefficients of enterprise scale, profitability, debt repayment ability, enterprise age, and independent innovation performance are positive, indicating that the larger the enterprise scale, the stronger the profitability and debt-paying ability, and the longer the production and operation time, the greater the performance of independent innovation. The estimated coefficients of enterprise nature, patent application, and main business income are all negative, indicating that the state-owned enterprise system is not conducive to enterprise independent innovation.

*3.5. The Impact of the Input of Enterprise on the Performance of Independent Innovation.* Capital and personnel are important input variables of independent innovation. According to the regression analysis results of model 1 and model 3, the correlation coefficients of capital input of independent innovation, personnel input of independent innovation, and patent application are all positive and pass the significance test. The two correlation coefficients are 1.0120 and 1.0918, respectively, which indicates that the patent application increases by 1.0120% and 1.0918% for every 1% increase in capital and personnel input. The correlation coefficients of the capital input of independent innovation, personnel input of independent innovation, and main business income are also positive and pass the significance test. The two correlation coefficients are 0.6051 and 0.0818, respectively, which indicates that the main business income increases by 0.6051% and 0.0818% for each 1% increase in capital input of independent innovation and personnel input of independent innovation. It can be seen from the research results that personnel input of independent innovation has a better promotion effect on enterprise independent innovation, which is because innovative talents provide new knowledge and technology for enterprises and are the developers and creators of enterprise R&D innovation.

According to the regression analysis results of model 2 and model 4, after adding financial support as a moderating variable, the estimated coefficients of capital input of independent innovation, personnel input of independent

innovation, and patent applications are still positive, but the estimation coefficient becomes larger, which indicates that financial support can increase the promotion effect of capital and personnel input of independent innovation on enterprise's independent innovation performance. The estimated coefficients of capital input of independent innovation, personnel input of independent innovation, and main business income are positive, and the estimated coefficient becomes smaller, which indicates that financial support reduces the promotion effect of capital and personnel input of independent innovation on the independent innovation performance of enterprises.

The reason is that capital input and personnel input of independent innovation are directly applied to enterprise R&D activities, which has a direct impact on enterprise patent application. However, the impact of capital and personnel input of independent innovation on the enterprise's main business income needs to go through the production and operation stage of the enterprise, and the impact effect becomes weaker.

#### 4. Conclusions and Suggestions

The main conclusions of this paper are as follows: Firstly, the performance of enterprise independent innovation has a lag effect and there is an inverted U-shaped relationship. The patent application and main business income with a lag period have a significant incentive effect on the current patent application and main business income. At the same time, after adding the square term of independent innovation output of the lag phase, there is an inverted U-shaped relationship between the patent application of the lag phase and the current patent application, but there is no inverted U-shaped relationship between the main business incomes. Secondly, there are differences in the effects of various means of government intervention on independent innovation performance of strategic emerging industry enterprises. Government subsidies, R&D expenses' additional deduction, and value-added tax incentives have a significant effect on increasing the number of patent applications, while the reduction of income tax burden can improve the main business income. Thirdly, after the financial support is added, short-term loans from banks have no significant effect on the patent application of enterprises, but it has a significant effect on the main business income of enterprises. Fourthly, capital input and personnel input can significantly increase the number of patent applications and increase the main business income of enterprises.

On how to promote the effect of government intervention on independent innovation of strategic emerging industry enterprises, we put forward the following suggestions:

First of all, the government intervention means of enterprise's technical performance and financial performance need to be differentiated, such as increasing government subsidies, R&D expenses' additional deduction, and value-added tax incentives to improve enterprise's technical performance and reducing enterprise's income tax burden to improve enterprise's financial performance.

Secondly, we should improve the innovation of financial support tools and encourage enterprises in strategic emerging industries to carry out independent innovation. Financial support is an important regulatory variable for the government to intervene in the independent innovation of strategic emerging industries and an important signal for the government to release policy dividends. The innovation of financial support tools can solve the problem of financing difficulties of strategic emerging industry enterprises, so as to mobilize the enthusiasm of independent innovation of enterprises.

Finally, we should pay close attention to the short-term goal of government intervention, closely monitor the independent innovation activities of strategic emerging industry enterprises, and prevent patent duplication and product duplication.

#### Data Availability

The data used to support the findings of this study are included within the article.

#### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

#### Acknowledgments

This work was supported by the project of Anhui University Humanities and Social Sciences Research (Grant no. SK2020A0011) and Key Project of National Social Science Foundation (Grant no. 17AJY018).

#### References

- [1] R. R. Nelson, *National Innovation Systems: A Comparative Analysis*, Oxford University Press, Oxford, UK, 1993.
- [2] E. Henry and L. Loet, "The dynamics of innovation: from national systems and mode 2 to a triple helix of university-industry-government relations" *Reserch Policy*, vol. 29, no. 2, pp. 109–123, 2000.
- [3] N. Rao, "Do tax credits stimulate R&D spending? The effect of the R&D tax credit in its first decade," *Journal of Public Economics*, vol. 140, no. 8, pp. 1–12, 2016.
- [4] J. R. Brown, G. Martinsson, and B. C. Petersen, "What promotes R&D? Comparative evidence from around the world," *Research Policy*, vol. 46, no. 2, pp. 447–462, 2017.
- [5] X. H. Zhang and Y. D. Lu, "An empirical study on the impact of government subsidies on enterprise R&D Investment," *Science and Technology Management Research*, vol. 34, no. 15, pp. 204–209, 2014.
- [6] Y. Y. Chen, "The impact of additional deduction policy on R&D investment of enterprises—an empirical analysis based on panel data of enterprises in a city," *Taxation*, vol. 11, pp. 88–93, 2015.
- [7] J. T. Zhan, X. J. Shao, and M. Xu, "The effect of government subsidies on R&D investment behavior of agricultural enterprises," *Scientific Research Management*, vol. 40, no. 4, pp. 103–111, 2019.

- [8] F. Israel and Y. Kislev, "Taxes and subsidies in a polluting and politically powerful industry," *Journal of Asian Economics*, vol. 15, no. 3, pp. 481–492, 2004.
- [9] J. Edler and L. Georghiou, "Public procurement and innovation-Resurrecting the demand side," *Research Policy*, vol. 36, no. 7, pp. 949–963, 2007.
- [10] K. Torani, R. Gordon, and Z. David, "Innovation subsidies versus consumer subsidies: a real options analysis of solar energy," *Energy Policy*, vol. 92, no. 5, pp. 255–269, 2006.
- [11] M. Evan, "Global kerosene subsidies: an obstacle to energy efficiency and development," *World Development*, vol. 99, no. 11, pp. 463–480, 2017.
- [12] L. L. Zuo and Y. F. Zhang, "Political connections, government subsidies and innovation performance," *Shanghai Management Science*, vol. 41, no. 4, pp. 39–45, 2019.
- [13] J. Z. Yan and N. N. Qi, "Empirical analysis on the influencing factors of innovation performance of China's strategic emerging industries—taking energy conservation and environmental protection industry as an example," *Journal of Henan University of Technology (Social Science Edition)*, vol. 15, no. 1, pp. 71–78, 2019.
- [14] Z. X. Zheng and H. Y. Chen, "Impact of fiscal policy on innovation performance of strategic emerging industries: based on empirical analysis," *Journal of Chongqing University of Technology and Technology (Social Science Edition)*, vol. 36, no. 2, pp. 7–12, 2019.
- [15] S. Filipe and C. Carreira, "Do financial constraints threaten the innovation process? Evidence from Portuguese firms," *Economics of Innovation and New Technology*, vol. 21, no. 8, pp. 701–736, 2011.
- [16] N. Droste, B. Hansjürgens, P. Kuikman et al., "Steering innovations towards a green economy: understanding government intervention," *Journal of Cleaner Production*, vol. 135, no. 1, pp. 426–434, 2016.
- [17] P. Jason and K. Tim, "Economics of innovation in Australian agricultural economics and policy," *Economic Analysis and Policy*, vol. 54, no. 6, pp. 96–104, 2017.
- [18] M. J. Li and M. N. Zheng, "Substantive innovation or strategic innovation? The impact of macro industrial policies on micro enterprise innovation," *Economic Research*, vol. 51, no. 4, pp. 60–73, 2016.
- [19] H. J. Huang, C. J. Lv, and X. W. Zhu, "Second generation involvement and enterprise innovation: evidence from Chinese family listed companies," *Nankai Management Review*, vol. 21, no. 1, pp. 6–16, 2018.
- [20] Y. Yang, J. Wang, and Z. Lin, "Influence mechanism and optimization path of government subsidies on business performance of listed cultural enterprises: an empirical analysis based on panel data of 191 listed cultural enterprises," *Cultural Industry Research*, vol. 1, pp. 168–183, 2018.
- [21] R. Y. Han and H. T. Ma, "Tax incentives and enterprise R&D Investment: an empirical test based on double difference model," *Journal of Central University of Finance and Economics*, vol. 3, pp. 3–10, 2019.
- [22] C. Y. Wang, "Does tax incentives stimulate enterprise R&D investment?" *Scientific Research*, vol. 35, no. 2, pp. 255–263, 2017.
- [23] X. Wang and Y. He, "Government subsidies, tax incentives and enterprise R&D investment based on dynamic panel system GMM analysis," *Technical Economy and Management Research*, vol. 4, pp. 92–96, 2017.
- [24] X. J. Li and N. He, "Is tax incentive conducive to enterprise technological innovation?" *Economic Science*, vol. 1, pp. 18–30, 2019.
- [25] Y. L. Song and X. Chen, "Incentive or inhibition: policy environment and R&D Investment—based on the perspective of equity structure," *Science and Technology Management Research*, vol. 38, no. 24, pp. 134–143, 2018.
- [26] L. M. Zhen and D. L. Luo, "Credit rent-seeking, financial mismatch and their impact on enterprise innovation behavior," *Industry Economic Review*, vol. 10, no. 4, pp. 68–80, 2019.
- [27] H. Liu, *Research on Financial Support to Promote the Development of Strategic Emerging Industries*, Shandong University, Jinan, China, 2014.
- [28] Z. B. Tan and Y. Zhao, "Crowding in and crowding out effect of Foreign direct investment: the impact of financial development," *Financial Research*, vol. 9, pp. 69–83, 2014.
- [29] X. Y. Tan, *Research on the Relationship between Government Subsidies and Innovation Performance of High-Tech Enterprises*, Xi'an University of Electronic Science and Technology, Chengdu, China, 2019.
- [30] R. Blundell and S. Bond, "GMM estimation with persistent panel data: an application to production functions," *Econometric Reviews*, vol. 19, no. 3, pp. 321–340, 2000.



## Research Article

# Locating and Tracking Model for Language Radiation Transmission Based on Neural Network and FAHP

SongGui Zhu,<sup>1</sup> Hailang He ,<sup>2</sup> and Yuanyuan Zheng<sup>3</sup>

<sup>1</sup>College of Civil and Environmental Engineering, Anhui Xinhua University, Hefei, China

<sup>2</sup>Lanzhou Institute of Seismology, China Earthquake Administration, Lanzhou 730000, China

<sup>3</sup>College of Mechanical and Electrical Engineering, Anhui Jianzhu University, Hefei, China

Correspondence should be addressed to Hailang He; hhlmw@163.com

Received 15 August 2020; Revised 30 September 2020; Accepted 13 October 2020; Published 27 October 2020

Academic Editor: Shaohui Wang

Copyright © 2020 SongGui Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of internationalization, the distribution of languages and the office addresses of multinational companies are changing constantly. This paper makes the following research and exploration on this phenomenon: impact on the development of languages around the world. This paper studies the changes of native and second-language users and uses the historical data to predict the development trend by using the gray number series prediction model. Get the types of factors that affect the second language. Then, use fuzzy analytic hierarchy process to calculate the score of each factor. Finally, the global language trend equation is simulated: predictions for the development of language. In this paper, radiation propagation is calculated, and the method of CNN neural network is used to train big data, and the language trend positioning equation is drawn. Finally, the optimal language is obtained by using wavelet analysis and linear programming at different addresses. About model checking, according to the model's internal prediction ability and the significance of internal parameters, it is concluded that the model has high practicability, sensitivity, and stability.

## 1. Introduction

*1.1. Background.* There are currently about 6900 kinds of languages spoken on Earth. Nearly half of the world's population is using the following ten languages: Mandarin (including Standard Chinese), Spanish, English, Hindi, Arabic, Bengali, Portuguese, Russian, Punjabi, and Japanese [1]. Therefore, these ten languages are called the top ten languages in the world. Much of the world's population can speak the second language. And the development of language tends to be more unified [2].

In the context of economic globalization, many languages are on the brink of extinction [3]. Language is not only a part of culture but also the carrier of culture and the accumulation of traditional culture. If any one language in the world is endangered or even extinct, then the diversity of the world's languages and cultures will be undermined. And it will bring enormous irreparable damage to the colorful language and cultural images of humankind [4].

*1.2. Our Tasks.* Since a large multinational service company needs the true internationalization, the company proposes to open more international offices, which has already setup offices in New York City in the United States and Shanghai, China.

Here are our tasks:

- (1) Determine the influencing factors and the extent of language development. According to the data of the global population and language users over the years, simulate the distribution of users in different languages over time.
- (2) Predict the changing trend of the total number of native speakers and language users in the next 50 years. And analyze whether the top 10 languages in the world will be replaced.
- (3) According to the established model, predict the global population and population migration patterns in the next 50 years, and determine whether the

geographical distribution of languages will change in the same period.

- (4) Analyze the establishment of six international offices under the identified model. Determine the language in which they are spoken (including English), judging whether there are different impacts in the short and long term.
- (5) To save client company resources, analyze whether it is possible to reduce the number of international offices and design management methods.

## 2. Assumptions

By analyzing the article data, we propose the following assumptions to complete our model:

- (1) Unpredictable high-impact and low-probability events will not occur in the next 50 years.
- (2) For convenience, we assume that a country has only one mother tongue, only the mother tongue or the second-language speaker, and does not consider the third language, etc.
- (3) Various languages do not lead to catastrophic jumps in evolution over time.
- (4) Except for the factors we considered, other factors have minimal or negligible impact on the model.
- (6) Because there are so many languages in the world, we only consider the impact of the main language. Because of its large proportion of the main language, its development can identify the reliability of the model.

## 3. Symbol Description and Noun Explanation

- (1) GM (1, 1): a single-sequence line dynamic model in the gray prediction model is mainly used for time series prediction
- (2) Train: the data is circulated once
- (3) Filter: data-processing weight matrix
- (4) Wavelet operation: an efficient algorithm for graphics compression and recognition, which is widely used in various fields where compression of data is required.

## 4. The Model

### 4.1. Basic Model (Determination of Influencing Factors)

4.1.1. *Determination of the Influencing Factors.* There are mainly two factors that affect the distribution of language users: one is a native speaker and the other is the number and distribution of second-language users [5].

The former is determined by two factors: the number of countries that speak the language as their mother tongue and the population of these countries [6].

The latter is determined by the following factors:

- (1) The number of people using the language
- (2) The economic strength of the country that uses the language
- (3) The literary and social status of the country where the language is used
- (4) The degree of importance of language of science and diplomacy (if it is the official language of the United Nations, then we will increase its importance)

For convenience, we depict the following diagram, as shown in Figure 1.

4.1.2. *The Proportion of the Influential Factors.* Weigh the evaluation criteria in Section 4.1.1, considering the use of fuzzy analytic hierarchy processes [7]. We have the following scoring criteria:

- (1) Number of native speakers in this language: highest score 40.
- (2) Number of people who speak the language as their second language: highest score 60.
  - (1) The economic strength of the country using that language, with a maximum score of 20
  - (2) The importance of the language in science and diplomacy, with a maximum score of 15.
  - (3) The number of countries and populations using that language, with a maximum score of 15.
  - (4) The social and literary status of the language, with a maximum score of 10.
  - (5) If the language is for the United Nations, we will add 5.

The results of the top ten languages in use are shown in Table 1:

4.1.3. *The Equation of the Trend of Language Development.* The development of future language trends depends on two aspects: native speakers and second-language speakers [8].

The development of future language trends depends on two aspects. The formula is

$$Y = \lambda_1 X_1 + \lambda_2 X_2. \quad (1)$$

In formula (1), the number of native speakers is  $X_1$  and the number of secondary speakers is  $X_2$ .

- (a) The number of native speakers is mainly determined by the changing trend of the population of countries in which the language is the mother tongue. We use the gray number sequence prediction model to describe the language trend. In the gray prediction model, GM (1, 1) is a linear dynamic model of first-order single sequence, which is mainly used for time series prediction.



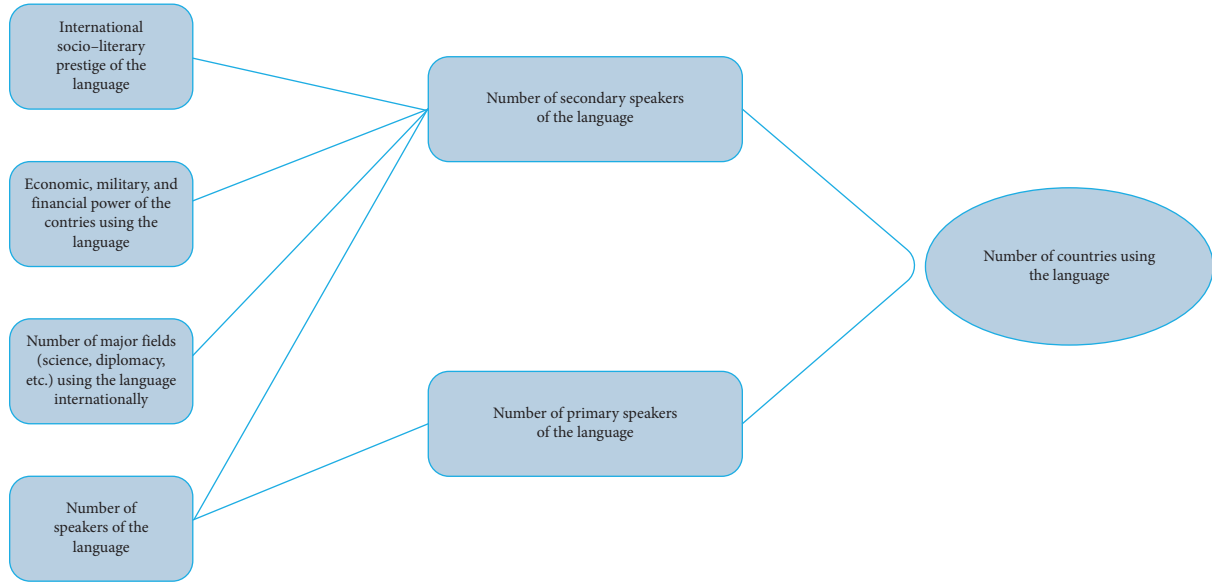


FIGURE 1: Relationship between influential factors.

TABLE 1: The ranking and occupying ratio of language usage.

Language	Native speakers	Total	The proportion (%)	Second-language speakers	Number of countries using the language	The total weight (%)	Total ranking
Mandarin Chinese	8.97	75	11.96	1.93	10.9	14.53	1
Spanish	4.36	75	5.81	0.91	5.27	7.03	2
English	3.71	75	4.95	6.11	9.82	13.09	3
Hindustani (Hindi/Urdu)	3.29	75	4.39	2.15	5.44	7.25	4
Arabic	2.9	75	3.87	1.32	4.22	5.63	5
Bengali	2.42	75	3.23	0.19	2.61	3.48	7
Portuguese	2.18	75	2.91	0.11	2.29	3.05	6
Russian	1.53	75	2.04	1.13	2.66	3.55	8
Punjabi	1.48	75	1.97	0	1.48	1.97	10
Japanese	1.28	75	1.71	0.01	1.29	1.72	9

Let GM(1, 1) be the original form:

$$x_1^{(0)}(k) + ax_1^{(1)}(k) = b,$$

$$x_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n)), \quad (2)$$

$$x_1^{(1)} = (x_1^{(1)}(1), x_1^{(1)}(2), \dots, x_1^{(1)}(n)).$$

Set the basic form of GM(1, 1) to the following formula:

$$x_1^{(0)}(k) + a_2 z^{(1)}(k) = b, \quad (3)$$

where

$$z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)), \quad (4)$$

$$z^{(1)}(k) = \frac{(x_1^{(1)}(k), x_1^{(1)}(k-1))}{2} \quad (k = 2, 3, \dots, n). \quad (5)$$

We define the matrix vector as follows:

$$u = \begin{bmatrix} a \\ b \end{bmatrix}, \quad (6)$$

$$F = \begin{bmatrix} x_1^{(0)}(1) \\ x_1^{(0)}(2) \\ \dots \\ x_1^{(0)}(n) \end{bmatrix}, \quad (7)$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \dots & \dots \\ -z^{(1)}(n) & 1 \end{bmatrix}. \quad (8)$$

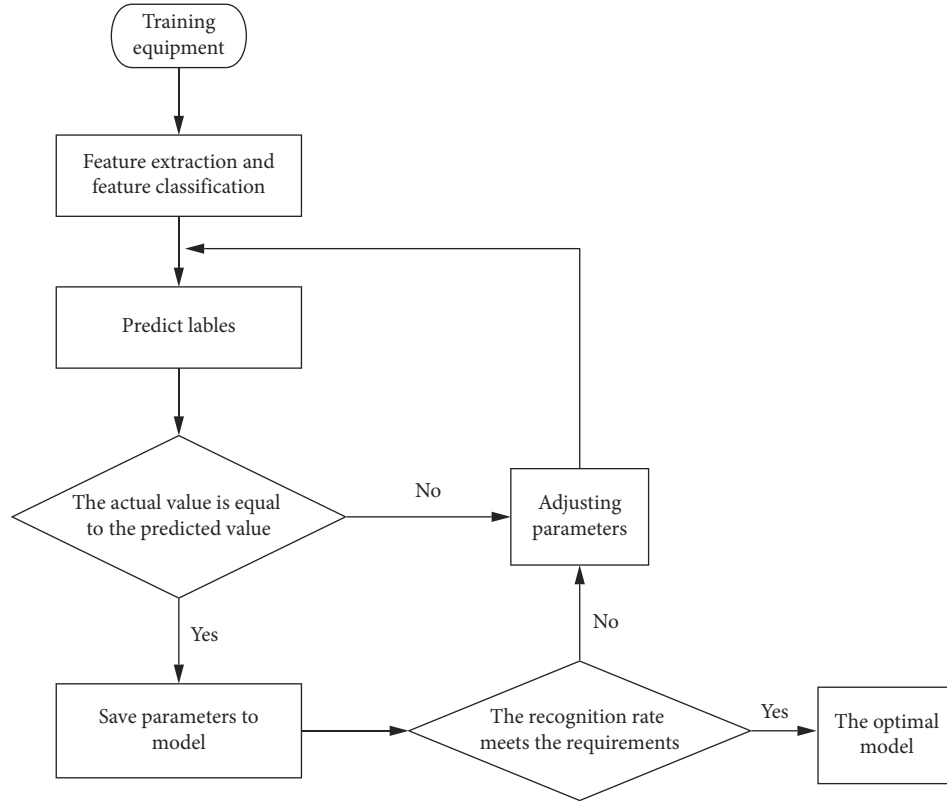


FIGURE 2: Flowchart of CNN neural network.

Putting these data into the above formula, we get a, b.

(b) For the second-language trend forecast and from the above factors, we have

$$x_2 = \eta x_1, \quad (9)$$

Where  $x$  is the language synthesis coefficient.

Therefore, from the above formula, we have the language trend equation:

$$Y = (\lambda_1 + \lambda_2 \eta) F. \quad (10)$$

**4.1.4. Accurate Positioning through the Radial Communication of the Global Language.** The spread of language is affected by the above factors, but we think the root cause is the radiative propagation of the source language. We use CNN neural network to do big data simulation [9] and depict a flowchart to indicate the exact location of the steps, as shown in Figure 2.

We simulate the distribution, assuming the location coordinates  $(x, y)$  and the time factor  $t$ . Big data simulation is used for filtering to find the specific feature  $N$  in the input. Eigenvalues are used for prediction equations and high precision calculations [10].

We have the following equations:

$$Y(\eta, t, s) = \frac{\xi}{\sum_{n=0}^{x_1-1} ((x_2/t)^n/n!) + ((\eta/t)^n/n!)}, \quad (11)$$

$$\xi = \frac{1}{xt(1 - (\eta/yt))}. \quad (12)$$

## 4.2. Application of the Model

**4.2.1. Distribution of Language Users.** Based on the global language trend positioning model, as time goes by, due to the variety of languages, here we select only the top 5 languages of the world to simulate the global rough distribution of speakers of this language in the next 50 years. With the office as the point and the contact between each office and the outside world as the degree, the topological relationship can be used for analysis [11].

It is clear from the pictures that the above five languages all have the following development trends:

- (1) The population of countries that are native speakers of the language is on the rise
- (2) The number of countries in that language as a second language is also on the rise

**4.2.2. Future Language Development Trend.** According to the change of the number of language users in the first 50

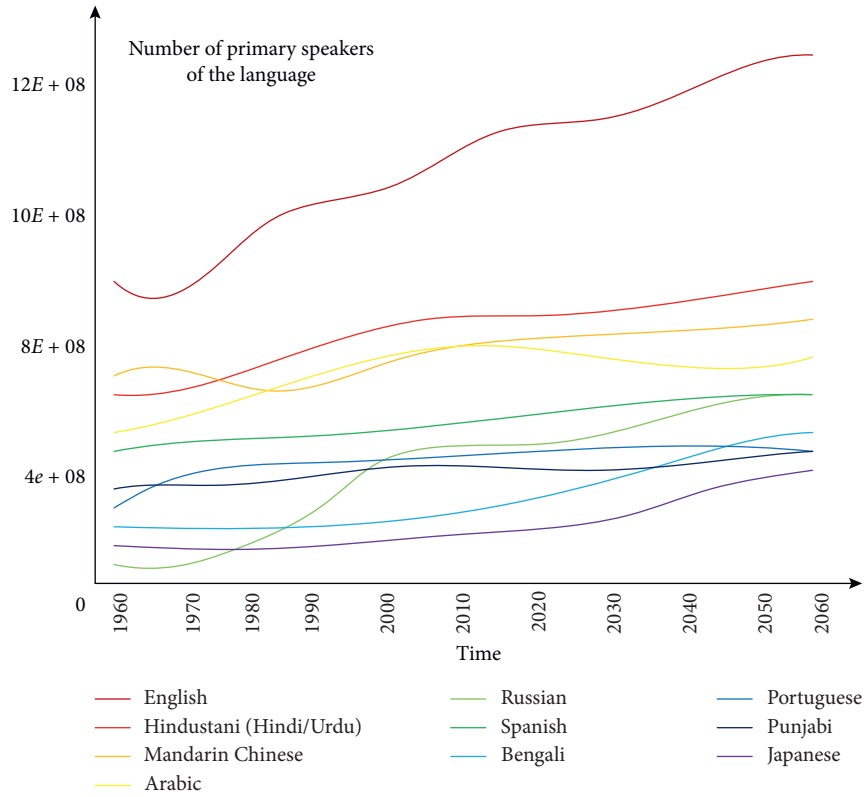


FIGURE 3: Number of native speakers from 1960 to 2060.

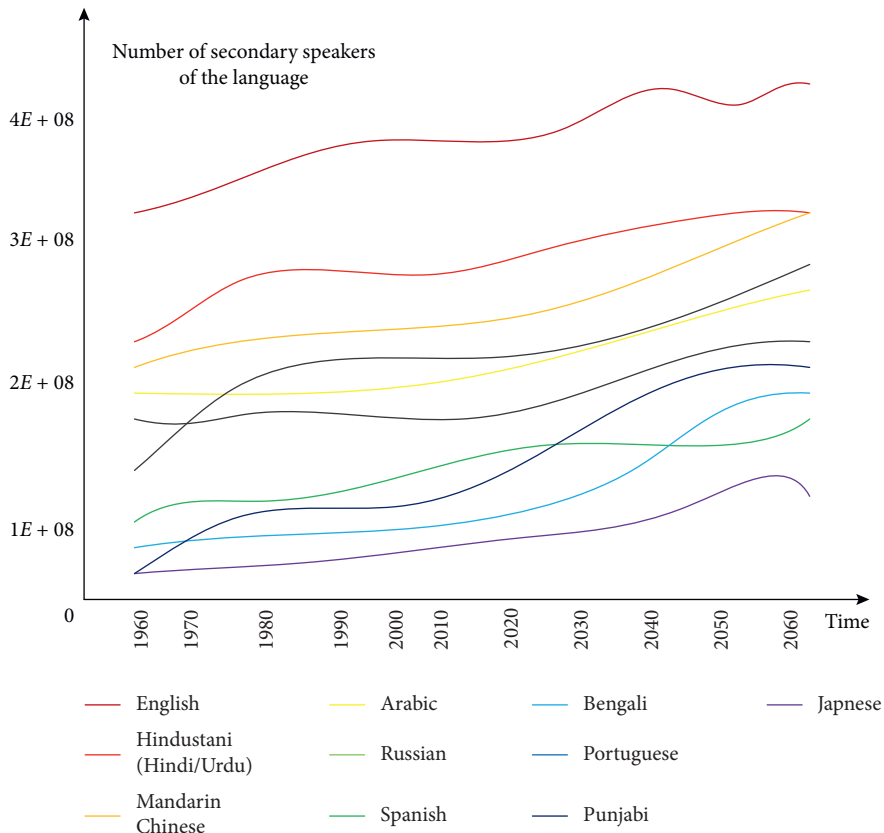


FIGURE 4: Number of secondary speakers from 1960 to 2060.

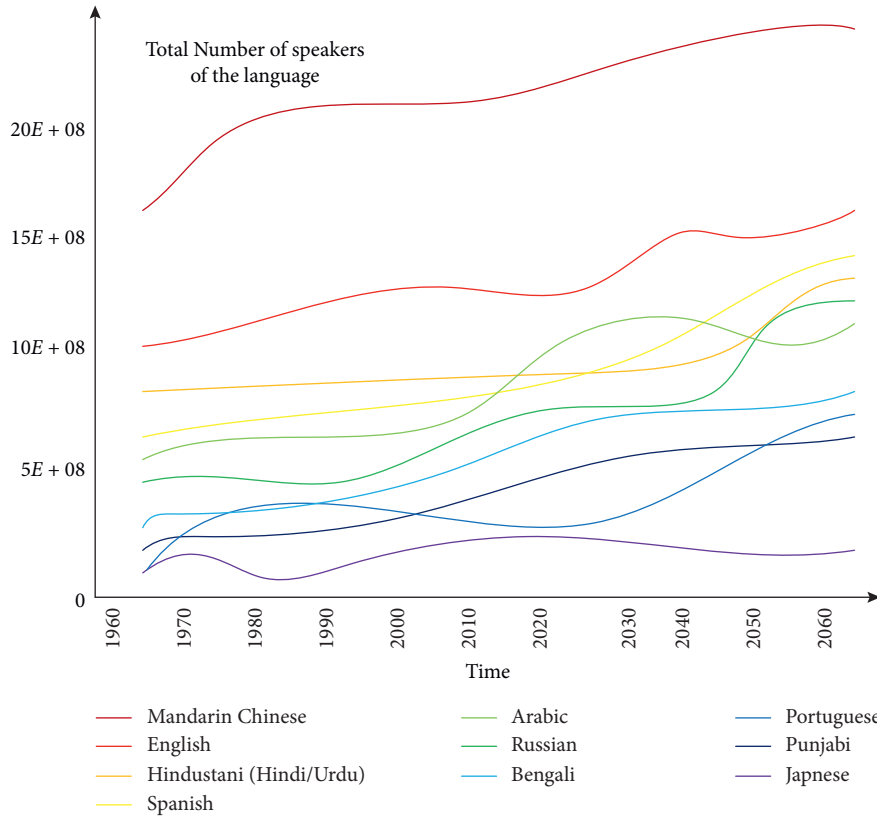


FIGURE 5: Total number of speakers from 1960 to 2060.

years and according to the established formula, the coefficient matrix remains unchanged and the time coefficient increases so that the ratio of coordinate position to trend coefficient will increase significantly. Therefore, the total number of people who speak the language will increase dramatically.

To describe the trend of language development, we depict the following diagram, as shown in Figures 3–5.

Similarly, we can substitute the remaining five languages in Figure 5 to predict the distribution in the same way. By predicting the top ten languages, we think there will be languages in other ten languages of the world that are replaced by other languages. The reason is as follows.

Although the world’s total population will eventually stabilize, it can be seen from Figure 5 that the number of speakers with languages continues to increase, great instability will occur in both countries and regions, and the number of second languages will continue to increase. And there is a tendency to approach the number of native speakers. So, anyone of the top ten languages may be replaced.

### 5. Sensitivity Analysis

Based on the internal fitting degree of the model, the significance of the parameters in the model, and the internal and external predictive ability of the model, we established a predictive model with good fit, strong stability, and good predictive ability.

In the first part, we establish a multifactorial global language localization prediction model. Due to the large number of factors, it is not conducive to the verification of the sensitivity of the variable change. Therefore, we use the data internally prediction test, using the data of 1960–2007 global languages as the database. We reanalyzed the languages’ comprehensive ability coefficient  $\eta$  for that period; putting it into the model we set up, we predicted the situation for the next decade, compared it with the actual total language population for 2007–2017, and plotted it in the same figure to compare, as shown in Figure 6.

The error is

$$m = \frac{\sum_{i=0}^n |x'_i/x_i - 1|}{n}. \tag{13}$$

Take the same distance 100000 number set to  $n$ , as shown in Table 2.

As can be seen from the above table, the error  $m$  is less than 0.5%, so the model is more closely related to the language trend forecast in the short term.

#### 5.1. The Advantages of the Model

- (1) Based on the correct analysis of the influencing factors, we use the correct data processing method to solve the multifactor optimization problem well and establish a predictive model with high fitness, good stability, and good predictive ability.

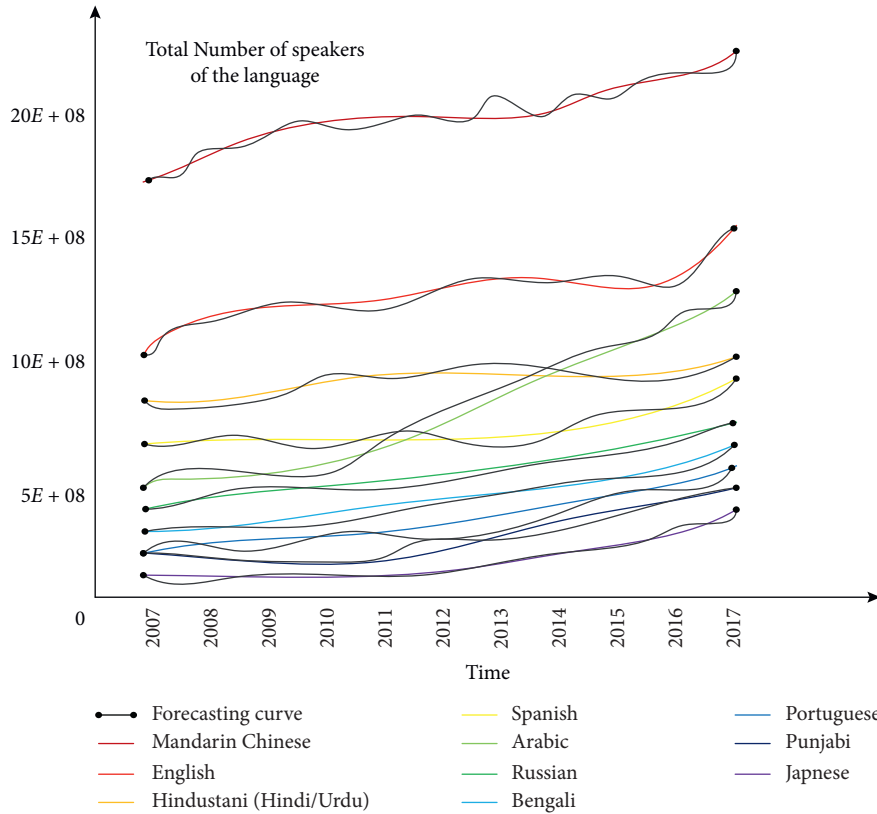


FIGURE 6: Analysis chart of actual and forecast results.

TABLE 2: Error analysis table.

Languages	Mandarin Chinese	English	Hindustani	Spanish	Arabic	Russian	Bengali	Portuguese	Punjabi	Japanese
$M$ (%)	0.48	0.35	0.32	0.24	0.18	0.15	0.25	0.10	0.08	0.06

(2) We conduct a sensitivity analysis based on the internal fitting degree of the model, the significance of the parameters in the model, and the internal and external predictive ability of the model. The internal parameters of the prediction test and the fitting test obtained Table 2, and the internal and external parameters of the test of significance test is stable, so our team model has good applicability.

5.2. The Disadvantages of the Model

- (1) Introducing too many variables in the process of model establishment can easily lead to “dimensionality disaster,” which is not conducive to programming
- (2) This model has a parameter hypothesis, but it cannot avoid the model deviation caused by some actual factors, which may not be completely in line with the actual situation

Hopefully, the above reference suggestions will be very helpful to the readers.

**Data Availability**

The data used to support the findings of the study are available from the corresponding author upon request.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Acknowledgments**

This work was supported by Key Projects of Support Program for Outstanding Young Talents in Anhui Province Colleges and Universities (no. gxyq 2018119) and Key Project Funds from Anhui Education Ministry (no. 2019 rcsfjd089).

## References

- [1] W. G. Mitchener, "A stochastic model of language change through social structure and prediction-driven instability," *SIAM Journal on Applied Mathematics*, vol. 77, no. 6, pp. 2272–2293, 2017.
- [2] B. J. Seggewiss, T. Straatmann, K. Hattrup, and K. Mueller, "Testing interactive effects of commitment and perceived change advocacy on change readiness: investigating the social dynamics of organizational change," *Journal of Change Management*, vol. 19, no. 2, 2019.
- [3] D. Sayers, "The mediated innovation model: a framework for researching media influence in language change," *Journal of Sociolinguistics*, vol. 18, no. 2, pp. 185–212, 2014.
- [4] G. Baxter and W. Croft, "Modeling language change across the lifespan: individual trajectories in community change," *Language Variation and Change*, vol. 28, no. 2, pp. 129–173, 2016.
- [5] S. J. Greenhill, C.-H. Wu, X. Hua, M. Dunn, S. C. Levinson, and R. D. Gray, "Evolutionary dynamics of language systems," *Proceedings of the National Academy of Sciences*, vol. 114, no. 42, pp. 8822–8829, 2017.
- [6] J. A. Jódar-Sánchez, "The handbook of historical sociolinguistics," *Language in Society*, vol. 44, no. 5, 2015.
- [7] K. Stadler, R. A. Blythe, K. Smith, and S. Kirby, "Momentum in language change," *Language Dynamics and Change*, vol. 6, no. 2, pp. 171–198, 2016.
- [8] R. Borges, "The role of extralinguistic factors in linguistic variation and contact-induced language change among Suriname's Coppename Kwinti and Ndyuka Maroons," *Acta Linguistica Hafniensia*, vol. 45, no. 2, pp. 228–246, 2014.
- [9] A. D. M. Smith, "Models of language evolution and change," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 5, no. 3, pp. 281–293, 2014.
- [10] F. Fu and M. A. Nowak, "Global migration can lead to stronger spatial selection than local migration," *Journal of Statistical Physics*, vol. 151, no. 3-4, pp. 637–653, 2013.
- [11] J.-B. Liu, J. Zhao, H. He, and Z. Shao, "Valency-based topological descriptors and structural property of the generalized sierpiński networks," *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.



## Research Article

# Multiobjective Optimization Design of Toll Plaza

Kai Wang<sup>1</sup>, Peng Wang,<sup>2</sup> Xin Chen,<sup>2</sup> and Lu-Ting Zhao<sup>2</sup>

<sup>1</sup>School of Science, Anhui Agricultural University, Hefei 230036, China

<sup>2</sup>School of Engineering, Anhui Agricultural University, Hefei 230036, China

Correspondence should be addressed to Kai Wang; [kaywang@163.com](mailto:kaywang@163.com)

Received 7 August 2020; Accepted 30 September 2020; Published 27 October 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Kai Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper mainly studies the optimization design of toll plaza, including the determination of the number of tollbooths and the design of the shape and size of the toll plaza. The optimization objectives are the construction cost of toll plaza, the throughput, and the accident rate. Through the analysis on the structure of toll plaza and the traffic flow, we determine the impact factors for optimization targets and select the number of tollbooths, the length of entrance and exit queue area, and the ingress and egress angles as decision variables and then build the function relationship between construction cost, accident rate, throughput, and decision variables. Then based on those functions to build the mathematical programming model, so as to get the optimal design plan.

## 1. Introduction

The design of the toll plaza of highways may directly affect the passage of vehicles and the accident occurrence possibility and at the same time determine the construction investment of the toll plaza. The design with insufficient consideration may not only lead to large construction cost but also result in such problems as traffic jam, high accident rate, and environment pollution. Especially nowadays with rapid increase of holding quantity of vehicles and the innovation of driving technology, under the condition that it cannot change the original highway lanes, how to design the toll plaza reasonably and guarantee the smooth passage of vehicles are the problems to be solved.

The progress of science and technology has an impact on charging methods. On the one hand, there was only one toll collection type in the original design, which was human-staffed manner. However, there are three kinds of toll collection manners existing nowadays: human-staffed, automated, and electronic toll collection. Hence, the design of tollbooths shall be changed. On the other hand, the number of autopilot cars is increasing. Usually, the autonomous vehicles may adopt the electronic toll collection manner. This trend will necessarily affect the design of toll plaza.

In order to guarantee the throughput in entrance area, in the original design, it basically adopted the method to set up more tollbooths. But, too many tollbooths result in the increase of the land area of the toll plaza which increases the construction cost. On the other hand, it increased the lane merging times for vehicles in the exit area, which is easier to cause the occurrence of accident.

So, there is a dilemma in the design of toll plaza. On the one hand, in order to improve the throughput, toll stations need to be increased. On the other hand, the increase of toll station will increase the construction cost and may increase the accident rate.

In this paper, the throughput of vehicles, the construction cost, and the occurrence possibility of accident are selected as the indicators to be optimized. Through the analysis on the structure of toll plaza and the traffic flow, the number of tollbooths, the length of entrance and exit queue area, and the ingress and egress angle are determined as decision variables. Based on the function relationship between construction cost, throughput, accident rate, and decision variables, we build the mathematical programming model and achieve the minimum value of the construction cost and accident rate under the condition of ensuring the throughput of the toll plaza, and the optimal design scheme of toll plaza is obtained.

## 2. Related Work

The research on toll plaza mainly includes four aspects: driver behavior, security of toll plaza, service level of toll plaza, and design of toll plaza.

*2.1. Research on Driver Behavior.* The most common driver behavior is the shortest path selection, which can be applied by using spanning tree in graph theory. Some literatures use polynomial logit model [1], mixed model of random effects [2], and XCS learning algorithm [3] to study the driver's lane selection [4] and the influence of road signs [5]. It is found that both dynamic information signs and road signs can reduce unsafe driving behaviors.

*2.2. Research on the Security of Toll Plaza.* By use of the time-varying mixed model and logit model, Lu et al. [6, 7] found that the crash risk of vehicles in the diversion area of toll plaza is the highest. Valdés et al. [8] and Abuzwidah and Abdel-Aty [9, 10] compared the security of different types of toll plaza. Jehad et al. [11] used VIS-SIM simulation model to design toll plaza to improve security.

*2.3. Research on Service Level of Toll Plaza.* Abdul Majid et al. [12], Mahdi et al. [13, 14], and Ozmen-Ertekin et al. [15] used the Markov chain system to study the factors that cause congestion and service level decline in Toll Plaza and found that the main factors include waiting queue length, delay time, and heavy vehicles. Jack and Haitham [16] used analytic hierarchy process to evaluate the service level of eight different toll plazas. Lin et al. [17] divided the service level of toll plaza into six levels according to the queue length and passing time of toll plaza.

*2.4. Research on the Design of Toll Plaza.* There are two aspects in the design of toll plaza. One is based on network topology index calculation and uses information technology to design toll station management system, including Internet of things [18], XBee wireless transceiver [19], electronic toll collection system [20], and vehicular ad hoc network [21] to observe and adjust the traffic flow. The other one is to target the throughput of toll plaza [22, 23] or cost and security [24], used discrete model [25], nonlinear integer programming [26], cellular automata [27, 28], ALINEA [29, 30], and multiple linear regression [31] to optimize the lane width, lane configuration, and queue length of toll plaza [32].

## 3. Model

The structure of toll plaza is shown in Figure 1.

Obviously, in one side of the isolation zone, vehicles may only move towards one direction. Therefore, it is enough to design the model of toll plaza in one side of the isolation zone; as to the other side, what needs we do is only change the parameters to produce a similar design (both sides are not necessarily symmetrical).

Let us select the lower part to analyze. Vehicles move from the left to the right, so the left part is the entrance of the toll plaza and the right part is the exit. The toll plaza is divided into five parts from the left to the right: fan-out area, entrance queue area, tollbooths area, exit area, and fan-in area. That is to say, we need to design those five parts to maximize the throughput, minimize the construction cost, and the accident rate.

Within a certain period of time, a certain number of different vehicles move into the toll plaza from the left, arrive at fan-out area, queue up to pay in entrance queue area based their own types, pass the tollbooth after paying toll, and then drive through the exit queue area and fan-in area to enter the lane.

During that driving course, there are two constraint nodes on vehicles: tollbooths and exit lanes. First of all, from the direction of entrance, if the tollbooths are too few, the vehicles will wait in the entrance queue area, increasing the queue length and the waiting time; hence, it needs a larger queue area, increasing the land acquisition expense of toll plaza. Therefore, the increase of tollbooths may shorten the queue length and reduce the land area of toll plaza. However, if there are too many tollbooths, on the one hand, the construction cost of tollbooth will be increased, and the width of toll plaza will be increased too due to the increase of tollbooth; hence, the land acquisition expense will be increased. On the other hand, too many tollbooths may increase the confluence extent of vehicles in fan-out area and hence increases the accident rate. At the same time, from the direction of entrance, the increase of tollbooth may lead to the increase of traffic flow in entrance area within the unit time, while the number of egress lanes is usually less than that of tollbooth, which may result in the confluence of more vehicles. As a result, it may cause a traffic jam, increase the waiting time and queue length, and reduce the throughput or increase the land acquisition expense due to the addition of the size of toll plaza. On the other hand, there will be more vehicles joining in the fan-in area, which may improve the accident rate.

*3.1. Construction Cost.* Construction cost mainly includes land acquisition expense and tollbooth construction cost. Where there are less tollbooths, there will be a longer queue in entrance area, hence reducing the throughput and increasing the land acquisition expense of entrance area. While if there are more tollbooths, on the one hand, the construction cost of tollbooth is directly proportional to the number of tollbooths, and on the other hand, it may reduce the throughput and increase the land acquisition expense of tollbooth, and at the same time, it may result in a longer queue in exit area, hence increasing the land acquisition expense of exit area. Therefore, whether there are too many or too few tollbooths, the construction cost of the toll plaza will be increased.

The construction cost is determined by the number of tollbooths, the length of queue area, and its angle, that is, the function:

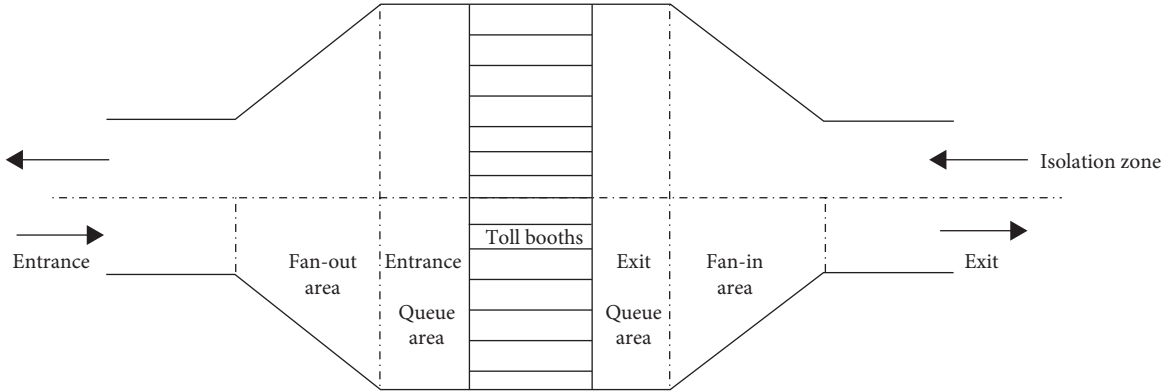


FIGURE 1: The structure of the toll plaza.

$$\text{Cost}_B = f(B_1, B_2, B_3, L_{EN}, L_{EX}, \theta_{EN}, \theta_{EX}), \quad (1)$$

where  $\text{Cost}_B$  is the construction cost of toll plaza and  $B_i$  ( $i = 1, 2, 3$ ) is the number of three kinds of tollbooth. As shown in Figure 2,  $L_{EN}$  and  $L_{EX}$  are the length of queue area in fan-out area and fan-in area and  $\theta_{EN}$  and  $\theta_{EX}$  are the angles of fan-out area and fan-in area.

Assuming the unit price for land acquisition is  $c$  US\$/sq.m, the land acquisition expense is

$$c[(L_{EN} + L_{EX}) \cdot H + (H + W_L)(H - W_L)(\cot \theta_{EN} + \cot \theta_{EX})H_{TB}W_{TB}B], \quad (2)$$

where  $H_{TB}$  is the length of the tollbooth,  $W_{TB}$  is the width of the tollbooth,  $B$  is the total number of the tollbooth,  $H$  is the total width of the toll plaza, and  $W_L$  is the width of the single direction of highways.

Assume the construction cost of three kinds of tollbooths is  $c_{TB}^i$  ( $i = 1, 2, 3$ ), then the construction cost of the toll booth is

$$\sum_{i=1}^3 (c_{TB}^i \cdot B_i). \quad (3)$$

Hence, the construction cost function is

$$\begin{aligned} \text{Cost}_B = & c[(L_{EN} + L_{EX}) \cdot H + (H + W_L)(H - W_L)(\cot \theta_{EN} \\ & + \cot \theta_{EX}) + H_{TB}W_{TB}B] + \sum_{i=1}^3 (c_{TB}^i \cdot B_i). \end{aligned} \quad (4)$$

**3.2. Throughput of Toll Plaza.** Throughput refers to the number of vehicles passing through the toll plaza within certain period of time, and it is determined by the passing speed, time of vehicles, density of traffic flow, etc. It is evident that the longer vehicles waiting in toll plaza, the lower the throughput. Because we also need to consider the construction cost of the toll plaza, when the waiting time is directly proportional to the queue length, the throughput may be determined by the queue length, and the longer the queue length is, the longer vehicles will wait and the lower the throughput will be. Vehicles mainly queue at two points. The first is in the entrance

area before the tollbooth, and the more the tollbooths are, the shorter each queue and the larger the throughput will be. The second point is in the exit area before vehicles entering into lanes. The more the tollbooths are, the more vehicles entered into exit area, the longer the queue and the waiting time will be, hence the smaller the throughput will be. If the exit queue area cannot accommodate the queue, it will cause the traffic jam in entrance area, so as to reduce the throughput. Therefore, too few tollbooths may result in the increase of queue length in entrance area, while too many tollbooths may lead to the increase of queue length in exit area, both of which will reduce the throughput.

When vehicles enter into the fan-out area of the toll plaza, they shall pay the toll to pass the tollbooth. The more the tollbooths, the higher the total service rate. The shorter the queue length, the larger the throughput. On the contrary, the longer the queue length, the smaller the throughput. Thus there exists the following function relationship between the number of tollbooths and the queue length:

$$D_{EN} = g_{EN}(\lambda_{EN}, B_1, B_2, B_3), \quad (5)$$

where  $D_{EN}$  is the number of vehicles queuing in fan-out area and  $\lambda_{EN}$  is the average arriving rate of vehicles moving into fan-out area.

While if the reserved length of entrance queue area is shorter than the actual queue length, it will cause the vehicles followed have to stop on the highways with fewer lanes, which will lead to more traffic jam. Therefore, the length of entrance queue area shall meet:

$$L_{EN} \geq D_{EN} \cdot h = g_{EN}(\lambda_{EN}, B_1, B_2, B_3) \cdot h, \quad (6)$$

where  $L_{EN}$  is the length of queue area in fan-out area and  $\lambda_{EN}$  is the average arriving rate of vehicles moving into fan-out area. Assume the arriving number of vehicles within certain period of time is  $N$ , then the average arriving rate is  $\lambda_{EN}$ . Assume the ratio of three kinds of vehicles are  $k_1, k_2$ , and  $k_3$ , the service rate of three kinds of toll booths are  $\mu_1, \mu_2$ , and  $\mu_3$ , respectively. Then, based on  $M/M/B$  models, the function relationship between the number of

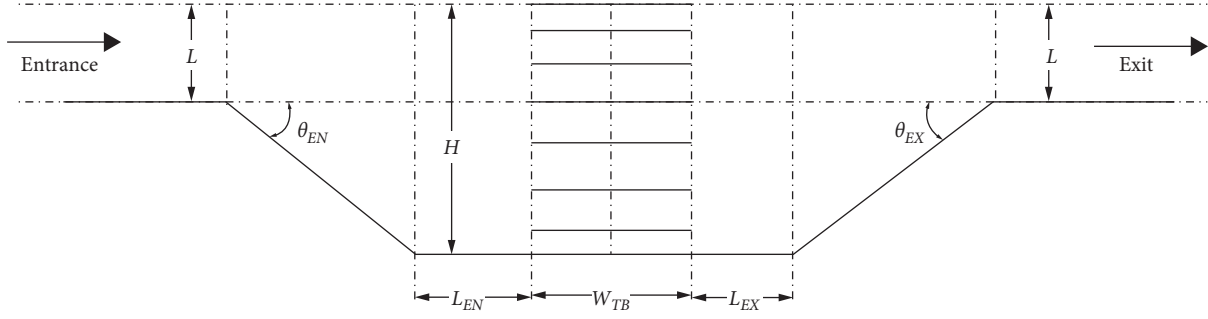


FIGURE 2: Figure of symbol.

each kind of vehicles queuing up and the number of tollbooths is

$$D_{EN}^i = \frac{(B_i \rho_i)^{B_i} \rho_i}{B_i! (1 - \rho_i)^2} P_0^i + B_i \rho_i, \quad (7)$$

where

$$P_0^i = [\sum_{k=0}^{B_i-1} (1/k!) (B_i \rho_i)^k + (1/B_i!) \cdot (1/(1 - \rho_i))] \cdot (B_i \rho_i)^{B_i}$$

$$\text{and } \rho_i = (\lambda_{EN}^i / \mu_i), \lambda_{EN}^i = k_i \cdot \lambda_{EN}.$$

Because the  $M/M/B$  model is a single queue, the number of actual queue of vehicles shall be divided by the number of the toll booth, so the actual number of vehicles queuing up is

$$AD_{EN}^i = \left[ \frac{D_{EN}^i}{B_i} + 1 \right]. \quad (8)$$

In order to save up the land area, let us take the length of queue area as the actual length of vehicles queuing up, then

$$L_{EN} = h \cdot \max_{1 \leq i \leq 3} \{AD_{EN}^i\}. \quad (9)$$

In the fan-in area, the incoming vehicles have paid the toll. If there are more tollbooths, then there will be more vehicles moving into the fan-in area in unit time. When driving into the lanes, the large traffic flow will form a "bottleneck," cause traffic jam, increase the queue length, and reduce the throughput. Consequently, there exists a function relationship between the number of tollbooths and the queue length:

$$D_{EX} = g_{EX}(\lambda_{EX}, B_1, B_2, B_3), \quad (10)$$

where  $D_{EX}$  is the number of vehicles queuing in fan-in area, and  $\lambda_{EX}$  is the average arriving rate of vehicles moving into fan-in area.

Obviously, if the reserved length of exit queue area is shorter than the actual queue length, it will cause the vehicles following cannot pass through the tollbooth, which will reduce the throughput further. Therefore, the length of exit queue area shall meet:

$$L_{EX} \geq D_{EX} \cdot h = g_{EX}(\lambda_{EX}, B_1, B_2, B_3) \cdot h, \quad (11)$$

where  $L_{EX}$  is the length of queue area in fan-in area and  $h$  is the length of vehicles.

When vehicles leave the toll plaza, there is no need to distinguish their types. Hence, under the condition that the egress lanes are definite, the queue length in fan-in area is determined by the total incoming traffic flow. While the vehicles in fan-in area are those having paid the toll, so the traffic flow is related to the number of tollbooths.

If the arriving number of certain type of vehicles before the entrance  $k_i X$  is not more than the total service rate of this kind of tollbooth, which means the work of the tollbooth is not at full load, then the number of vehicles entering into fan-in area is the number of vehicles at the entrance  $k_i X$ .

If the arriving number of certain type of vehicles before the entrance  $k_i X$  is more than the total service rate of this kind of tollbooth, which means the work of the tollbooth is at full load, then the number of vehicles entering into fan-in area is  $c_i \mu_i$ .

Hence, the probability distribution sequence of vehicles arriving at the fan-in area is as follows:

$Y$	$X$	$(k_1 + k_2) X_3 + c_3 \mu_3$	$\dots$	$c_1 \mu_1 + c_2 \mu_2 + c_3 \mu_3$
$P$	$P_1^0 P_2^0 P_3^0$	$P_1^0 P_2^0 P_3^1$	$\dots$	$P_1^1 P_2^1 P_3^1$

(12)

where  $P_i^0 = P(k_i X < c_i \mu_i)$  and  $P_i^1 = 1 - P_i^0$  ( $i = 1, 2, 3$ ).

Thus, the average arriving rate of vehicles in fan-out area is

$$\lambda_{EX} = E(Y) = P_1^0 P_2^0 P_3^0 \cdot X + \dots + P_1^1 P_2^1 P_3^1 \cdot (c_1 \mu_1 + c_2 \mu_2 + c_3 \mu_3). \quad (13)$$

We have known the number of lanes is  $L$ , whose total service rate is the arriving rate of vehicles to the entrance  $\lambda$ , based on  $M/M/L$  model, there will be

$$D_{EX} = \frac{(L\rho)^L \rho}{L! (1 - \rho)^2} P_0 + L\rho, \quad (14)$$

where  $P_0 = [\sum_{k=0}^{L-1} (1/k!) (L\rho)^k + (1/L!) \cdot (1/(1 - \rho))] \cdot (L\rho)^L$  and  $\rho = (\lambda_{EX}/\lambda)$ .

In the same way with fan-out area, the queue length in the fan-in area is

$$L_{EX} = h \cdot \left[ \frac{D_{EX}}{B + 1} \right]. \quad (15)$$

**3.3. Accident Rate.** In the toll plaza, the occurrence of accident is mainly due to the confluence of vehicles. As shown in the Figure 1, vehicles confluence mainly happens

in the fan-out area after the entrance and the fan-in area before the exit. Apparently, if there are more tollbooths, the confluence extent of vehicles in fan-out area and fan-in area will be higher, and then the possibility of accident occurrence will be larger. And, if the ingress and egress angle in entrance fan-out area and exit fan-in area is larger, then the buffer extent will be smaller, and the possibility of accident occurrence will be larger. So, the increase of tollbooth, and the enlarging of ingress and egress angle may increase the possibility of accident occurrence.

So, the accident rate in fan-out area and fan-in area may be described by the following function:

$$\begin{aligned} I_{EN} &= \varphi_{EN}(B-L, \theta_{EN}), \\ I_{EX} &= \varphi_{EX}(B-L, \theta_{EX}), \end{aligned} \quad (16)$$

where  $B = \sum_{i=1}^3 B_i$ .

Assume one toll booth corresponds to one lane, obviously, when vehicles move into corresponding type of tollbooth, they only need to move straightly. So, the upper limit of the times to change the lane is  $B-L$ .

Then, the upper limit of merging times in unit time and unit area in fan-out area and fan-in area are

$$\begin{aligned} \text{entrance: } I_{EN} &= \frac{\lambda_{EN}(B-L)}{S_{EN}T_{EN}}, \\ \text{exit: } I_{EX} &= \frac{\lambda_{EX}(B-L)}{S_{EX}T_{EX}}, \end{aligned} \quad (17)$$

where  $T_{EN}$  and  $T_{EX}$  are the waiting time for vehicles in fan-out area and fan-in area, respectively, and they meet:

$$\begin{aligned} T_{EN} &= \frac{D_{EN}}{\lambda_{EN}}, \\ T_{EX} &= \frac{D_{EX}}{\lambda_{EX}}. \end{aligned} \quad (18)$$

Thus,

$$\begin{aligned} I_{EN} &= \frac{\lambda_{EN}^2(B-L)}{S_{EN} \cdot D_{EN}}, \\ I_{EX} &= \frac{\lambda_{EX}^2(B-L)}{S_{EX} \cdot D_{EX}}. \end{aligned} \quad (19)$$

Assume the probability of accident occurrence with each lane merging is  $P_I$  and the loss for each accident is  $C_I$ . The useful life of the toll plaza is  $T$  (hours), then the possible loss due to the traffic accident happened in fan-out area and fan-in area is

$$\begin{aligned} \text{Lost}_{EN} &= I_{EN}P_IC_IT = \frac{\lambda_{EN}^2(B-L)}{S_{EN}(\theta_{EN}) \cdot D_{EN}}P_IC_IT, \\ \text{Lost}_{EX} &= I_{EX}P_IC_IT = \frac{\lambda_{EX}^2(B-L)}{S_{EX}(\theta_{EX}) \cdot D_{EX}}P_IC_IT, \end{aligned} \quad (20)$$

respectively.

When the average arriving rate of vehicles  $\lambda$ , the number of lanes  $L$ , and the number of tollbooth  $B$  are definite, apparently the  $\text{Lost}_{EN}$  and  $\text{Lost}_{EX}$  are determined only by  $\theta_{EN}$  and  $\theta_{EX}$  and hence may minimize the loss of traffic accident through the decision of  $\theta_{EN}$  and  $\theta_{EX}$ .

To sum up, the optimization design model of the toll plaza is

$$\begin{aligned} \text{Min } Z &= c[(L_{EN} + L_{EX}) \cdot H + (H + W_L)(H - W_L)(\cot \theta_{EN} + \cot \theta_{EX}) + H_{TB}W_{TB}B] \\ &+ \sum_{i=1}^3 (c_{TB}^i \cdot B_i) + \frac{\lambda_{EN}^2(B-L)}{S_{EN}(\theta_{EN}) \cdot D_{EN}}P_IC_IT + \frac{\lambda_{EX}^2(B-L)}{S_{EX}(\theta_{EX}) \cdot D_{EX}}P_IC_IT, \\ \text{s.t. } &\begin{cases} L_{EN} = h \cdot \text{Max}_{1 \leq i \leq 3} \{AD_{EN}^i\}, \\ L_{EX} = h \cdot \left[ \frac{D_{EX}}{B} + 1 \right], \\ B_i, L_{EN}, L_{EX} \geq 0, \\ 0 \leq \theta_{EN}, \theta_{EX} \leq \frac{\pi}{2}. \end{cases} \end{aligned} \quad (21)$$

#### 4. Solution

Because  $B_1, B_2$ , and  $B_3$  are only the integral values,  $B_1, B_2$ , and  $B_3$  shall be calculated within  $[1, 20]$  based on the step size of 1, which is to calculate the length of queue area in entrance area  $L_{EN}$  under different mix of  $(B_1, B_2, B_3)$  according to the formula (9).

Similarly, calculating within  $[1, 20]$  based on the step size of 1 to  $B_1, B_2$ , and  $B_3$ , we may get the traffic flow in entrance

area  $\lambda_{EN}$ , the number of vehicles queuing up, and the length of queue area in exit area  $L_{EX}$ .

Based on the results  $B_1, B_2, B_3, L_{EX}, D_{EX}$  and the formula (4), we may get the construction cost function of the toll plaza  $\text{Cost}_B(\theta_{EN}, \theta_{EX})$ . Based on the formula (21), we may get the loss function for accident in entrance and exit area  $\text{Lost}_{EN}(\theta_{EN})$  and  $\text{Lost}_{EX}(\theta_{EX})$ .

So, as to the objective function,

$$\begin{aligned}
 Z &= \text{Cost}_B(\theta_{EN} + \theta_{EX}) + \text{Lost}_{EN}(\theta_{EN}) + \text{Lost}_{EX}(\theta_{EX}) \\
 &= c \left( L_{EN} \cdot H + \frac{(W_L + H)(H - W_L)}{2 \tan \theta_{EN}} \right) + c \left( L_{EX} \cdot H + \frac{(W_L + H)(H - W_L)}{2 \tan \theta_{EX}} \right) + c \cdot S_{TB} + c_{TB} \cdot B \\
 &\quad + \frac{\lambda_{EN}^2 (B - L) \cdot P_I C_I T}{(L_{EN} \cdot H + ((W_L + H)(H - W_L)) / (2 \tan \theta_{EN})) \cdot D_{EN}} + \frac{\lambda_{EX}^2 (B - L) \cdot P_I C_I T}{(L_{EX} \cdot H + ((W_L + H)(H - W_L)) / (2 \tan \theta_{EX})) \cdot D_{EX}}, \\
 F_{EN} &= \frac{\lambda_{EN}^2 (B - L) \cdot P_I C_I T}{D_{EN}}, \\
 F_{EX} &= \frac{\lambda_{EX}^2 (B - L) \cdot P_I C_I T}{D_{EX}}, \\
 G &= \frac{(W_L + H)(H - W_L)}{2}, \\
 \text{Letting } \begin{cases} \frac{\partial Z}{\partial \theta_{EN}} = -G \csc^2 \theta_{EN} \left( c - \frac{F_{EN}}{(L_{EN} H + G \cot \theta_{EN})^2} \right) = 0, \\ \frac{\partial Z}{\partial \theta_{EX}} = -G \csc^2 \theta_{EX} \left( c - \frac{F_{EX}}{(L_{EX} H + G \cot \theta_{EX})^2} \right) = 0, \end{cases} \\
 \text{we have } \begin{cases} \tan \theta_{EN} = \frac{G}{\sqrt{(F_{EN}/c) - L_{EN} \cdot H}}, \\ \tan \theta_{EX} = \frac{G}{\sqrt{(F_{EX}/c) - L_{EX} \cdot H}}, \end{cases}
 \end{aligned} \tag{22}$$

**4.1. Simulation.** According to the highway toll station and toll plaza design specifications and survey results, the following assumptions are made for the initial data, as shown in Table 1.

According to the above algorithm, we can get 8000 kinds of results, some of which are listed in Table 2.

Table 2 shows, along with the increase of the number of tollbooths, the length of exit queue area increases too, leading to the increase of the land acquisition expense. Therefore, after comprehensive comparison, the result of  $(B_1, B_2, B_3) = (3, 3, 1)$  not only may guarantee the preferable throughput but also minimize the total expense.

Thus, when the arriving rate of vehicles is 1800 vehicles/hour, the type ratio of vehicles is 1 : 1 : 1 and the optimized design plan is as follows: the number of human-staffed, exact-change, and electronic tollbooths are 3, 3, and 1, respectively, and 7 in total; the length of entrance queue area and that of the exit queue area are both 5 meters; the angle of ingress is  $14.43^\circ$  and that of the egress is  $6.53^\circ$ . In this design plan, the total expense is minimum, which is US\$ 29.73 million.

**4.2. Sensitivity Analysis.** Because the traffic flow is a random variable, under the condition that the ratio of different kinds



TABLE 1: Initial data.

Toll type	Conventional toll booths	Exact-change toll booths	Electronic toll collection booths
Service rate	350 vehicles/h	550 vehicles/h	1200 vehicles/h
Arriving rate	600 vehicles/h	600 vehicles/h	600 vehicles/h
Road width		3.5 m	
Width of tollbooth		5.8 m	
Length of tollbooth		4.2 m	
Length of vehicles		5.0 m	
The land acquisition expense		2200\$ (m <sup>2</sup> )	
Useful life		20 years	
Construction expense	58400\$ (per)	71700\$ (per)	83300\$ (per)

TABLE 2: Part of design plan.

No.	( $B_1, B_2, B_3$ )	$L_{EN}$ (m)	$L_{EX}$ (m)	$\theta_{EN}$	$\theta_{EX}$	Cost <sub>B</sub> (million)	Lost <sub>EN</sub> + Lost <sub>EX</sub> (million)
1	(3, 2, 1)	5	5	12.00°	6.32°	12.88	17.39
2	(3, 3, 1)	5	10	14.43°	6.53°	15.34	14.39
3	(4, 2, 1)	5	5	14.43°	6.90°	14.88	21.45
4	(4, 3, 1)	5	15	17.03°	7.10°	17.67	13.70
5	(5, 2, 1)	5	10	17.03°	7.46°	17.15	16.42

TABLE 3: The sensitivity analysis on traffic flow.

$\lambda_{EN}$	800	900	1000	2000	2500	2600
( $B_1, B_2, B_3$ )	(3, 2, 1)	(3, 3, 1)	(3, 3, 1)	(3, 3, 1)	(3, 3, 1)	(4, 3, 1)

TABLE 4: The sensitivity analysis on vehicles ratio.

Vehicles ratio	( $B_1, B_2, B_3$ )	$L_{EN}$ (m)	$L_{EX}$ (m)	$\theta_{EN}$	$\theta_{EX}$	Cost <sub>B</sub> (million)	Lost <sub>EN</sub> + Lost <sub>EX</sub> (million)
1:1:2	(3, 2, 1)	10	5	18.9°	6.32°	8.95	15.61
1:1:2	(3, 3, 1)	10	10	22.6°	6.53°	10.76	12.35
1:1:2	(4, 3, 1)	10	15	26.4°	6.90°	12.57	11.40
1:1:10	(1, 1, 2)	10	5	13.43°	5.01°	5.22	9.84
1:1:10	(2, 2, 2)	10	15	18.92°	11.90°	9.73	14.03

of vehicles is 1 : 1 : 1, changing the average arriving rate of vehicles and observing its effect on the optimized design plan may produce Table 3:

Table 3 shows, when the average arriving rate of vehicles within [900, 2500], the setup of (3, 3, 1) is the optimized plan, which proves this plan has a wide permissible range of traffic flow. Even the traffic flow is small, as long as it is not less than 900 vehicles/hour, there will no tollbooth in idle; as to the large traffic flow, as long as it is not more than 2500 vehicles/hour, there will no traffic jam in the toll plaza to affect the throughput.

Along with the progress of technology, the autonomous vehicles increase, and inevitably, the autonomous vehicles must use electronic tollbooth. Therefore, under the condition that the arriving rate of vehicles still is 1800 vehicles/hour, changing the ratio of different types of vehicles may produce the results as shown in Table 4.

Table 4 shows, with the arriving rate of vehicles unchanged, when the ratio of using electronic toll-paying manner increases to 1 : 1 : 2, if still sticking to set 1 electronic toll collection booth, there will be traffic jam at the entrance, and the throughput is reduce. This is because that the vehicles using electronic toll-paying manner usually

have high speed, if there is fewer tollbooth, the vehicles will absolutely queue up. At the same time, the table shows traffic jam will increase the length of entrance queue area, but the angle of entrance may be increased, because the electronic tollbooth is mainly set on the straight lane, most vehicles may keep their way in straight lane. As a result, increase the entrance angle may reduce the construction cost. In comprehensive, the optimized design for three types of tollbooths is still the set of (3, 3, 1). When the ratio of vehicles changes to be 1 : 1 : 10, it may add the electronic toll collection booths and reduce the human staffed and exact-change tollbooths, which may reduce the width of the toll plaza and the times for vehicles to change the lanes without enlarging the length of entrance and exit queue area; thus, it greatly reduce the construction cost and the accident loss, in which the optimized design of three kinds of tollbooths is (1, 1, 2).

### 5. Conclusion

By describing the function relationships between the construction cost, throughput, accident rate, and number of tollbooths, the length of entrance and exit queue area, the

ingress and egress angle, the mathematical programming model built by us has a large universality. Only inputting such initial data as the arriving rate of vehicles, the ratio of different type of vehicles, the service rate of tollbooths, and the land acquisition expense etc., you may get a complete design plan for the toll plaza, including the number of tollbooths and their ratio, the shape, the size of the plaza, etc.

Under certain initial conditions, the design plan of the toll plaza we get from the model also has strong robustness. It may guarantee the throughput of the plaza and avoid the idling of tollbooths in a large permissible range of traffic flow, which proves that the model we build is solid.

### Data Availability

The data are used to support the study are available within the article.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

### Acknowledgments

This work was supported by the Major Quality Engineering Projects in Anhui Province (2018jyxm1403).

### References

- [1] H. Parmar, P. Chakroborty, and D. Kundu, "Modelling automobile drivers' toll-lane choice behaviour at a toll plaza using mixed logit model," *Procedia-Social and Behavioral Sciences*, vol. 104, pp. 593–600, 2013.
- [2] M. Saad, M. Abdel-Aty, and J. Lee, "Analysis of driving behavior at expressway toll plazas," *Transportation Research Part F*, vol. 61, pp. 1–15, 2018.
- [3] B. Bartin, "Use of learning classifier systems in microscopic toll plaza simulation models," *IET Intelligent Transport Systems*, vol. 13, no. 5, pp. 860–869, 2019.
- [4] A. Dubedi, P. Chakroborty, D. Kundu, and K. Harikishan Reddy, "Modeling automobile drivers' toll-lane choice behavior at a toll plaza," *Journal of Transport Engineering*, vol. 138, no. 12, pp. 1350–1357, 2012.
- [5] S. Gaca, M. Kieć, S. Pogodzińska, and A. Wontorczyk, "Impact of toll plazas road marking and signs on drivers' behaviour," *MATEC Web of Conferences*, vol. 231, 8 pages, Article ID 02003, 2018.
- [6] X. Lu, J. He, M. Abdel-Aty, Y. Wu, and J. Yuan, "Time-varying analysis of traffic conflicts at the upstream approach of toll plaza," *Accident Analysis and Prevention*, vol. 141, no. 4, pp. 1–14, 2020.
- [7] L. Xing, J. He, M. Abdel-Aty, Q. Cai, Y. Li, and O. Zheng, "Examining traffic conflicts of up stream toll plaza area using vehicles' trajectory data," *Accident Analysis and Prevention*, vol. 125, no. 1, pp. 174–187, 2019.
- [8] D. Valdés, B. Colucci, M. Knodler et al., "Comparative analysis of toll plaza safety features in Puerto Rico and Massachusetts with a driving simulator," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2663, no. 1, pp. 1–11, 2017.
- [9] M. Abuzwidah and M. Abdel-Aty, "Safety assessment of the conversion of toll plazas to all-electronic toll collection system," *Accident Analysis & Prevention*, vol. 80, no. 3, pp. 153–161, 2015.
- [10] M. Abuzwidah and M. Abdel-Aty, "Crash risk analysis of different designs of toll plazas," *Safety Science*, vol. 107, no. 2, pp. 77–84, 2018.
- [11] A. E. Jehad, A. Ismail, M. N. Borhan, and S. Z. Ishak, "Modelling and optimizing of electronic toll collection (ETC) at Malaysian toll plazas using microsimulation models," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 2304–2308, 2018.
- [12] K. Abdul Majid, Z. Yusoff, and A. Aziz Jemain, "Queue analysis at toll plazas (inbound): a basic model for traffic systems towards a study of the effect on toll plazas with different arrival patterns," *International Conference on Information and Communication Technology*, pp. 5–9, 2016.
- [13] M. B. Mahdi and L. V. Leong, "Assessment of queue length and delay at toll plaza using microscopic traffic simulation," *Applied Mechanics and Materials*, vol. 802, pp. 387–392, 2015.
- [14] M. B. Mahdi, V. L. Lee, and M. S. Ahmad Farhan, "Use of microscopic traffic simulation software to determine heavy-vehicle influence on queue lengths at toll plazas," *Arabian Journal for Science and Engineering*, vol. 12, 2019.
- [15] D. Ozmen-Ertekin, K. Ozbay, S. Mudigonda, and A. M. Cochran, "Simple approach to estimating changes in toll plaza delays," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2047, no. 1, pp. 66–74, 2008.
- [16] K. Jack and M. Al-D. Haitham, "Proposed level-of-service methodology for toll plazas," *Transportation Research Record*, vol. 1802, pp. 86–96, 2002.
- [17] F. B. Lin and C.-W. Su, "Level-of-service analysis of toll plazas on freeway main lines," *Journal of Transportation Engineering*, vol. 120, no. 2, pp. 246–263, 1994.
- [18] H. Manjunath Singh, V. Velantina, G. Veda Bai, M. N. Varshini, and S. Padmavathi, "Reduction of traffic at toll plaza by automatic toll collection using rfid and gsm technology," *International Journal of Current Engineering and Scientific Research*, vol. 6, no. 6, pp. 187–192, 2019.
- [19] R. Deepika, R. A. Kalpana, S. Meera, and R. Sharmikha Sree, "Advanced toll plaza monitoring by utilization of WSN modules," *International Journal of Engineering & Technology*, vol. 7, no. 3, pp. 168–171, 2018.
- [20] E. Gordin, J. Klodzinski, and C. D. Santos, "Safety benefits from deployment of open road tolling for main-line toll plazas in Florida," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2229, no. 1, pp. 85–92, 2011.
- [21] S. Hussain, D. Wu, S. Memon, and N. K. Bux, "Vehicular ad hoc network (VANET) connectivity analysis of a highway toll plaza," *Data*, vol. 4, pp. 1–18, 2019.
- [22] R. Neuhold, F. Garolla, O. Sidla, and M. Fellendorf, "Predicting and optimizing traffic flow at toll plazas," *Transportation Research Procedia*, vol. 37, pp. 330–337, 2019.
- [23] C. Kim, D.-K. Kim, S.-Y. Kho, S. Kang, and K. Chung, "Dynamically determining the toll plaza capacity by monitoring approaching traffic conditions in real-time," *Applied Sciences*, vol. 87, no. 6, pp. 1–11, 2016.
- [24] H. Dong, "The design of a toll plaza," in *Proceedings of the 5th International Conference on Computer-Aided Design, Manufacturing, Modeling and Simulation*, pp. 1–5, Busan, South Korea, April 2017.

- [25] M. A. d. F. Caldas and K. T. Sacramento, "Simulation model of discret events applied to the planning and operation of a toll plaza," *Journal of Transport Literature*, vol. 10, no. 3, pp. 40–44, 2016.
- [26] S. Kim, "The toll plaza optimization problem: design, operations, and strategies," *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 1, pp. 125–137, 2009.
- [27] H. Lin, J. Zhu, Y. Xiao, and L. Han, "Optimal design of toll plaza based on traffic flow cellular automata," in *4th International Conference on Information Science and Control Engineering*, pp. 1066–1072, 2017.
- [28] Y.-J. Liu, J. Cao, X.-Y. Cao, and Y.-B. Zhang, "Optimization of design scheme for toll plaza based on M/M/C queuing theory and cellular automata simulation algorithm," *Modern Applied Science*, vol. 11, no. 7, pp. 1–12, 2017.
- [29] M. Papageorgiou, I. Papamichail, A. D. Spiliopoulou, and A. F. Lentzakis, "Real-time merging traffic control with applications to toll plaza and work zone management," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 535–553, 2008.
- [30] D. Anastasia, "Spiliopoulou, ioannis papamichail, and markos papageorgiou. "Toll plaza merging traffic control for throughput maximization," *Journal of Transportation Engineering*, vol. 136, no. 1, pp. 67–76, 2010.
- [31] R. F. Perry and S. M. Gupta, "Response surface methodology applied to toll plaza design for the transition to electronic toll collection," *International Federation of Operational Research Societies*, vol. 8, no. 3, pp. 707–726, 2001.
- [32] D. R. McDonald Jr., R. E. Stammer Jr., and ASCE. Members, "Contribution to the development of guidelines for toll plaza design," *Journal of Transportation Engineering*, vol. 127, no. 3, pp. 215–222, 2001.

## Research Article

# The Approach of Hierarchical Linear Model to Exploring Individual and Team Creativity: A Perspective of Cultural Intelligence and Team Trust

Chenhan Huang,<sup>1,2</sup> Changqing He,<sup>3</sup> and Xuesong Zhai <sup>4</sup>

<sup>1</sup>School of Public Affairs, University of Science and Technology of China, 96 Jinzhai Road, Hefei 230026, Anhui, China

<sup>2</sup>School of Art, Anhui University of Finance and Economics, 962 Caoshan Road, Bengbu 233030, Anhui, China

<sup>3</sup>College of Economics and Management, Nanjing University of Aeronautics and Astronautics, 29 Jiangjun Ave., Nanjing 211106, Jiangsu, China

<sup>4</sup>College of Education, Zhejiang University, 866 Yuhangtang Road, Hangzhou, Zhejiang, China

Correspondence should be addressed to Xuesong Zhai; xszhai@zju.edu.cn

Received 19 August 2020; Revised 1 October 2020; Accepted 8 October 2020; Published 24 October 2020

Academic Editor: Shaohui Wang

Copyright © 2020 Chenhan Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to promote the creativity of interorganizational teams has always been the focus among scholars and management practitioners. From the perspective of leadership, this study explores the influence of shared leadership on creativity in interorganizational teams. Specifically, this study integrates leadership perspective with trust perspective and explores the mediating role of team trust between shared leadership and creativity at both team and individual level. In addition, this study examines the moderating effect of the leader's cultural intelligence between shared leadership and team trust based on the perspective of leadership situation. The data comes from 275 employees within 54 interorganizational teams. The results show that shared leadership will promote team trust and team trust plays a key mediating role between shared leadership and creativity. Moreover, the relationship between shared leadership and team trust is moderated by the cultural intelligence of leader, such that the positive relationship will be stronger with high cultural intelligence and weaker with low cultural intelligence.

## 1. Introduction

In order to cope with the complex and fierce market competition, more and more organizations tend to seek external cooperation; it is in this context that interorganizational teams emerge as the times require [1–3]. Through the establishment of interorganizational teams, the two partners can complement each other's resources and strengthen the sharing and exchange of information and knowledge, which can effectively promote the generation of creativity [4]. Among the many factors or mechanisms that influence the creativity of the team (or team members), scholars unanimously agree that team trust is one of the most critical factors [5–7]. However, interorganizational teams have many characteristics (see Table 1); for example,

members of interorganizational teams come from different organizations and therefore team members have diverse knowledge, skills, behavioral patterns, cultural backgrounds, etc. [8]. This difference will make it more difficult for team members to communicate, exchange knowledge, and establish mutual trust. At the same time, interorganizational teams are built on the link of common tasks and goals. Only when they have mutual trust, sincere cooperation, and interactivity will they be able to accomplish the task and achieve win-win and cooperative creativity. Thus, it can be seen that, on the one hand, interorganizational teams urgently need to build effective trust among members; on the other hand, the establishment of this kind of team trust has become increasingly difficult in the context of interorganizational cooperation. In fact, in the established

TABLE 1: The features of interorganizational teams and shared leadership.

Characteristics of interorganizational teams	Characteristics of shared leadership
(1) Team members from different organizations have great differences in organizational culture, values, knowledge skills, behavior patterns, etc.	(1) Sharing of power: team members share leadership, so they will have higher autonomy and decision participation.
(2) Team members from different organizations rarely have informal communications and therefore lack emotional communication and contact.	(2) Influence mode: different from the traditional vertical leadership that focuses on authority, shared leadership pays more attention to the horizontal mutual effect among team members.
(3) Interorganizational teams linked by common tasks and goals often need to cooperate with each other to achieve complementary resources and advantages in order to effectively complete the task.	(3) Team members tend to form a common or shared vision.

interorganizational teams, a significant part is due to the lack of trust between partners and members of interorganizational teams, ultimately leading to the failure of creativity. In the cooperation between Yidao Yongche and LeTV, for example, it is precisely due to mutual suspicion between the two parties that the relationship has broken down and deteriorated, and ultimately the cooperation has come to an end. The contradictory and cruel reality lies before us: how can we build team trust and promote creativity in interorganizational teams? Solving this problem not only helps to bridge the gaps in existing research on interorganizational teams, but has important guiding significance for effectively promoting the creativity of interorganizational teams.

It is precisely for this purpose that this study attempts to, from the perspective of leadership, explore the important antecedents that influence team trust and then affect the creativity of interorganizational teams. In fact, Burke et al. [9] have suggested in a review study that leadership is one of the most striking determinants of trust; besides, some scholars have argued that leadership is also an important factor influencing team trust [10]. In the general organizational situation, some scholars have empirically studied the influence of different dimensions of transformational leadership style on team trust [5]. In particular, in interorganizational teams, this paper proposes a shared leadership style in which the leadership of a team is shared by multiple members, and they negotiate together, make joint decisions, and jointly take responsibility for the team's tasks [11], which will be an important factor influencing team trust and then creativity. This is also due to the characteristics of shared leadership: shared leadership style lays more emphasis on mutual influence and motivation among team members, and this way of horizontal leadership helps team members understand and trust each other and thus enhances the sense of team trust. However, the existing research has paid little attention to the role of shared leadership in interorganizational teams, especially the integration of the two different perspectives of shared leadership and team trust. In addition, "innovation projects across different organizations often involve independent key persons on a voluntary basis with the absence of a formal structure and no hierarchical controls..." [2]; p. 391); in this situation, team members have a common leadership role or executive leadership behavior, which is especially necessary for creativity [11].

In addition, the success of interorganizational creativity depends not merely on the results of team creativity, but also on the creativity of individual team employees [12, 13]. Different from team creativity, individual creativity mainly refers to the novel and useful ideas related to work flow, product, and services [14]. Although previous literature has pointed out that two creativity types (i.e., team creativity and individual creativity) are related to each other (e.g., [15]), are the influencing mechanisms of the same leadership style (i.e., shared leadership) on both team creativity and individual creativity consistent? It is still largely unknown [16]. Therefore, this study will propose a multilevel model to explore how shared leadership affects team creativity and individual creativity by investigating the key team process of team trust. Furthermore, the existing literature also calls for the need to simultaneously study the impact of cross-level leadership on the results of individual creativity [17]. In view of this, in the context of interorganizational teams, this study will integrate two different perspectives, namely, shared leadership and team trust, to explore whether and how shared leadership can affect team trust and then influence both team and individual creativity.

In addition to exploring the mediating role of team trust between shared leadership and creativity, this paper will further explore the boundary conditions that influence the effects of shared leadership. In the past, only a few studies have explored the contextual factors that affect the role of shared leadership from the perspective of task characteristics, such as task interdependence [18]. So far as the literature is concerned, there is little literature on the boundary conditions of the role of shared leadership from the perspective of individual characteristics of team leaders. In fact, scholars have called for strengthening the exploration of contextual factors of shared leadership. In 2007, for example, Carson et al. suggested that "future work should focus on a more detailed understanding of the nature of shared leadership, its development, and boundary conditions on its effectiveness" (p. 1230). In order to make up for the deficiency of this research, this study considers cultural intelligence of team leaders as a key moderator. The leadership theory has pointed out that the mechanism of leadership is not in a vacuum [19]; thus, the exertion of shared leadership also depends on the specific organizational situation. In interorganizational teams, cultural intelligence—the ability of collecting and processing information in different cultural

contexts, making the right judgments, and taking corresponding actions—is just such a specific organizational situation [20]. In interorganizational teams, as team members come from organizations with different cultural backgrounds, when team leaders have relatively strong cultural intelligence, it will be conducive to identify the characteristics and expertise of different team members, accelerate power sharing, promote the exertion of effect of shared leadership, and then affect team trust and creativity. Therefore, this study will explore the moderating role of leaders' cultural intelligence. To sum up, the hypothesis model of this study is shown in Figure 1.

Thus, the main contributions of this study are as follows. First of all, in the context of interorganizational teams, it makes clear that shared leadership will be the key to solving the problem of “how to enhance team trust” as an arrangement mechanism of new leadership and team power. This not only enriches the research concerning interorganizational teams, but greatly expands the scope of application of shared leadership, deepening our understanding of shared leadership. It also echoes the appeal of Burke et al. [9] regarding the integration of two different perspectives, namely, leadership and trust. Secondly, in the context of interorganizational teams, it views the impact of trust on two different types of creativity from the perspective of the team, which deepens our influence on trust in different contexts. Besides, based on the perspective of leadership situation and the characteristics of shared leadership, this paper initiatively proposes the important situational role of leaders' cultural intelligence, which deepens our understanding of cultural intelligence and the effectiveness of shared leadership.

## 2. Literature Review and Hypotheses Development

**2.1. Team Trust and Creativity.** Team trust is a kind of emotional and cognitive team process in which team members have good faith in each other and have strong confidence in the ability of the whole team and trust in other team members in an interorganizational situation [5, 21]. In recent years, more and more scholars have begun to pay attention to the impact of team trust on creativity. For example, some scholars believe that both goodwill trust and competence trust will produce a positive impact on creativity [22]; Chen et al. [23] also confirmed that, in an R&D team, trust among team members can effectively promote creativity. Based on previous studies on team trust and creativity, this study specifically points out that team trust is an important source of creativity in the situation of interorganizational teams.

Interorganizational team members' knowledge structures are relatively varied since their prior working experience is different. However, their different knowledge structures may attract other members when team trust is high, which may generate their passion to exchange mind and thinking. In this situation, their trust is conducive to the production of creative thinking and enhance the level of creativity of the whole team.

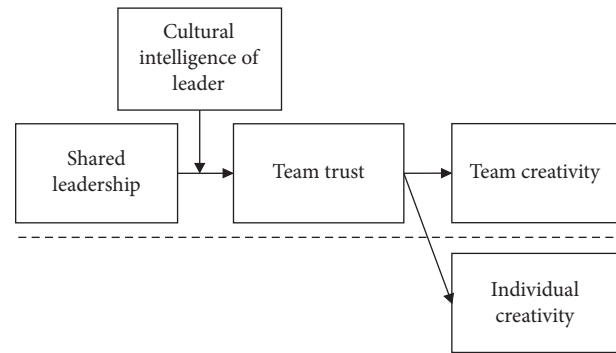


FIGURE 1: Hypothesis model.

In the next place, emotionally, team trust in interorganizational teams will promote effective communication and exchange as well as the process of knowledge sharing among team members, leading to higher team creativity [4, 5]. The natural features of interorganizational teams cause team members to often have feelings of estrangement with lack of communication and exchange. When team trust is high, it is often able to provide such a mechanism for communication and exchange among team members, which makes team members exchange ideas and viewpoints without reservation, accelerates the process of knowledge sharing, and ultimately improves team creativity [25]. In conclusion, this study presents the following hypotheses.

H1a: in interorganizational teams, team trust will be positively related to team creativity.

In a similar way, in the situation of interorganizational teams, when team trust is at a high level, it will benefit individual employees and promote their individual creativity. When team trust is high, the collision of this opinion and sharing of knowledge will also enable employees to obtain more creative inspirations, leading to the improvement of their individual creativity [6]. Moreover, when interorganizational team members trust each other, individual employees will feel the collective strength. This sense of trust will greatly enhance the psychological security of employees and encourage them to carry out creative activities more boldly.

H1b: in interorganizational teams, team trust will be positively related to individual creativity.

**2.2. Shared Leadership and Team Trust.** At present, a lot of literature has confirmed that leadership style is an important factor to promote team creativity [26, 27]. For example, Braun et al. [28] pointed out the positive role of transformational leadership style in team trust. Nonetheless, due to the specificity of interorganizational teams, team trust of interorganizational teams seems to be more difficult and complex. Therefore, choosing the type of leadership style or leadership mode is the key to building team trust. Shared leadership puts more emphasis on the sharing of power, and mutual influence among team members [11] makes it easier for team members to build mutual trust and common vision.



Therefore, this study believes that shared leadership will positively influence team trust, specifically as follows.

First, shared leadership helps to promote the recognition and trust of each other's ability among team members, form a relatively consistent view and cognition, and promote the generation of team trust. Carson et al. [11] emphasize that shared leadership is more of a horizontal leadership style, in which team members interact and make decisions among themselves, rather than a traditional top-down leadership style. Hence, shared leadership can lead to communication and cooperation among team members, and this kind of full communication and sharing is conducive for team members to the appreciation and recognition of each other's ability, ultimately resulting in the generation of team trust.

Moreover, shared leadership can boost emotional communication and exchange among team members, enhance mutual trust, and promote team trust. In other words, shared leadership gives team members greater power and encourages team members to have more opportunities for internal communication and exchange [11]. Frequent and close contact is beneficial to make original team members from different organizations get goodwill from each other and form a close emotional bond. Therefore, shared leadership is beneficial to the generation of team trust. To sum up, this study puts forward the following hypothesis.

H2: in interorganizational teams, shared leadership will be positively related to team trust.

*2.3. The Mediating Role of Team Trust.* At present, some scholars have begun to pay attention to the mediating mechanism of shared leadership on creativity; for instance, Gu et al. [18] revealed the influence mechanism of shared leadership on team creativity from the perspective of knowledge sharing. Based on the context of interorganizational teams, this study argues that team trust plays a mediating role in the relationship between shared leadership and creativity (at both team and individual level). Shared leadership is conducive to mutual trust among team members, which will provide intrinsic motivation and information communication for creativity at both team and individual level [29]. This is particularly important for interorganizational teams because team members from different cultural backgrounds urgently need to build trust to solve a challenging creative task [30].

From hypothesis 1 and hypothesis 2, we can find that, in interorganizational teams, team trust is the key to influencing creativity, and the establishment of such team trust often needs shared leadership. Specifically, in interorganizational teams, the authorized and horizontal leadership mode of shared leadership can accelerate the frequency of communication, exchange, and sharing among team members and enable team members to form emotionally mutual trust and cognitively mutual trust, while this emotional and cognitive trust will ultimately be the basis for creativity. As a matter of fact, scholars have pointed out that leadership style such as shared leadership can affect team-level and individual-level behavioral outcomes, such as creativity [31] by influencing team process (such as team

trust). Therefore, this study presents the following hypotheses.

H3a: team trust mediates the relationship between shared leadership and team creativity in interorganizational teams.

H3b: team trust mediates the relationship between shared leadership and individual creativity in interorganizational teams.

*2.4. The Moderating Role of Cultural Intelligence.* Previous literature has pointed out that the relationship between shared leadership and creativity is likely to be influenced by some situational factors [11, 32]. For example, some scholars have discussed the moderating role of task interdependence between shared leadership and creativity [18]. However, most of the existing studies on the boundary conditions of the role of shared leadership focus on task characteristics and organizational structures, and few studies have explored the moderating effect of individual traits of leaders in interorganizational teams. Actually, in interorganizational teams, although power is authorized to each team member, whether the power-sharing mechanism can come into play to a great extent still cannot leave leaders' personal guidance and promotion [11]. Specially, in the situation of interorganizational teams, when leaders themselves have relatively high cultural intelligence, that is, for employees from different organizational and cultural backgrounds, all leaders can interact with them more effectively and adapt to employees in different cultural backgrounds, which will give full play to the role of shared leadership. Hence, this study argues that leader's cultural intelligence will affect the action process of shared leadership in creativity.

This study suggests that cultural intelligence will positively moderate the positive effect of shared leadership on team trust. Firstly, when leader's cultural intelligence is at a high level, the leaders of interorganizational teams can calmly deal with the conflicts and frictions among team members caused by the power-sharing mechanism, accelerating the role of emotional bond brought by shared leadership and enhancing the positive role of shared leadership in team trust [33]. When leaders' cultural intelligence is at a high level, they can not only effectively identify the emotional needs of different team members, but also, when members from different backgrounds have conflicts and frictions, intercede and resolve conflicts and frictions by virtue of their high cultural intelligence, enhance emotional exchange and communication among team members, and realize team trust to the maximum extent. For example, when the members of interorganizational teams from different organizations have disagreements because of their different ideas and goals, the leaders with high cultural intelligence can not only find such differences in a timely manner, but also use their high cultural intelligence for intercession when matters are not serious, enabling team members to live in harmony and enhancing team trust. In the next place, when cultural intelligence is stronger, leaders can be more effective in identifying the ability of team

employees from different organizations. This will accelerate the cognitive consistency of shared leadership, speed up the understanding and collaboration among team members, and increase the positive role of shared leadership in team trust [34]. The leaders with high cultural intelligence, for instance, can effectively identify the strengths and weaknesses of different employees, which can tilt the team's resources to those who need them more; besides, the power-sharing mechanism can be carried out more smoothly, which can accelerate mutual appreciation and recognition among team members, enhance the role of shared leadership, and promote team trust. Therefore, this study proposes the following hypothesis.

H4: leader's cultural intelligence will moderate the relationship between shared leadership and team trust, such that this positive relationship is stronger when leader's cultural intelligence is higher.

### 3. Method

**3.1. Sample and Procedures.** The research team cooperates with a university in China to issue survey questionnaires. With its assistance, the research team distributes a total of 68 questionnaires of interorganizational teams. These 68 teams come from 19 different companies, and their industries include information technology, manufacturing industry, medical treatment, and scientific research institutions. When conducting the study, the research team takes a variety of approaches to ensure that respondents are clear about the purpose of the study: first of all, there will be a special researcher in this research team to contact the selected interorganizational teams and explain the purpose of this study; secondly, the cover of the paper questionnaire will have a special page to introduce the purpose of this study and fill in the points for attention; in addition, every variable measurement item in the questionnaire is preceded by an introduction to remind the respondents that the investigation objects are the interorganizational cooperation teams. "Among team members of interorganizational cooperative creativity team. . .," for example, the purpose of doing so is to repeatedly remind the respondents that the investigation objects are the interorganizational cooperation teams rather than the original department teams of their own work units.

In order to increase the recovery rate of questionnaires, we adopt two ways. In the first place, before the research, the research team will communicate and exchange information with the tested interorganizational teams, try to make explanation and illustration to make the investigated interorganizational teams fully understand the purpose of this study, and invite them to participate in this study. Next, after the questionnaires are issued, the research team also follows up the filling of questionnaires by means of subsequent telephone contacts or e-mail contacts. Finally, the research team recovers effective questionnaires from 54 teams, including 54 team leaders and 275 employees, with a recovery rate of 79.4%. Among these 54 sample teams, those whose team size is less than or equal to 5 persons take up 7.4%, those whose team size is from 6 persons to 10 persons

account for 40.7%; 27.8% of the teams have 11 persons to 15 persons; and 24.1% of the teams have 16 persons and above. Those with the establishment period of 1 to 6 months take up 16.6%; those with the establishment period of 7 to 12 months account for 24.1%; 24.1% of the teams have the establishment period of 13 to 36 months; and those with the establishment period of 36 months and above account for 35.2%.

**3.2. Measurement.** In order to ensure the validity and reliability of measurement tools, this study tries to adopt a mature scale of the existing literature and then make appropriate revision according to the purpose and specific situation of the research. Prior to the formal finalization and investigation of the questionnaires, one of the authors makes presurvey questionnaires to evaluate the appropriateness of the design and wording of questionnaires and then revises the questionnaires according to the advice provided by the presurvey respondents. Each variable is measured with 5-point Likert scale.

- (1) Shared leadership. There are 10 items in the scale of shared leadership, which come from Shane Wood and Fields [35]. These items are mainly used to measure the degree to which team members participate in the decision-making process of teams. The scale is filled by team members, for example, "team members in interorganizational teams share in deciding on the best course of action when a problem faces the team." In order to ensure the reliability and validity of the scale, this study finally retains seven items with the highest loading value.
- (2) Team trust. It is derived from the trust scale prepared by Dakhli and De Clercq [36]. The scale is filled in by team members and includes 3 items, for example, "team members in interorganizational teams have mutual trust and are supportive of change."
- (3) Cultural intelligence. With reference to the cultural intelligence scale developed by Ang et al. [37], the scale measures the cultural intelligence of interorganizational team leaders from different aspects such as metacognition, cognition, motivation, and behavior. The scale is completed by the leaders of interorganizational teams with 20 items, for example, "I enjoy interacting with people from different cultures."
- (4) Individual creativity. The variable of individual creativity adopts the scale developed by Dul et al. [38] with three items. These items are used to measure the extent to which individuals come up with new ideas. For example, "in my work, I often have new and creative ideas" (1 = strongly disagree; 5 = strongly agree).
- (5) Team creativity. The variable scale is selected from Shin and Zhou [14] with four items. These items are used to measure the extent to which new ideas are generated by interorganizational teams. This scale is filled out by leaders in the interorganizational teams,

for instance, “how well does your team produce new ideas?” (1 = strongly poorly; 5 = very much).

In addition, with reference to previous studies by scholars, we controlled for gender, age, and educational level at the individual level [39], and team’s tenure and size at the team level [40].

**3.3. Data Aggregation.** To determine if individual-level data can be aggregated to the team level, this study examines intragroup consistency and intergroup variance [41]. Intragroup consistency is tested by Rwg index while intergroup variance is tested by ICC(1) and ICC(2). Previous literature has pointed out that shared leadership is a team-level variable [42]. Data analysis indicates that ICC(1) = 0.17, ICC(2) = 0.52, so shared leadership can be aggregated into a team-level variable. Furthermore, previous literature has revealed that team trust is also a variable at the team level (e.g., [5]). The result shows that ICC(1) and ICC(2) of team trust are 0.21 and 0.57, respectively. The ICC result demonstrates that shared leadership and team trust can indeed be aggregated into a team-level variable (e.g., [16]). In addition, in order to further determine that shared leadership and team trust can be aggregated into team-level variables from employees’ individual level, this study examines internal consistency within the group. The result suggests that the average Rwg scores of shared leadership and team trust are both 0.92, and all scores exceed the threshold of 0.70 [41]. The above result manifests that the data of shared leadership and team trust can be aggregated into team-level variables.

## 4. Data Analysis and Results

**4.1. Analysis of Reliability and Validity.** Before hypotheses testing, we have first performed a confirmatory factor analysis (CFA) on the main variables. The loading values, coefficient of internal consistency (Cronbach’s alpha), composite reliability, and AVE values of the main study variables are shown in Table 2. Table 2 shows that the reliability (Cronbach’s alpha) and composite reliability of various variables are between 0.733 and 0.903, which are higher than the threshold of 0.70. The scores of average variance extraction value (AVE) of all variables are between 0.505 and 0.766, which exceed the threshold of 0.50.

To further verify the discriminant validity of various variables, this study adopts the comparison method of average variation extraction (AVE), and the result shows that the AVE value of each variable is higher than the critical value of 0.5. Furthermore, as shown in Table 3, the AVE square root of each variable is greater than the correlation coefficient between the variables, which further explains that there is good discriminant validity between the variables.

**4.2. Descriptive Statistics of Variables.** Table 3 summarizes the mean value, standard deviation, and correlation coefficient of all variables. We found that shared leadership and team trust are both significantly correlated to team creativity.

TABLE 2: Composite reliability and average variance extracted (AVE).

Variables	Items	Factors loading	Cronbach’s alpha	Composite reliability	AVE
Shared leadership	7	0.585~0.726	0.798	0.803	0.505
Team trust	3	0.795~0.815	0.733	0.849	0.652
Cultural intelligence	20	0.766~0.773	0.769	0.887	0.662
Individual creativity	3	0.873~0.877	0.846	0.907	0.766
Team creativity	4	0.724~0.788	0.759	0.885	0.659

Note. ML: moral leadership, LMX: leader-member exchange, TI: team identification, EC: employee creativity. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .  $N = 275$ .

### 4.3. Hypothesis Test

**4.3.1. Team-Level Analysis Result.** To test the team-level hypotheses, we use hierarchical regression analyses.

H1a assumes that team trust will positively affect team creativity. As shown in Model 5 of Table 4, team trust positively affects team creativity ( $\beta = 0.374$ ,  $p < 0.01$ ), thereby supporting H1a.

At the same time, model 1 in Table 4 suggests that shared leadership positively affects team trust ( $\beta = 0.642$ ,  $p < 0.001$ ). Hence, we support hypothesis 2.

The result of Table 4 also shows that the coefficient for shared leadership on team creativity decreases when the effect of team trust is included in the regression equation. The coefficient decreases from 0.407 ( $p < 0.01$ , model 8) to 0.279 ( $p < 0.05$ , model 9). Thus, H3a is supported.

H4 assumes that leader’s cultural intelligence can moderate the relationship between shared leadership and team trust, where such positive relationship will become stronger under higher cultural intelligence. The result of model 3 in Table 4 suggests that interaction between shared leadership and cultural intelligence significantly relates to team trust ( $\beta = 0.222$ ,  $p < 0.05$ ), supporting H4.

We further calculate the simple slopes for the relationship between shared leadership and team trust at higher (mean + s.d.) and lower (mean – s.d.) level of cultural intelligence [43]. Figure 2 shows that cultural intelligence amplifies the positive relationship between shared leadership and team trust. As expected, the relationship between shared leadership and team trust is significant (simple slope = 0.432,  $p < 0.05$ ) at low levels of cultural intelligence and becomes stronger (simple slope = 0.876,  $p < 0.01$ ) at high levels of cultural intelligence.

**4.3.2. Hierarchical Analysis Result.** In this study, the Hierarchical Linear Model (HLM) is used to examine the hierarchical research hypothesis.

Table 5 summarizes the analysis results of HLM. In step 1, model 1 shows that shared leadership is positively related to individual creativity ( $\beta = 0.757$ ,  $p < 0.001$ ). In step 2, team trust should be positively related to individual creativity

TABLE 3: Means, standard deviations, and correlation.

Variables	Mean	s.d.	1	2	3	4	5	6
<i>Individual-level variables</i>								
(1) Gender	0.72	0.47						
(2) Age	2.47	0.62	0.07					
(3) Education level	3.12	0.67	0.04	-0.04				
(4) Individual creativity	3.97	0.55	-0.03	0.131*	-0.08	( 0.875 )		
<i>Team-level variables</i>								
(1) Team size	6.76	1.59						
(2) Team tenure	3.72	1.10	-0.23+					
(3) Shared leadership	3.88	0.34	0.13	0.21	( 0.711 )			
(4) Team trust	4.16	0.30	-0.11	0.21	0.63***	( 0.807 )		
(5) Cultural intelligence	3.97	0.35	0.16	0.16	0.34	0.19	( 0.814 )	
(6) Team creativity	3.94	0.31	-0.10	0.29*	0.43***	0.42**	0.60***	( 0.812 )

Note. N (employees) = 275; N (leaders/teams) = 54. Square roots of AVE are in parentheses on the diagonal. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

TABLE 4: Regression results for team-level analysis.

Level and variables	Trust				Team creativity			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
<b>Level 2</b>								
Team size	-0.185	-0.185	-0.125	-0.035	-0.010	-0.134	-0.086	0.005
Team tenure	0.029	0.026	0.052	0.281*	0.210	0.111	0.027	0.012
SL	0.642***	0.632***	0.654***				0.407**	0.279*
Trust					0.374**	0.276*		0.200*
CI		0.029	-0.006			0.550***		
<b>Interaction item</b>								
SL × CI			0.222*					
R <sup>2</sup>	0.428	0.429	0.474	0.085	0.218	0.491	0.530	0.667
ΔR <sup>2</sup>	0.428	0.001	0.044	0.085	0.133	0.272	0.530	0.137
F	12.496***	9.211***	8.641***	2.372	4.654**	11.802***	18.822***	24.588***
ΔF	12.496***	0.062	4.058*	2.372	8.519**	26.207***	18.822***	20.202***

Note. N = 54. SL: shared leadership; CI: cultural intelligence of leader.

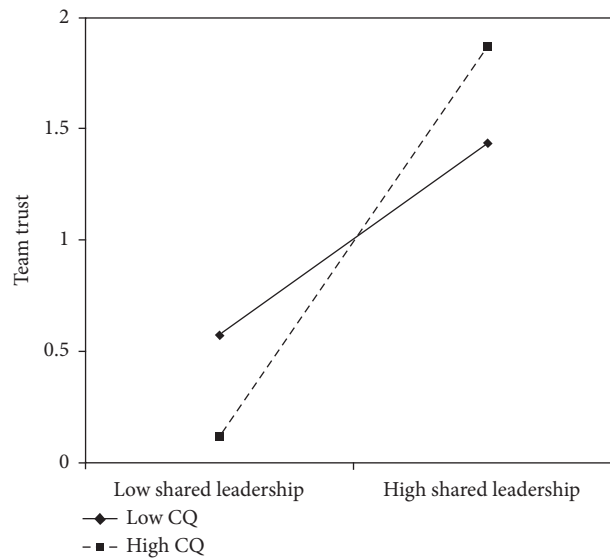


FIGURE 2: Moderating effect of leader’s cultural intelligence on the relationship between shared leadership and team trust.

(H1b). The result of model 3 demonstrates that team trust is positively related to individual creativity ( $\beta = 0.667$ ,  $p < 0.001$ ), thereby supporting hypothesis 1b. In step 3, the

coefficient for shared leadership on individual creativity decreases when the effect of team trust is included in the regression equation. The coefficient decreases from 0.757

TABLE 5: HLM results for cross-level analysis.

Level and variables	Individual creativity		
	Model 1	Model 2	Model 3
Level 1			
Gender	0.006	0.007	0.004
Age	0.111	0.113	0.106
Education level	0.022	0.022	0.017
Level 2			
Team size	-0.025	-0.015	0.009
Team tenure	0.030	0.029	0.053
Shared leadership	0.757***	0.602***	
Team trust		0.257*	0.667***

Note.  $N$  (subordinates) = 275;  $N$  (leaders/teams) = 54. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

( $p < 0.001$ , model 1) to 0.602 ( $p < 0.001$ , model 2), supporting hypothesis 3b.

## 5. Conclusion and Discussion

Interorganizational teams proved to be a very effective way of cooperative creativity. Therefore, how to explore the generation mechanism of creativity in the context of interorganizational teams seems to be especially necessary. In particular, based on the characteristics of interorganizational teams, this study proposes that team trust will be the key for promoting interorganizational creativity; then, from the perspective of leadership, it suggests that shared leadership as an emerging power-sharing system and arrangement will be the key factor for building team trust. On this basis, this study also investigates the situational moderating role of leaders' cultural intelligence in the relationship between shared leadership and team trust.

**5.1. Theoretical Implications.** The greatest contribution of this study lies in its confirmation of how and when shared leadership affects creativity at both team and individual level. Hence, important theoretical implications to the literature of shared leadership and creativity are offered in this study.

First, we propose and empirically test a theoretical model that provides a nuanced depiction of how and under what conditions shared leadership influences creativity. Our study also enriches the literature regarding the consequences of shared leadership by verifying a positive association linking shared leadership to creativity at both team and individual level. In this process, we contribute to shared leadership literature by demonstrating that shared leadership, either at the individual or team level, is valuable. To our knowledge, no prior research has considered team trust as a critical mediating mechanism linking shared leadership and creativity. Previous research has also shown that some potential team processes (such as knowledge sharing) can mediate the relationship between shared leadership and creativity [18]. Other scholars also try to interpret the relationship between leadership and creativity from the perspective of trust [6]. Therefore, the results of this study are also further expansion of the literature in this aspect. We go further in this theoretical model by highlighting the moderating role of cultural

intelligence in strengthening the above link. Indeed, shared leadership, cultural intelligence, and team trust constitute three important areas in organizational behavior research that have not been integrated until now as a means of predicting creativity at both team and individual level.

Second, our findings contribute to the creativity literature by identifying the mediating role of team trust. This study has confirmed that team trust is a key mediating mechanism between shared leadership and creativity, which expands our understanding of the creativity of interorganizational teams. Although the existing research has proven that, in general team situations, team trust is the key to unlocking team creativity (e.g., [5]), and some scholars have also pointed out that team trust may play a key role in interorganizational alliances (such as knowledge sharing, information exchange, and team learning), there is still a great shortage of empirical research in interorganizational team context. In this context, this study empirically tests the key impact of team trust on the creativity in interorganizational teams, which helps us have a better understanding of the source of creativity in interorganizational teams.

Third, we probe into the situational variables that affect the process of shared leadership [44]. We draw the conclusion that leader's cultural intelligence can strengthen the positive relationship between shared leadership and team trust. The result indicates that team trust and creativity are based not only on shared leadership, but also on leader's cultural intelligence. This finding responds to the previous scholars' call for exploring the situational factors that play a role in shared leadership [11].

**5.2. Practical Implications.** The results of this paper are of great enlightenment significance in respect of how to carry out better creativity in interorganizational teams.

First, interorganizational teams should establish mutual trust among each other. The research shows that team trust is a key factor in creativity. Hence, the key to the success of creativity is to build trust between the two partners and among team members. In the process of interorganizational team creativity, only by offering mutual help to make up for what the other lacks rather than "holding back a trick or two," is it possible to establish mutual trust among members to the maximum extent, boost the sharing of knowledge and information, and accelerate the production of creativity.

What is more, the shared leadership model should be adopted and applied widely. The research result proves that shared leadership has a positive impact on both team trust and cooperative creativity. Therefore, in the context of interorganizational teams, it is necessary to use this new power-sharing mechanism to mobilize the enthusiasm of every member in the teams to the maximum extent, give play to their creativity, and promote the formation of cooperative creativity.

Finally, when considering the interorganizational team leadership, we must pay special attention to examining its cultural intelligence. Although research has shown that the use of shared leadership is an effective

power distribution model and institutional arrangement, the research conclusion reveals that leaders' personal ability traits, especially cultural intelligence, still play a vital role in it. Hence, when selecting and appointing interorganizational team leaders, what calls for special attention is to examine the cultural intelligence of leaders and preferably choose those with high cultural intelligence, which helps to accelerate the exertion of the role of shared leadership and achieve the success of interorganizational teamwork.

**5.3. Limitations.** This study has some limitations as well.

In the first place, this study takes a self-evaluation approach to measure employees' individual creativity. Although most of the current literature uses the method of leadership assessment to evaluate employees' individual creativity [16], the method of employees' self-evaluation can be accepted as well [38, 45]. Specifically, in the context of this study, the method of employees' individual creativity may be more applicable to the measurement of individual creativity. Firstly, in the context of interorganizational teams, it is very difficult for a team leader to have an accurate and comprehensive observation of each member's creativity. Besides, creativity itself is a subtle behavior, and only the innovators themselves are more aware of how much creative ideas they have play a role in their work assignments [38, 45]; spectators may not necessarily fully and accurately understand the creative ability of a certain employee, so they may not be able to accurately evaluate the performance of an employee's individual creativity [46]. In the next place, although the self-assessment method can bring subjective errors, previous literature has pointed out that individual creativity evaluated by oneself and individual creativity evaluated by others are highly relevant [47]; moreover, subsequent empirical studies also suggest that individual creativity evaluated by oneself also has a good validity [38]. Nonetheless, future research may adopt a more objective approach or a combination of subjective and objective approach with multiple evaluation subjects to measure employees' individual creativity so as to avoid possible problems.

Furthermore, it belongs to cross-sectional research in terms of research design. Since cross-section data cannot provide strong evidence for the causal relationship between variables, future research can employ the method of longitudinal research to test the influence of mediating and causal relationship between mediating variables.

Finally, when examining situational factors, we only consider cultural intelligence—individual traits of leaders. However, in future studies, other situational factors, such as the personal qualities of employees, also need to be taken into consideration.

### Data Availability

All data, models, or code generated or used during the study are proprietary or confidential in nature and may only be provided with restrictions.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

### Authors' Contributions

Chenhan Huang and Changqing He contributed equally to this work.

### Acknowledgments

This research work was supported by the 2020 Humanities and Social Science Projects of the Ministry of Education (Grant no. 20YJC880118).

### References

- [1] D. Faems, B. Van Looy, and K. Debackere, "Interorganizational collaboration and innovation: toward a portfolio approach," *Journal of Product Innovation Management*, vol. 22, no. 3, pp. 238–250, 2005.
- [2] A. Rese, H.-G. Gemünden, and D. Baier, "Too many cooks spoil the broth': key persons and their roles in inter-organizational innovations," *Creativity and Innovation Management*, vol. 22, no. 4, pp. 390–407, 2013.
- [3] R. M. Stock and N. A. Zacharias, "Patterns and performance outcomes of innovation orientation," *Journal of the Academy of Marketing Science*, vol. 39, no. 6, pp. 870–888, 2011.
- [4] N. Panteli and S. Sockalingam, "Trust and conflict within virtual inter-organizational alliances: a framework for facilitating knowledge sharing," *Decision Support Systems*, vol. 39, no. 4, pp. 599–617, 2005.
- [5] K. Boies, J. Fiset, and H. Gill, "Communication and trust are key: unlocking the relationship between leadership and team performance and creativity," *The Leadership Quarterly*, vol. 26, no. 6, pp. 1080–1094, 2015.
- [6] N. Y. Jo, K. C. Lee, D. S. Lee, and M. Hahn, "Empirical analysis of roles of perceived leadership styles and trust on team members' creativity: evidence from Korean ICT companies," *Computers in Human Behavior*, vol. 42, pp. 149–156, 2015.
- [7] J. Wu, H. Zhao, and D. Pan, "Effects of different trust on team creativity: taking knowledge sharing as a mediator," in *Proceedings of the International Symposium on Knowledge and Systems Sciences*, Springer Singapore, Kobe, Japan, pp. 44–56, November 2016.
- [8] H. C. Dekker, "Control of inter-organizational relationships: evidence on appropriation concerns and coordination requirements," *Accounting, Organizations and Society*, vol. 29, no. 1, pp. 27–49, 2004.
- [9] C. S. Burke, D. E. Sims, E. H. Lazzara, and E. Salas, "Trust in leadership: a multi-level review and integration," *The Leadership Quarterly*, vol. 18, no. 6, pp. 606–632, 2007.
- [10] P. Lee, N. Gillespie, L. Mann, and A. Wearing, "Leadership and trust: their effect on knowledge sharing and team performance," *Management Learning*, vol. 41, no. 4, pp. 473–491, 2010.
- [11] J. B. Carson, P. E. Tesluk, and J. A. Marrone, "Shared leadership in teams: an investigation of antecedent conditions and performance," *Academy of Management Journal*, vol. 50, no. 5, pp. 1217–1234, 2007.
- [12] M. Hoegl and K. P. Parboteeah, "Creativity in innovative projects: how teamwork matters," *Journal of Engineering and Technology Management*, vol. 24, no. 1–2, pp. 148–166, 2007.
- [13] H. Liao, D. Liu, and R. Loi, "Looking at both sides of the social exchange coin: a social cognitive perspective on the joint effects of relationship quality and differentiation on creativity," *Academy of Management Journal*, vol. 53, no. 5, pp. 1090–1109, 2010.



- [14] S. J. Loi and J. Zhou, "When is educational specialization heterogeneity related to creativity in research and development teams? Transformational leadership as a moderator," *Journal of Applied Psychology*, vol. 92, no. 6, pp. 1709–1721, 2007.
- [15] A. Pirolo-Merlo and L. Mann, "The relationship between individual creativity and team creativity: aggregating across people and time," *Journal of Organizational Behavior*, vol. 25, no. 2, pp. 235–257, 2004.
- [16] Y. Gong, T.-Y. Kim, D.-R. Lee, and J. Zhu, "A multilevel model of team goal orientation, information exchange, and creativity," *Academy of Management Journal*, vol. 56, no. 3, pp. 827–851, 2013.
- [17] D. T. Yoshida, S. Sendjaya, G. Hirst, and B. Cooper, "Does servant leadership foster creativity and innovation? A multi-level mediation study of identification and prototypicality," *Journal of Business Research*, vol. 67, no. 7, pp. 1395–1404, 2014.
- [18] J. Gu, Z. Chen, Q. Huang, H. Liu, and S. Huang, "A multilevel analysis of the relationship between shared leadership and creativity in inter-organizational teams," *The Journal of Creative Behavior*, vol. 52, no. 2, pp. 109–126, 2018.
- [19] D. V. Day, P. Gronn, and E. Salas, "Leadership in team-based organizations: on the threshold of a new era," *The Leadership Quarterly*, vol. 17, no. 3, pp. 211–216, 2006.
- [20] P. C. Earley and E. Mosakowski, "Cultural intelligence," *Harvard Business Review*, vol. 82, no. 10, pp. 139–158, 2004.
- [21] C. W. Langfred, "Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams," *Academy of Management Journal*, vol. 47, no. 3, pp. 385–399, 2004.
- [22] A. Brattström, H. Löfsten, and A. Richtnér, "Creativity, trust and systematic processes in product development," *Research Policy*, vol. 41, no. 4, pp. 743–755, 2012.
- [23] M. H. Chen, Y. C. Chang, and S. C. Hung, "Social capital and creativity in R&D project teams," *R & D Management*, vol. 38, no. 1, pp. 21–34, 2008.
- [24] S. B. Sitkin, D. M. Rousseau, R. S. Burt, and C. Camerer, "Special topic forum on trust in and between organizations," *Academy of Management Review*, vol. 23, no. 3, pp. 387–531, 1998.
- [25] F. Bidault and A. Castello, "Trust and creativity: understanding the role of trust in creativity-oriented joint developments," *R&D Management*, vol. 39, no. 3, pp. 259–270, 2009.
- [26] K. T. Dirks and D. L. Ferrin, "Trust in leadership: meta-analytic findings and implications for research and practice," *Journal of Applied Psychology*, vol. 87, no. 4, pp. 611–628, 2002.
- [27] M.-J. Shen and M.-C. Chen, "The relationship of leadership, team trust and team performance: a comparison of the service and manufacturing industries," *Social Behavior and Personality: An International Journal*, vol. 35, no. 5, pp. 643–658, 2007.
- [28] S. Braun, C. Peus, S. Weisweiler, and D. Frey, "Transformational leadership, job satisfaction, and team performance: a multilevel mediation model of trust," *The Leadership Quarterly*, vol. 24, no. 1, pp. 270–283, 2013.
- [29] Z. Song, Q. Gu, and F. L. Cooke, "The effects of high-involvement work systems and shared leadership on team creativity: a multilevel investigation," *Human Resource Management*, vol. 59, no. 3, pp. 201–213, 2020.
- [30] A. Serban and A. J. B. Roberts, "Exploring antecedents and outcomes of shared leadership in a creative context: a mixed-methods approach," *The Leadership Quarterly*, vol. 27, no. 2, pp. 181–199, 2016.
- [31] V. C. Nicolaides, K. A. LaPort, T. R. Chen et al., "The shared leadership of teams: a meta-analysis of proximal, distal, and moderating relationships," *The Leadership Quarterly*, vol. 25, no. 5, pp. 923–942, 2014.
- [32] D. S. DeRue, "Adaptive leadership theory: leading and following as a complex adaptive process," *Research in Organizational Behavior*, vol. 31, pp. 125–150, 2011.
- [33] D. C. Thomas, "Domain and development of cultural intelligence," *Group & Organization Management*, vol. 31, no. 1, pp. 78–99, 2006.
- [34] H. C. Triandis, "Cultural intelligence in organizations," *Group & Organization Management*, vol. 31, no. 1, pp. 20–26, 2006.
- [35] M. Shane Wood and D. Fields, "Exploring the impact of shared leadership on management team member job outcomes," *Baltic Journal of Management*, vol. 2, no. 3, pp. 251–272, 2007.
- [36] M. Dakhli and D. De Clercq, "Human capital, social capital, and innovation: a multi-country study," *Entrepreneurship & Regional Development*, vol. 16, no. 2, pp. 107–128, 2004.
- [37] S. Ang, L. Van Dyne, C. Koh et al., "Cultural intelligence: its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance," *Management and Organization Review*, vol. 3, no. 3, pp. 335–371, 2007.
- [38] J. Dul, C. Ceylan, and F. Jaspers, "Knowledge workers' creativity and the role of the physical work environment," *Human Resource Management*, vol. 50, no. 6, pp. 715–734, 2011.
- [39] T. M. Amabile, "A model of creativity and innovation in organizations," *Research in Organizational Behavior*, vol. 10, no. 1, pp. 123–167, 1988.
- [40] D. A. Harrison and K. J. Klein, "What's the difference? diversity constructs as separation, variety, or disparity in organizations," *Academy of Management Review*, vol. 32, no. 4, pp. 1199–1228, 2007.
- [41] L. R. James, R. G. Demaree, and G. Wolf, "Estimating within-group interrater reliability with and without response bias," *Journal of Applied Psychology*, vol. 69, no. 1, pp. 85–98, 1984.
- [42] C. L. Pearce and H. P. Sims Jr., "Shared leadership: toward a multi-level theory of leadership," in *Advances in Interdisciplinary Studies of Work Teams: Team Leadership*, M. Beyerlein, D. Johnson, and S. Beyerlein, Eds., pp. 115–139, JAI Press, Greenwich, CT, USA, 2000.
- [43] L. S. Aiken and S. G. West, *Multiple Regression: Testing and Interpreting Interactions*, Sage Publications, NewburyPark, CA, USA, 1991.
- [44] M. D. Ensley, K. M. Hmieleski, and C. L. Pearce, "The importance of vertical and shared leadership within new venture top management teams: implications for the performance of startups," *The Leadership Quarterly*, vol. 17, no. 3, pp. 217–231, 2006.
- [45] C. E. Shalley, L. L. Gilson, and T. C. Blum, "Interactive effects of growth need strength, work context, and job complexity on self-reported creative performance," *Academy of Management Journal*, vol. 52, no. 3, pp. 489–505, 2009.
- [46] J. Zhou, S. J. Shin, and A. A. Cannella, "Employee self-perceived creativity after mergers and acquisitions: interactive effects of threat-opportunity perception, access to resources, and support for creativity," *Journal of Applied Behavioral Sciences*, vol. 44, no. 4, pp. 397–421, 2008.
- [47] C. M. Axtell, D. J. Holman, K. L. Unsworth, T. D. Wall, P. E. Waterson, and E. Harrington, "Shopfloor innovation: facilitating the suggestion and implementation of ideas," *Journal of Occupational and Organizational Psychology*, vol. 73, no. 3, pp. 265–285, 2000.

## Research Article

# The Development Law of the Freezing Temperature Field of a Calcareous Clay Layer

Zhi Wang,<sup>1</sup> Chuanxin Rong ,<sup>1</sup> Meng Du,<sup>1</sup> Maoyan Ma,<sup>2</sup> and Xiangyang Liu <sup>1,2</sup>

<sup>1</sup>College of Civil Engineering and Architecture, Anhui University of Science and Technology, Huainan 232001, China

<sup>2</sup>Anhui Province Key Laboratory of Building Structure and Underground Engineering, Hefei 230601, Anhui, China

Correspondence should be addressed to Chuanxin Rong; rongcx@ustc.edu

Received 30 July 2020; Accepted 17 September 2020; Published 24 October 2020

Academic Editor: Shaohui Wang

Copyright © 2020 Zhi Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The problem of “difficult” freezing of the calcareous clay layer fractures in freezing pipes has been investigated. Based on the engineering background of the deep calcareous clay in the Yangcun Mine, model tests were carried out in order to conduct in-depth research on the development law of the freezing temperature field of this clay layer. The test results have shown that the calcareous clay has a freezing point of  $-1.3^{\circ}\text{C}$  under the action of both the water and the soil’s chemistry and the supercooling temperature can be as low as  $-3.8^{\circ}\text{C}$  because of its complex mineral composition causing poor thermal conductivity. This means that the calcareous clay will freeze slowly than the other layers of the soil. The time taking for the temperature fields to intersect is 2.5 h, which is equivalent to 127.6 days in the actual engineering. In the three sections, each temperature measurement point in the temperature field had an irregular saddle shape in the temperature space at the same time, and the ratio of the time between the formation and total melting of the frozen wall was 1 : 1.91. The development speed of the thickness of the frozen wall from 5 h to 16 h was 17.9 mm/h, and the development speed from 16 h to 70 h was 1.96 mm/h; corresponding to the actual development speed of the thickness of the frozen wall which were 0.0123 m/d and 0.0014 m/d, respectively. These speeds were significantly slower than the development speed of the thickness of the freezing wall of the general sandy clay layer, which were 0.0515 m/d in the early stage and 0.02 m/d in the later stage. The thin thickness and low strength of the frozen wall of the calcareous clay layer cause the fracture of the frozen pipes, which should be paid attention to in actual engineering construction.

## 1. Introduction

The artificial ground freezing method entails burying a predetermined number of freezing pipes in order to realize heat exchange with the rock and soil through the circulation of a refrigerant. This is done to form a freezing curtain with a certain strength and a water-tight seal, thus creating good “dry ground” construction conditions for underground geotechnical engineering. As an important temporary support structure in the construction of underground frozen rock and soil, the freezing wall technique has unique “natural” advantages, and it also involves many scientific problems: the freezing of the water in the soil which produces frost heave and the formation law of the freezing temperature field, as well as the strength and thickness of the frozen wall, the difficulty of freezing soil with different properties, the influence of the flow velocity and salinity on the formation of

the frozen wall, the influence of frost heave and thaw settlement of the ground on the surrounding structures [1, 2], etc. The freezing temperature field is formed. The study of laws is the most basic. Zhang et al. [3–5], Xiangdong et al. [6–8], and Chuanxin et al. [9–11] idealized the temperature field into a trapezoid or parabola, based on the macroscopic characteristics of the freezing temperature field of multiple rows of tubes, a detailed theoretical analysis of the force state of the frozen wall. However, factors such as the difference in the stratum, the response of soil with different properties at low temperatures, the water content, and the latent heat of the phase change [12] make the process of forming the temperature field both changeable and complicated. In view of the difference in freezing process of different soil layers, He [13] considered the effect of different thicknesses of the sand layer on the temperature of the permafrost and carried out indoor experiments to study the degree of its influence. Haipeng [14]

studied clay using the double coiled tube model test and the frost heave force that was generated during the formation of the frozen wall. The results showed that the frost heave force was closely related to the temperature and that it manifested as spatiotemporal inhomogeneity. Wang and Zhou [15] considered the spatial variability of the soil parameters and combined thermal conductivity and the volumetric heat simulation in a random field, and the research showed that this method was more scientific and suitable. Jinhua [16] carried out model tests to investigate the deformation law of a frozen pipe under the condition of multiloop pipe freezing; they found that the frozen pipe was broken during the excavation stage. The deformation had been basically completed in the early stage of the freezing, and the amount of strain in the frozen pipe in the clay layer was greater than that in the sand layer; this has provided a basis for reducing excavation in the later period and therefore reducing the rate of damage to frozen pipes.

The depth of excavation shafts in coal mines keeps increasing year on year, and deep and water-rich loose alluvial strata have been encountered in many areas in China during the process of freezing shaft sinking. In the Lianghuai mining area, Shungeng Mountain is the boundary to the northwest and the thickness of the alluvial layer gradually deepens. This loose layer is mainly composed of clay, calcareous clay, silty clay, sand, and gravel layers. Freezing pipe fractures often occur when sinking freezing pipes in this soil layer. Distresses often occur, such as when the Pan Sandong air shaft was excavated in an expansive calcareous clay layer with a depth of 252.0~326.0 m, 22 pipes were broken in the process. 34 pipes were broken in a clay layer with a depth of 224.4~237.1 m during the construction of the Xieqiao gangue well [17–19]. When the auxiliary shaft of the Yangcun Mine was excavated to a depth of 404.0~412.0 m, fractures occurred in 15 freezing pipes. The statistics have shown that fractures of freezing pipes generally occurred at the junction of the calcareous clay layer and the sand layer. According to numerous inspections of the mine's site and related studies [20, 21], it was found that it is difficult to form an "effective" frozen wall in this layer of calcareous clay compared to other strata. Therefore, it is necessary to conduct in-depth research on the "hard" freezing characteristics of calcareous clay. In this paper, the deep calcareous clay layer (at a depth of 407.3~445.3 m) of the Yangcun Mine has been taken as the research object. Through indoor model tests and quantitative analysis of the expansion of the calcareous clay, the freezing temperature field and the development law of the frozen wall have been ascertained; these can therefore be used to provide a reference for engineering practice.

## 2. Mineral Composition and Thermal Conductivity of Calcareous Clay

Figure 1 shows the measured temperature curve for hole No. 2 during the artificial freezing process of the auxiliary shaft of the Yangcun Mine. From the figure, it can be seen that the measured temperature of the calcareous clay layer was both higher and lower than the two layers under the same freezing conditions (the clay layer and the fine sand layer). The

temperature of the calcareous clay layer was 6~8 higher, thus showing its "difficult" freezing characteristics.

The reason for the "difficult" freezing of calcareous clay is mainly due to its complex mineral composition, as shown in Table 1 (X-ray diffraction test (XRD) results of dried soil samples), containing Fe, Mn, Ti, Ca, K, S, P, Si, Al, Na, Mg, and other elements, the corresponding minerals are quartz, kaolinite, muscovite, illite, etc., which belong to calcareous kaolinite-illite mudstone, in which Al and Si elements are the main clay minerals and cement chemical composition, of which there is a high content, followed by  $\text{Fe}^{3+}$ ,  $\text{Ca}^{2+}$ , and  $\text{Mg}^{2+}$ , which are ion exchange products; this makes the  $\text{Na}^+$  content low. The calcareous clay in the original environment is in a water-soil system, and a series of geochemical reactions will occur, mainly including ion exchange, dissolution, and precipitation; the reaction process is jointly controlled by the pH and the  $\text{Ca}^{2+}$  [22] ions in the soil. The exchange of  $\text{Na}^+$  and  $\text{K}^+$  ions on the surface of the colloidal particles of the soil and the high-valency  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ , and  $\text{Fe}^{3+}$  ions in the aqueous solution increase the surface charge of the soil, and the specific surface area of the calcareous clay is large (the specific surface area of the calcareous clay can be obtained by the  $\text{N}_2$  adsorption method  $S_a = 74.07 \text{ m}^2/\text{g}$ ), and its adsorption strength is large. The water film on the surface of the particles becomes thicker, and their binding energy is high, so it is difficult to freeze them. Combined with the thickening of the water film and the presence of salt ions, the thermal conductivity of calcareous clay is affected for the following reasons: its freezing temperature is low (Figure 2); the freezing point of calcareous clay is  $-1.3^\circ\text{C}$ ; and its supercooling temperature is as low as  $-3.8$ . The presence of salt ions turns the frozen water between the soil particles into a salt solution; this lowers the freezing point of the clay and increases the adsorption strength of the water. If it is to be frozen, a lower temperature is required to overcome the adsorption potential energy; The second is that the thermal conductivity is 40% and 35% lower than that of sandy soil and sandy clay at room temperature, while the difference is 46% and 35% under freezing conditions, as shown in Table 2. The specific surface area of the particles is large, and the thickness of the combined water film means that most of the calcareous clay particles are not in direct contact with each other. The heat transfer path is the soil particles, water, and soil particles, and the thermal conductivity of water is less than that of the solid particles [23]; therefore, the thermal conductivity of calcareous clay is small. As a result, under the same cooling conditions, a clay layer has poor thermal conductivity, the frozen soil develops slowly, its strength is low, the strength of the frozen wall is uneven in the vertical direction, and shear failure is prone to occur at the interface of adjacent soil layers.

## 3. Experimental Design

### 3.1. Basic Assumptions

- (1) The experiment only considered the distribution law of the temperature field in the soil during the freezing process and did not analyze its mechanical

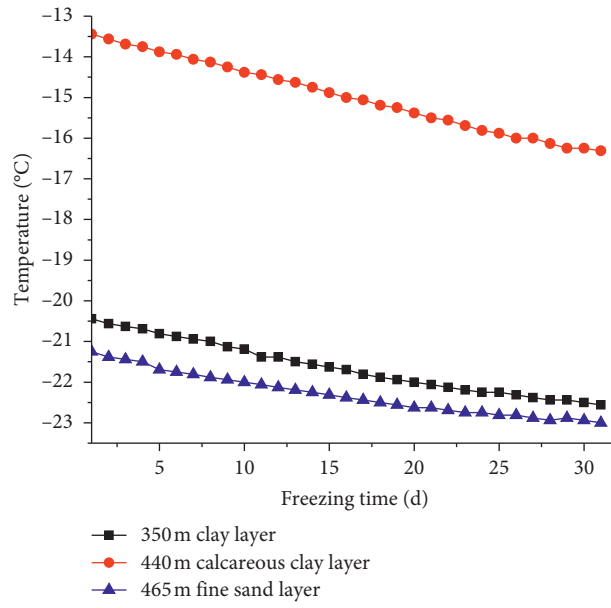


FIGURE 1: Curve of the freezing temperature and the time of the measured temperature of No.2 well in the Yangcun Mine.

TABLE 1: Mineral content of calcareous clay (unit: %).

Sample	Fe <sub>2</sub> O <sub>3</sub>	MnO	TiO <sub>2</sub>	CaO	K <sub>2</sub> O	SO <sub>3</sub>	P <sub>2</sub> O <sub>5</sub>	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Na <sub>2</sub> O	MgO
Undisturbed soil	7.69	0.063	0.76	3.72	2.02	0.10	0.085	55.32	18.91	0.25	2.87

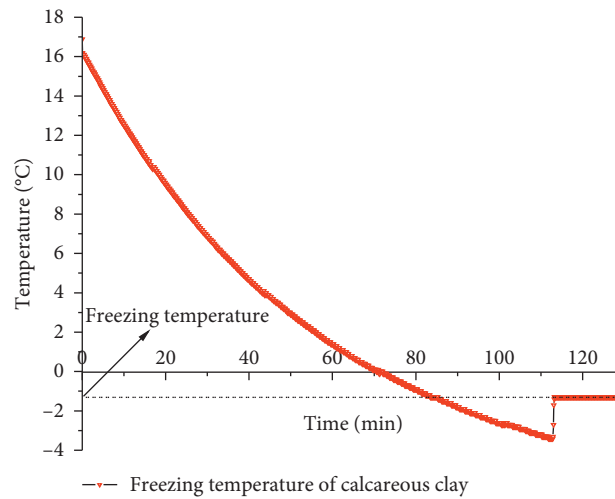


FIGURE 2: Yangcun calcareous clay freezing point experiment.

TABLE 2: Thermodynamic properties of Yangcun frozen soil.

Serial number	Sampling depth (m)	Lithology	Thermal conductivity W/(m·K)	
			Unfrozen soil	Frozen soil
1	406-420	Calcareous clay	0.9304	1.2705
2	260	Fine sand	1.55	1.85
3	\	Sandy clay	1.428	1.952

characteristics. At the same time, the thermal conductivity of the soil was not correlated with the stress.

- (2) The soil was considered to be a continuous medium.
- (3) The initial temperature of the soil was assumed to be an equivalent constant (the first type of boundary conditions), and the length of the frozen pipe was kept constant.

**3.2. Similarity Criterion for the Experiment.** According to the actual working conditions on-site in the Yangcun mine and the existing test conditions in the laboratory, the diameter of the frozen tube that was used in the test was a 6 mm diameter copper tube in order to ensure the geometric similarity constant of the test, and the design of the test system was carried out accordingly. The size of the box that was used in the experiment was  $H = 1.2$  m with a diameter  $D = 1.6$  m. The material used in the test was an undisturbed soil sample from the Yangcun site. The term  $\pi$  and the similarity criterion equation [24] have been listed according to the dimensions. From the similarity criterion, the time similarity ratio was  $C_\tau = C_l^2$  (that is, the time similarity ratio is the square of the geometric similarity ratio), and the temperature similarity ratio was  $C_\theta = 1$  (that is, the temperature at the corresponding points of the model and the prototype was equal), the thermal conductivity similarity ratio was  $C_{\alpha_n} = 1$ , and the thermal conductivity similarity ratio was  $C_{\lambda_n} = 1$ . Since the test model and the prototype used the same number of frozen tubes, the similarity ratio of the frozen tubes' heat dissipation coefficient and the similarity ratio of the number of frozen tubes in the frozen section were all 1, that is,  $C_q = C_N = 1$ . The similarity ratio of the outer diameter of the freezing tubes and the similarity of the freezing depth were  $C_D = C_H = 1$  such that the heat flow  $C_Q = C_l^2$  was similar to the refrigerant flow rate  $v' = C_l \cdot v$  (that is, the flow rate of the refrigerant in the test was  $C_l$ , which was higher than the actual flow rate of the project). The parameters of the layout and the filling parameters of the freezing pipes have been shown in Table 3 and Figure 3, respectively.

**3.3. Layout of the Measuring Points.** The temperature measuring points were arranged in two levels: the main level and the auxiliary level. Each level was arranged at three faces, which were the main face, the common interface, and the boundary's main face; there were 6 measuring points on the main surface of the boundary. The arrangement of the specific measuring points and the distance between each measuring point have been shown in Figure 4.

**3.4. Test Procedure.** The temperature of the test environment was  $13^\circ\text{C}$ , and the box was filled with the original calcareous clay from the Yangcun Mine. The clay was reshaped and filled according to the properties of the original soil. The density of the fill was  $\rho = 2070$  kg/m<sup>3</sup>, the moisture content was 25%, and the height of the fill was  $h = 0.9$  m. A 0.3 m calcareous clay cushion was placed on the bottom of an iron bucket, and then temperature measuring points were located

at intervals of 0.3 m, with a total of upper and lower layers. In order to prevent the temperature measurement line from twisting and deflecting during the filling process, a wooden stick was used to accurately position the temperature measurement points. The test's freezing system used an alcohol tank and a freezing control cabinet. The temperature measurement system that was used consisted of a temperature sensor and a data collector. The test process was divided into the freezing process and the thawing process. The freezing pipes were arranged as shown in Figure 5; the model of the freezing test has been shown in Figure 6.

## 4. Test Results and Analysis

**4.1. Time Effect of the Freezing Temperature Field.** Figures 7(a)–7(c) show the temperature drop at each temperature measurement point on the upper temperature measuring surface at the common principal surface, the common interface, and the principal boundary surface, respectively. As the temperature drop on the upper and lower temperature measurement surfaces was the same, the upper temperature measurement surface was selected for analysis.

The following can be observed from Figure 7(a): (1) The freezing section and the measuring points A3 and A4 on the common main surface (as shown in Figure 4) were located between the inner and middle rows, and the middle and outer rows, respectively. Close to the freezing tube, as the freezing progressed, the cooling capacity was superimposed to supply, and the temperature decreased the fastest; in the positive temperature zone, the temperature drop is basically linear and rapid. In the negative temperature zone, as the free water in the soil gradually freezes, latent heat is released, and the temperature drop rate is reduced; this process released latent heat, and the average temperature drop rate in the first 16 hours was  $2.66^\circ\text{C/h}$  and  $2.61^\circ\text{C/h}$ , respectively. The freezing process of the rock and the soil was accompanied by various thermophysical and chemical reactions and migration of both water and salt. In the stable freezing period, for the unfrozen water, the increase in the salt concentration meant it was extremely difficult to freeze the strong and weakly bound water. In addition, the temperature difference between the soil and the refrigerant decreased; as a result, the potential energy of the heat exchange decreased, and the rate at which the temperature decreased significantly slowed down. After 60 h, the lowest temperature had reached  $-26.1^\circ\text{C}$  and  $-23.2^\circ\text{C}$ , respectively; because the cold transfer path of A1 and A2 was farther than A3 and A4, the rate at which temperature decreased was slow,  $0.79^\circ\text{C/h}$  and  $1.28^\circ\text{C/h}$ , respectively. At the same time, with the continuous supply of cooling capacity, the minimum temperatures of the A1 and A2 measuring points reached  $-21.5^\circ\text{C}$  and  $-22^\circ\text{C}$ , respectively. The measuring point A5 and the measuring point A2 were the symmetrical measuring points of the outer and inner rows of the three rows of tubes, and these were close to the outer and inner rows of tubes, respectively. During the freezing period of the positive temperature zone, the curves of the temperature reduction of the two were basically the same. But entering the negative temperature zone, the



TABLE 3: Yangcun coal mining model and the prototype's geometric parameters (ratio of similitude was 35).

Parameter	Prototype	Model
Outer holes	Circle diameter (m)	29.7
	Number of holes (Pc)	55
	Hole spacing (m)	1.696
Middle row of holes	Circle diameter (m)	23.3
	Number of holes (Pc)	26/26
	Hole spacing (m)	1.407
Inner row of holes	Circle diameter (m)	18.3
	Number of holes (Pc)	26
	Hole spacing (m)	2.206
Frozen wall thickness (m)	10.7	0.306
Alcohol temperature (°C)	-32	-32
Freezing tube's outer diameter (mm)	159	6
Freezing wall's average temperature (°C)	-18	-18

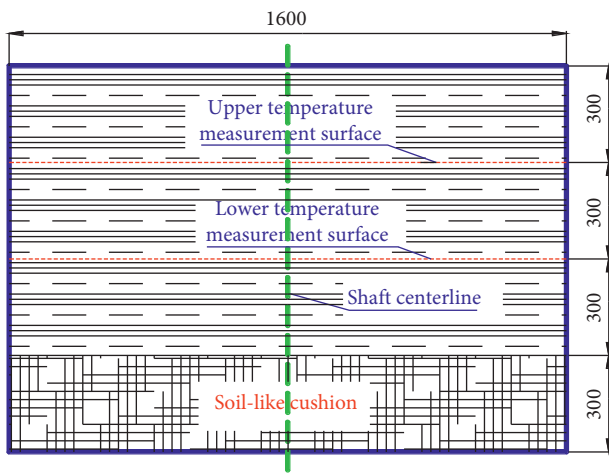


FIGURE 3: The model's fill map (label length unit: mm).

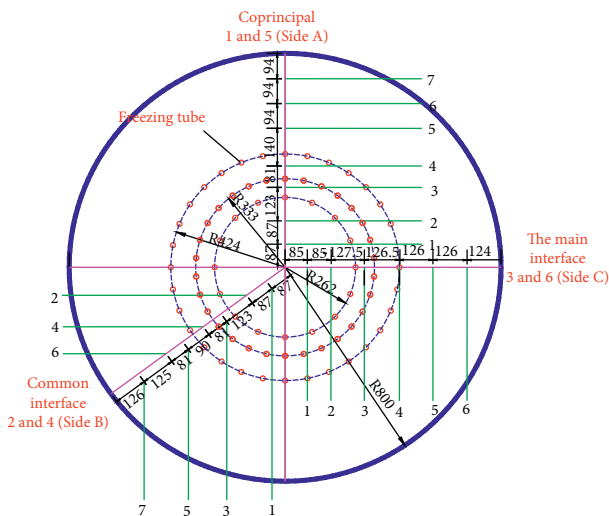


FIGURE 4: Horizontal distribution of the temperature measurement points (label length unit: mm).

temperature drop rate of A2 measuring point is much greater than that of A5 measuring point. The A5, A6, and A7 measuring points had a heat supply for longer time due to

the continuous heat supply, which lasted for 72 h; their minimum temperatures were  $-8.7^{\circ}\text{C}$ ,  $-1.7^{\circ}\text{C}$ , and  $3.3^{\circ}\text{C}$ , respectively. The temperature reductions rates of points A5, A6, and A7 were  $0.61^{\circ}\text{C/h}$ ,  $0.28^{\circ}\text{C/h}$ , and  $0.20^{\circ}\text{C/h}$ , respectively; compared with points A1~A4, the temperature was higher and the temperature reduction was slow, as shown in Table 4. The minimum temperature at each measuring point (B1~B7) for the common interface shown in Figure 7(b) was  $-25.8^{\circ}\text{C}$ ,  $-28.1^{\circ}\text{C}$ ,  $-30.9^{\circ}\text{C}$ ,  $-27.0^{\circ}\text{C}$ ,  $-16.5^{\circ}\text{C}$ ,  $-6.4^{\circ}\text{C}$ , and  $2.8^{\circ}\text{C}$ , respectively, and the temperature reduction rate was  $0.79^{\circ}\text{C/h}$ ,  $0.90^{\circ}\text{C/h}$ ,  $1.86^{\circ}\text{C/h}$ ,  $2.41^{\circ}\text{C/h}$ ,  $0.67^{\circ}\text{C/h}$ ,  $0.30^{\circ}\text{C/h}$ , and  $0.21^{\circ}\text{C/h}$ , respectively. Figure 7(c) shows the measurement points on the main surface (C1~C6); their minimum temperature was  $-25.8^{\circ}\text{C}$ ,  $-27.7^{\circ}\text{C}$ ,  $-29.9^{\circ}\text{C}$ ,  $-21.2^{\circ}\text{C}$ ,  $-6.5^{\circ}\text{C}$ , and  $1.0^{\circ}\text{C}$ , respectively, and their temperature reduction rates were  $0.80^{\circ}\text{C/h}$ ,  $0.88^{\circ}\text{C/h}$ ,  $1.37^{\circ}\text{C/h}$ ,  $1.75^{\circ}\text{C/h}$ ,  $0.30^{\circ}\text{C/h}$ , and  $0.19^{\circ}\text{C/h}$ , respectively. From the above analysis, it can be seen that the development trends of the freezing temperature field of the three sections were roughly the same. The temperature drop of the common main surface in the area of the freezing wall (measurement points 3 and 4 in each section) was the fastest, and the temperature drop of the common interface was the second fastest. The temperature reduction rate of the No. 1 and No. 2 measuring points for the three cross sections was basically the same, indicating that the freezing wall intersected the area of the three rows of tubes, and the freezing circular front faced the center of the well's bore; there was also stable convective heat transfer in the direction of the line. (2) The reheating section: the hysteresis characteristic of the melting process of the frozen soil [25–28]. In the reheating section, a slow reheating period of nearly 135 h could be seen, while the freezing section of the curve was due to the long time interval of the test. The cooling rate was too fast to show an obvious phase change or supercooling phenomena in the figure. The water in the soil at the three cross sections A1~A4, B1~B4, and C1~C4 melted from the solid phase (ice) turning into the liquid phase (water). The surrounding soil absorbed a lot of heat, and the soil at measuring points 6 and 7 showed no freezing or obvious freezing phenomenon, and the two measuring points of the three sections showed no slow temperature phase change recovery period.



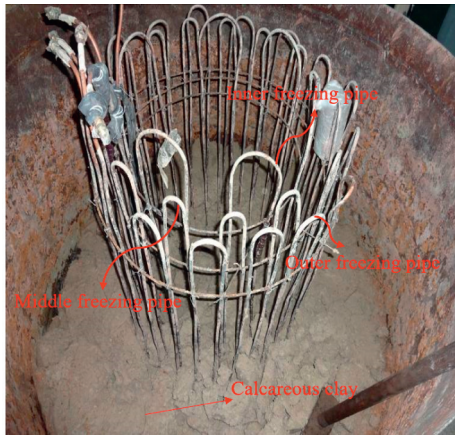


FIGURE 5: Arrangement of the freezing pipes in the experiment.

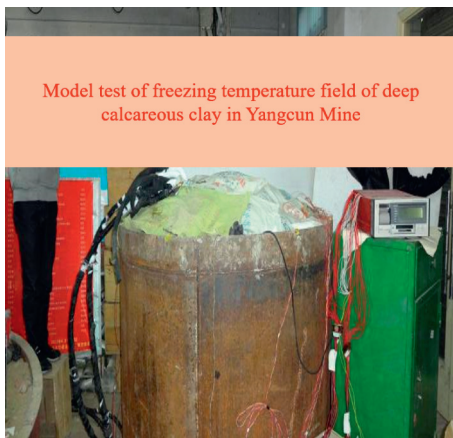


FIGURE 6: The model of the freezing experiment.

#### 4.2. Spatial Effect of the Freezing Temperature Field.

During the test, the spatial distribution of the temperature at the freezing point and the return temperature section of each temperature measuring point on the common principal surface, the common interface, and the principal boundary surface at the specified times have been shown in Figure 8.

Figures 8(a)–8(c) show the spatial distribution of the freezing temperature of the three cross sections of the freezing section at each temperature measurement point for a certain time interval. It can be seen from the figure that the distributions of the temperature values of the measuring points for the three cross sections were an irregular saddle shape. Taking the coprinciple as an example, the temperature drops at the measuring points A3 and A4 are basically synchronized within 16 hours before the freezing period. At the same time, the temperature difference between the two points is within  $0.9^{\circ}\text{C}$ , while the maximum temperature difference between points A3 and A4 between 16 h and 64 h was  $3^{\circ}\text{C}$ ; this was due to the continuous supply of heat from the original ground temperature; therefore, the outer row of pipes needs to provide more cooling to ensure that the outer edge of the frozen wall expands outward.

In the initial stage of the freezing (between 0 to 2 h), the temperature changes of the A1 and A2 measuring points

located in the excavation wasteland and the A5–A7 measuring points located outside the outer row of pipes were not obvious; this was because, in theory, the formation of a frozen front occurred first. On the “circle” locus of the row of pipes, when the temperature of the area reached a certain value, a sufficient temperature potential could form with the adjacent soil body in order to drive the cooling capacity to move inwards and outwards; the number of inner pipes was small, and the distance between the holes was large. At certain times in the positive temperature zone, the temperature of the outside measurement point 5 was lower than those of measurement points 1 and 2 that were located in the excavation barren path (Figures 8(a) and 8(b)). During the entire freezing period, the temperature difference between the A2 and A3, A4 and A5 measuring points was the most obvious. The temperature gradients between the A2 and A3, A4 and A5 measuring points were  $0.14^{\circ}\text{C}/\text{mm}$  and  $0.11^{\circ}\text{C}/\text{mm}$  at 2 h; at both 8 h and 16 h, the values were  $0.17^{\circ}\text{C}/\text{mm}$  and  $0.15^{\circ}\text{C}/\text{mm}$ , respectively. This indicates that there was stable cooling transfer between the points A2 and A3, and A4 and A5 within 2 to 16 hours, and the temperature difference between the A2 and A3, A4 and A5 measurement points during the subsequent freezing period gradually decreased. For example, the gradient between points A2 and A3 was  $0.097^{\circ}\text{C}/\text{mm}$  at 32 h,  $0.047^{\circ}\text{C}/\text{mm}$  at 48 h, and  $0.018^{\circ}\text{C}/\text{mm}$  at 64 h, indicating that the temperature difference decreased, the potential energy was small, and the cooling capacity transfer was slow. The values at the measuring points A5, A6, and A7 on the outside mirrored the measuring points A1 and A2. The difference was due to the existence of the original ground temperature and the better “cold” effect inside; this caused the temperature drop hysteresis of the measuring point A5 in the negative temperature zone and the measuring point A2, and its temperature is higher than A2. The results of the analysis of the spatial distribution of the common interface and the main surface were similar; the difference was that there was no freezing tube in the section where the common interface was located, and there were only a few internal frozen tubes with large spacing. Therefore, the temperature of the B4 measuring point in the first 16 h was lower than that of point B3. The “cold insulation” factor meant that the temperature of B3 was low and the main surface was the section where the inner row and the middle row were connected; the temperature difference was greater here due to the different cooling capacity. The temperature gradient between points C3 and C4 at 2 h and 5 h was  $0.034^{\circ}\text{C}/\text{mm}$ ; at 8 h, it was  $0.063^{\circ}\text{C}/\text{mm}$ ; and in the later period, it was basically stable at  $0.073^{\circ}\text{C}/\text{mm}$ .

Figures 8(d)–8(f) show the temperature distribution of each temperature measuring point during the return temperature period for the three sections. As mentioned above, during the freezing and reheating phases, the calcareous clay underwent a phase change around  $-3.8^{\circ}\text{C}$ , and the rate of the temperature decrease or reheating changed significantly. The three times of 128 h, 156 h, and 240 h were phase transition slow temperature recovery periods, and the temperature measurement points on the three cross sections corresponded to the temperature difference at these three times.

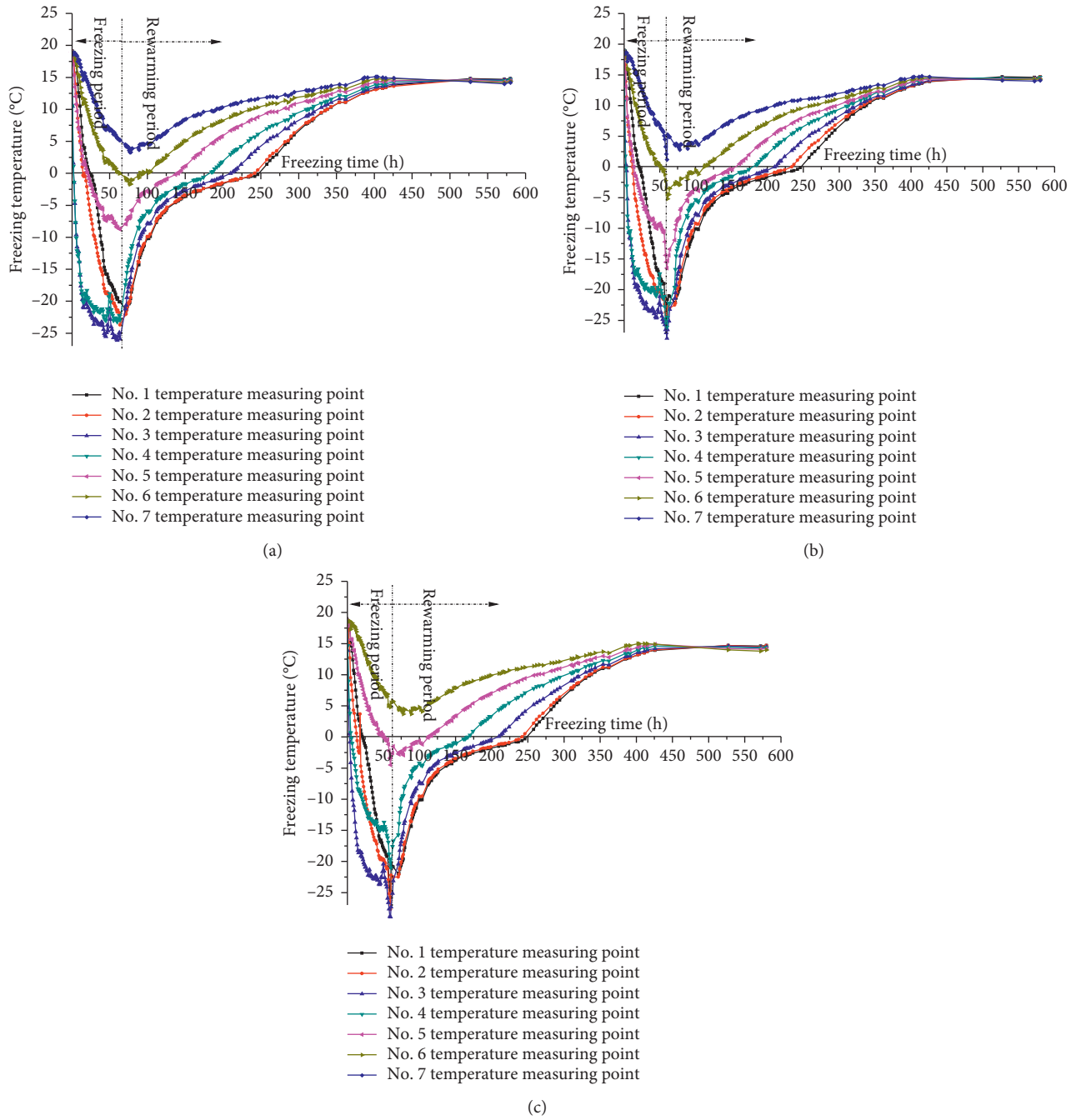


FIGURE 7: The three interface temperatures for the freezing point against time. (a) Coprincipal (side A). (b) Common interface (side B). (c) The main interface (side C).

TABLE 4: The minimum temperature of each measuring point of the three sections and the rate of temperature drop in the active freezing period.

Measuring point	1		2		3		4		5		6		7	
	$T_{min}$ (°C)	V (°C/h)	$T_{min}$ (°C)	V (°C/h)	$T_{min}$ (°C)	V (°C/h)	$T_{min}$ (°C)	V (°C/h)	$T_{min}$ (°C)	V (°C/h)	$T_{min}$ (°C)	V (°C/h)	$T_{min}$ (°C)	V (°C/h)
Co-principal	-21.5	0.79	-23.7	0.86	-26.1	2.66	-23.2	2.61	-8.7	0.61	-1.7	0.28	3.3	0.20
Common interface	-25.8	0.79	-28.1	0.90	-30.9	1.86	-27.0	2.41	-16.5	0.67	-6.4	0.30	2.8	0.21
The main interface	-25.8	0.80	-27.7	0.88	-29.9	1.37	-21.2	1.75	-6.5	0.30	1.0	0.19	—	—

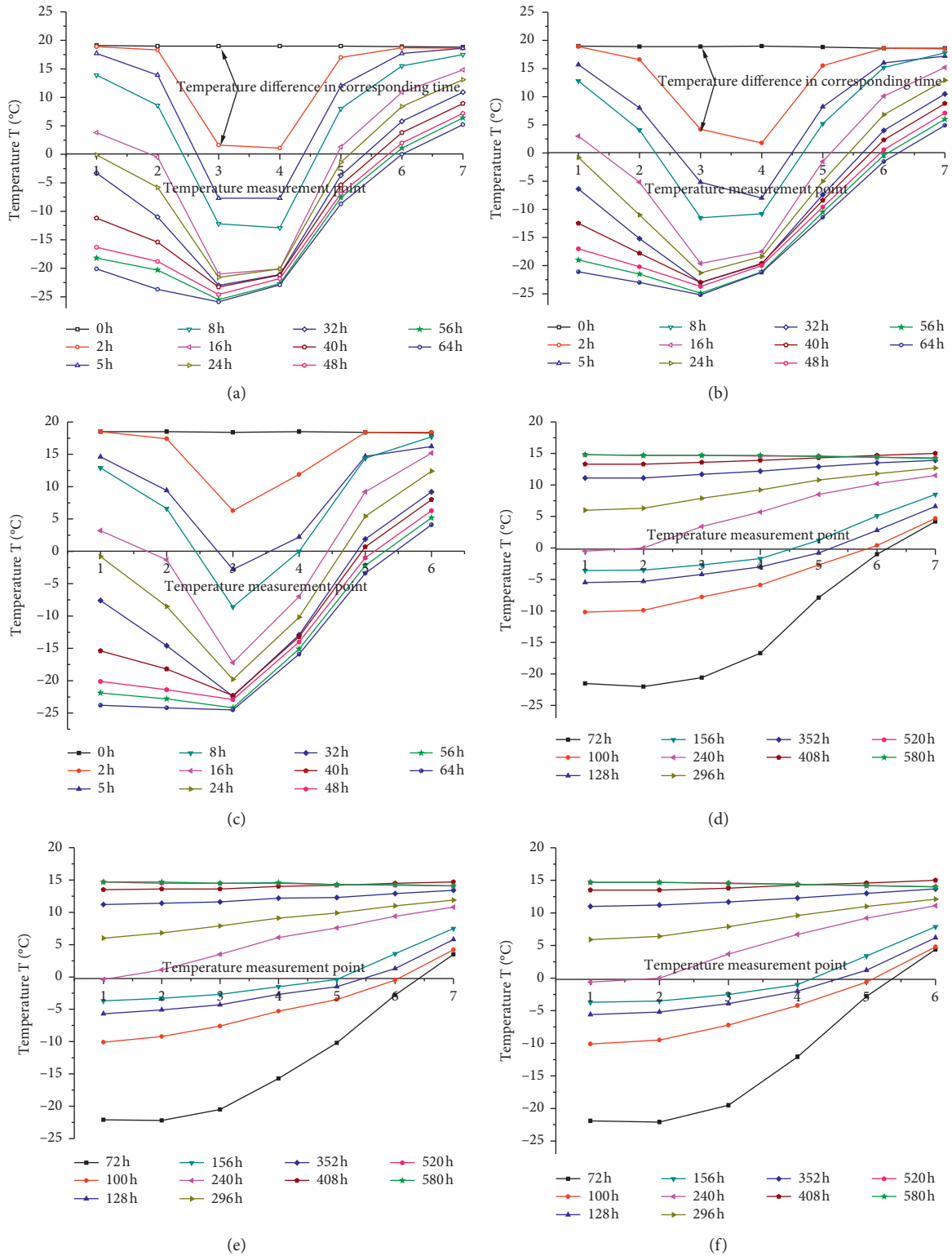


FIGURE 8: Temperature measurement and temperature cycle distribution map for the three interfaces. (a) Coprincipal. (b) Common interface. (c) The main interface. (d) Coprincipal temperature return. (e) Common interface temperature return. (f) Temperature return of the main interface.

At measuring point 7 of each section (the main boundary surface No. 6), due to the natural melting of the frozen soil, the phase body needed to absorb a large amount of heat from the environment through the outer wall of the model. At the measuring point of section No. 7 (the main boundary surface No. 6), the soil was not frozen, but a large “relative” change in the temperature could also be seen between 156 h and 240 h. Since the freezing point of calcareous clay is  $-1.3^{\circ}\text{C}$ , but its supercooling temperature can be as low as  $-3.8^{\circ}\text{C}$ , the temperature at points 2, 3, 4, and 5 was higher than its supercooling temperature. This is the node where the frozen wall completely melted; the total duration of the wall melting process was 210 h, and the freeze-thaw time ratio was 1 : 1.91.

*4.3. The Overall Development Law of the Frozen Wall.* By combining the test data of the three main test sections and the auxiliary test plane, based on the kriging difference method in the Surfer software, the freezing temperature field formed by the calcareous clay has been shown as an isotherm diagram, as shown in Figure 9. It should be noted that the measurement points of this test were divided into upper and lower planes, and each plane had three main sections in which the measurement points were arranged. The temperature field was constructed based on the temperature measurement points of the three main sections. The temperature could not be described completely due to the characteristics of the field distribution. The degree of compaction of the soil filling and the slight difference in the water content in all directions were ignored, and according to the layout characteristics of the frozen pipes as shown in Figure 4, a number of sections were selected at a certain angle from the semicircular plane and the corresponding measured section data were assigned to the plot. Among these, there were 2 main surfaces and 1 main boundary surface, and the total main surface was divided into  $11.25^{\circ}$  sections, for a total of 16 sections. Figures 9(a)–9(f) show the contour maps of the temperature field at the corresponding moments in the test freezing section. It can be seen from the figure that the temperature of the entire fill area was basically the same at the beginning of the freezing process; that is, at 0 h, there was a temperature difference of  $0.5^{\circ}\text{C}$ . As the freezing progressed, the temperature rapidly decreased, and the temperature of the central area of the three rows of pipes dropped to  $0^{\circ}\text{C}$  around the 2 h mark, as shown in Figures 9(b) and 9(c); it can be seen that the  $-1.3^{\circ}\text{C}$  line started to intersect the position of the middle row of pipes at about 2.5 h. It should be pointed out that according to the similarity theory, the flow rate of the refrigerant in the test was  $C_f$ , which was higher than the actual flow rate of the project, but the current test equipment could not achieve such a high flow rate; therefore in this test process, as much as possible, the refrigerant flow rate was guaranteed to freeze the tube and a constant temperature of  $-32^{\circ}\text{C}$  was maintained for cooling, reducing the test’s errors. Although the test could not achieve a similar flow rate, according to the relationship between the heat flow similarity ratio, the refrigerant flow rate’s similarity ratio, and the geometrical similarity ratio, it can be inferred from this that the

intersection time of the  $2.5\text{ h } 0^{\circ}\text{C}$  line in the test was equivalent to the actual engineering freezing time of 127.6 days. It can be seen that due to the complex mineral composition and hydrophilicity of the calcareous clay, the soil layer has “difficult” freezing characteristics. In the 5 h freezing period, a freezing wall was formed in the “true” sense in the area of the three rows of freezing tubes. After 72 h, the temperature in the freezing wall had been reduced to the minimum, and the freezing wall reached an average temperature of  $-18^{\circ}\text{C}$ , and the freezing wall reached its design thickness. Because of the lack of understanding of the “difficult” freezing characteristics of calcareous clay, when the actual wellbore was excavated, the wellbore was unloaded, and the frozen wall of the layer where the calcareous clay was located had not reached its design strength; as a result, the pressure exerted by the surrounding soil on the frozen wall was nearly below the radial in situ stress of 5 MPa (calcareous clay buried depth of 420 m), and the frozen calcareous clay layer was prone to rheological deformation; as a result of this, a freezing pipe was sheared and destroyed.

Figures 9(g)–9(i) show the contour maps of the temperature field of the frozen calcareous clay during the recuperation period. It can be seen from the figure that the melting cycle of frozen calcareous clay is longer, which lasts 508 h to warm to room temperature, which is nearly 7 times of the freezing period. Figures 9(g) and 9(h) show the contour line cloud maps corresponding to 100 h and 192 h during the slow-phase recovery period. The temperature rise in the frozen soil area during this 92 h period did not exceed  $9^{\circ}\text{C}$ ; this has shown that the thermal conductivity of the calcareous clay was poor, the heat absorption from the outside was slow, and it had a long melting time.

The average temperature and thickness of the frozen wall could be calculated from the temperature measurement data, as shown in Figure 10. It can be seen from the figure that due to the difference in the temperature development of the three sections, the effective thickness of the frozen wall is formed on the main surface of the 5 h time boundary, and the frozen wall formed a circle at that time; the development trend of the frozen wall at the three interfaces was consistent. A faster time meant that the temperature dropped more quickly in the initial stage of freezing; it tended to be gentle in the later stage, and the thickness of the frozen wall changed slowly. The development speed of the thickness of the freezing wall is the fastest on the coprincipal (side A), the second is the main interface (side C), and the common interface (side B) is the slowest. The development of the thickness of the frozen wall could be divided into two stages: the first stage was between 5 hours–16 hours; this period of time was the fastest stage of the development of the thickness of the frozen wall where the development speed was 17.9 mm/h. The second stage is from 16 hours to 70 hours, and the development speed of the thickness of the frozen wall is 1.96 mm/h. The average development speed of the thickness of the frozen wall was 5.4 mm/h; the actual development speed of the thickness of the frozen wall was 0.0123 m/d in the early stage, and 0.0014 m/d in the later stage, which was significantly less than the thickness of the frozen wall of the mine shaft and other



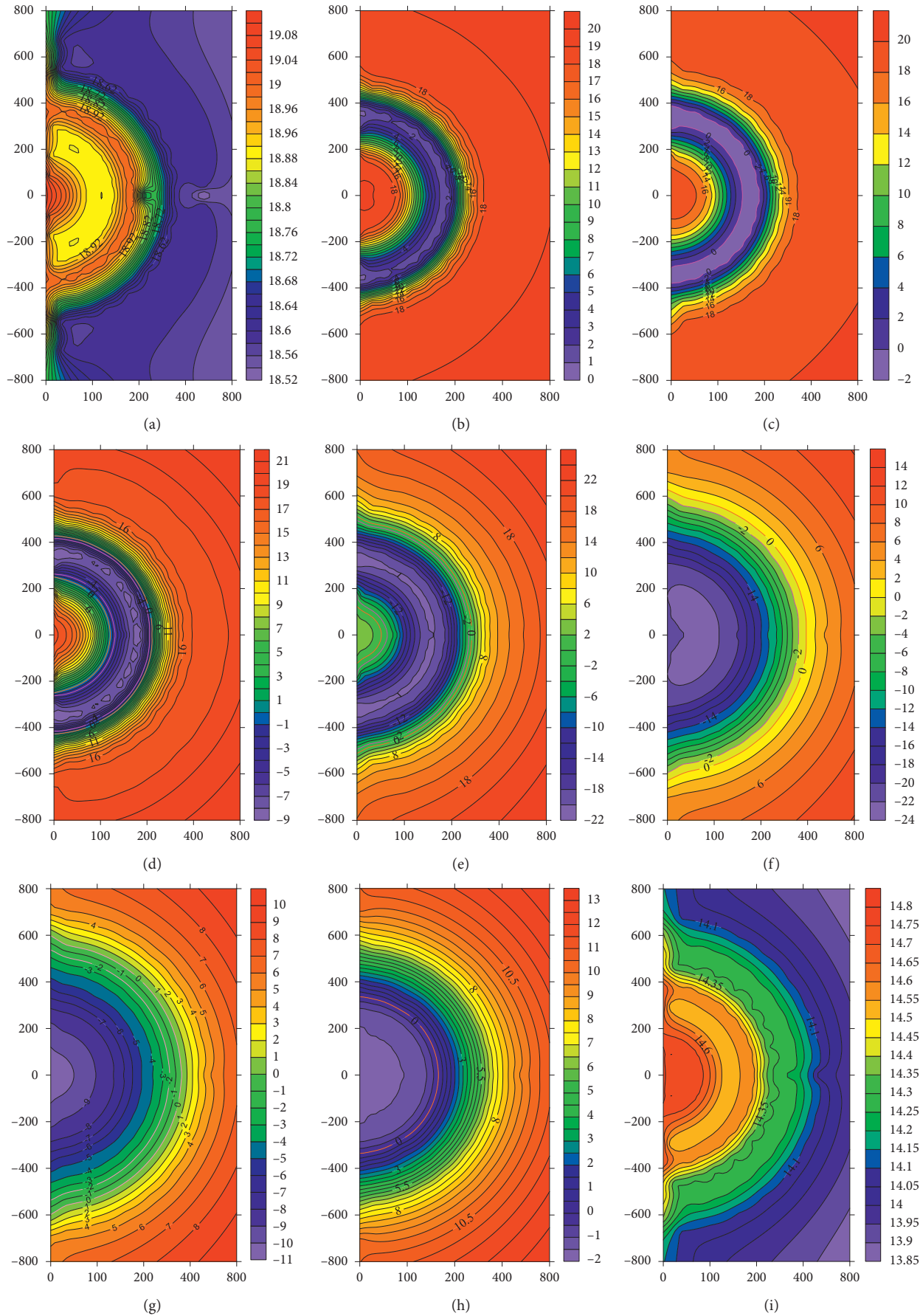


FIGURE 9: Cloud diagram of the freezing temperature field distribution of the three rows of pipes (small scale unit: °C). (a) 0 h. (b) 2 h. (c) 3 h. (d) 5 h. (e) 19 h. (f) 72 h. (g) 100 h. (h) 192 h. (i) 580 h.

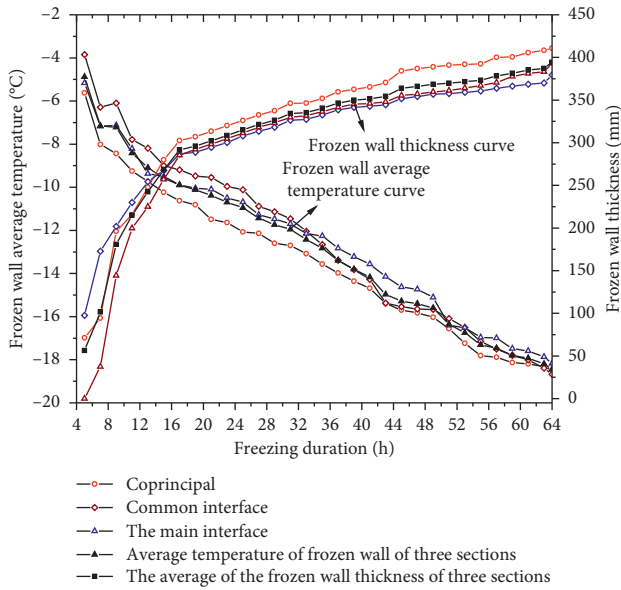


FIGURE 10: The relationship between the average temperature and the thickness of the frozen wall at the three interfaces and their average values with the freezing time.

mines in Zhangji’s North District which was 0.0515 m/d in the early stage and 0.02 m/d in the later stage [29]. The complex chemical composition of calcareous clay (the high content of montmorillonite in calcareous clay has been preliminarily explored; this can account for more than 40% of the total minerals in the clay, along with illite. The total amount reaches about 90%, see later research for details) and the hydrophilicity of the surface of the soil particles means they are attached to a thicker layer of bound water, coupled with frequent ion exchange and salt migration during the complex thermophysical and chemical interactions during freezing; this is the cause of the “difficult” freezing characteristics of calcareous clay. Therefore the frozen wall will develop slowly and the creep deformation is large; as a result, accidents such as breaks in the freezing pipes occur.

### 5. Conclusion

The “difficult” freezing characteristics of calcareous clay have caused many occurrences of pipe breakage during construction. In this paper, based on macromodel test research, an in-depth quantitative study on the temperature field expansion and development law of a freezing wall under the freezing conditions of calcareous clay has been carried out; from this, the following conclusions could be drawn:

- (1) The complex mineral composition of calcareous clay, as well as ion exchange, dissolution, precipitation, and the other chemical reactions that occur in the water and soil system, means that the freezing point of calcareous clay is  $-1.3^{\circ}\text{C}$ , its supercooling temperature can be as low as  $-3.8^{\circ}\text{C}$ , and its thermal conductivity is low. Therefore, under the same cooling conditions, the layer of clay has poor thermal conductivity, and it freezes more slowly than the

other layers of soil. The development speed of the frozen soil will then be slow, its strength will be low, and, as a result of this, the strength of the frozen wall will be uneven in the vertical direction, and it will be easy for shear to occur at the interface of an adjacent cracked soil layer.

- (2) The initial intersection time ( $-1.3^{\circ}\text{C}$ ) of the freezing temperature field in the experiment was 2.5 h, which was equivalent to 87.5 days in the actual project; this is nearly 23 days longer than the designed intersection time of the clay layer for the freezing method of the mine’s construction which was 65 days. In the current construction freezing method, the freezing time of the mine was nearly doubled to 45 days. A lack of understanding of the “difficult” freezing characteristics of the calcareous clay meant it was regarded as being the same as an ordinary clay layer. Excavation and unloading of a soil layer that has not completely formed an “effective” freezing wall that can resist high ground pressure can result in shear failure occurring in a freezing tube.
- (3) For the three interfaces that were investigated in the experiment, the temperature of each temperature measurement point in the temperature field that was formed by the calcareous clay layer changed at the same time in an irregular saddle shape, and the time ratio between the formation of the frozen wall and complete melting was 1 : 1.91. The growth rate of the thickness of the frozen wall of the calcareous clay layer from 5 hours to 16 hours was 17.9 mm/h, and the growth rate from 16 hours to 70 hours was 1.96 mm/h; this corresponded to the actual development rate of the thickness of the frozen wall of 0.0123 m/d and 0.0014 m/d, which was significantly smaller than the development speed of the freezing of the general sandy clay layer wall thickness of 0.0515 m/d in the early stage and 0.02 m/d in the later stage. This will cause the thickness of the freezing wall and its strength to be unevenly distributed in the vertical direction. This is the main cause of pipe fractures and therefore should be taken seriously in the construction of an actual project.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was funded by the National Natural Science Foundation of China (51374010 and 51474004), the Natural Science Foundation of Anhui Province (2008085ME165), and the Key Academic Funding Project for Top-Notch Talents (gxbjZD2016045).



## References

- [1] H. Ren, X. Hu, C. Jin, and J. Zhang, "Study on freezing effect and operation of profiled enhancing freezing-tube in freeze-sealing pipe roof," *Tunnel Construction*, vol. 35, no. 11, pp. 1169–1175, 2015.
- [2] H. U. Xiangdong, L. I. Xinyi, and Y. Wu, "etc. Effect of waterproofing in Gongbei Tunnel by freeze-sealing pipe roof method with field temperature data," *Chinese Journal of Geotechnical Engineering*, vol. 41, no. 12, pp. 2207–2214, 2019.
- [3] B. Zhang, W. Yang, and B. Wang, "Elastoplastic design theory for ultra-deep frozen wall considering large deformation features," *Chinese Journal of Geotechnical Engineering*, vol. 41, no. 7, pp. 1288–1295, 2019.
- [4] W. Yang, D. U. Zibo, and Y. Zhijiang, "Plastic design theory of frozen soil wall based on interaction between frozen soil wall and surrounding rock," *Chinese Journal of Geotechnical Engineering*, vol. 35, no. 10, pp. 1857–1862, 2013.
- [5] W. Yang and Y. Zhijiang, "Elastic-plastic design theory of frozen soil wall based on interaction between frozen wall and surrounding rock," *Chinese Journal of Geotechnical Engineering*, vol. 35, no. 1, pp. 175–180, 2013.
- [6] H. U. Xiangdong and S. H. U. Chang, "Stress field analysis of functionally graded material frozen soil wall in double-row-pipe shaft freezing," *Engineering Mechanics*, vol. 31, no. 1, pp. 145–153, 2014.
- [7] X. Hu, S. She, and R. Yu, "Average temperature calculation for straight single-row-piped frozen soil wall," *Sciences in Cold and Arid Regions*, vol. 3, no. 2, pp. 124–131, 2011.
- [8] H. U. Xiangdong, F. Zhao, S. H. E. Siyuan, and S. H. U. Chang, "Equivalent parabolic arch method of average temperature calculation for straight double-row-pipe frozen soil wall," *Journal of China Coal Society*, vol. 37, no. 1, pp. 28–32, 2012.
- [9] R. Chuanxin, B. Wang, and P. Shilong, "Study on mechanical characteristics of functionally graded material frozen soil wall basing on Druker-Prager strength criterion," *Journal of Guangxi University*, vol. 37, no. 40, p. 1344, 2015.
- [10] B. Wang, R. Chuanxin, and H. Cheng, "Theoretical analysis of the double-row-pipe frozen soil wall basing on drucker-prager strength criterion," *Science Technology and Engineering*, vol. 16, no. 25, pp. 44–50, 2016.
- [11] J. Lin, H. Cheng, H.-B. Cai, B. Tang, and G.-Y. Cao, "Effect of seepage velocity on formation of shaft frozen wall in loose aquifer," *Advances in Materials Science and Engineering*, vol. 2018, pp. 1–11, 2018.
- [12] H. B. Cai and R. Chuanxin, "Nonlinear analysis of freezing temperature field considering latent heat of phase change," *Low Temperature Architecture Technology*, vol. 37, no. 2, pp. 43–45, 2009.
- [13] Z. He, R. Zu, K. Zhang et al., "Laboratory experiment of the influence of aeolian sand accumulation on permafrost temperature," *Journal of Glaciology and Geocryology*, vol. 37, no. 1, pp. 156–161, 2015.
- [14] L. I. Haipeng, W. Yang, H. Jiahui, S. Lei, and T. Zhang, "Model test of frost heaving pressure during formation of clay freezing wall with two-cycle freezing-tube," *Journal of Glaciology and Geocryology*, vol. 33, no. 4, pp. 801–806, 2011.
- [15] T. Wang and G. Zhou, "Analysis of temperature field around a single freezing pipe considering variability of soil parameters," *Journal of China Society*, vol. 39, no. 6, pp. 1063–1069, 2014.
- [16] L. I. Jinhua, W. Yansen, L. I. Dahai, and C. Zhang, "Model tests on stress and deformation of freezing-pipes during multi-circle freezing before excavation," *Chinese Journal of Geotechnical Engineering*, vol. 33, no. 7, pp. 1072–1077, 2011.
- [17] Y. U. Chuhou and P. Yang, "A discussion on the cause of freezing pipe breaking and the technology to avoid the breaking," *Journal of Anhui University of Science and Technology (Natural Science)*, vol. 10, pp. 40–49, 1987.
- [18] S. Zhang and L. I. Gongzhou, *Special Well Drilling Technology for Deep Alluvium in Yongxia Mining Area*, Coal Industry Press, Beijing, China, 2003.
- [19] P. Yang, "Disposal of freezing pipe fracture and study of freezing temperature detection," *Coal Technology of Northeast China*, vol. 4, no. 4, pp. 8–10, 1996.
- [20] M. Ma, H. Cheng, and R. Chuanxin, "Freezing construction technology for main shaft of Yangcun Coal Mine in deep and thick calcareous clay layer," *Coal Engineering*, vol. 49, no. 9, pp. 14–18, 2017.
- [21] M. Ma, H. Cheng, and R. Chuanxin, "A numerical analysis of freezing pressure in frozen soils around a deep shaft based on the modified Nishihara model," *Rock and Soil Mechanics*, vol. 36, no. 10, pp. 3015–3022, 2015.
- [22] Y. Liu, H. Liang, and T. Zhaohui, "Inverse modeling of geochemical behavior of  $\text{Ca}^{2+}$  in landslide water-soil interaction system near the Three Gorges Reservoir," *Hydrogeology & Engineering Geology*, vol. 39, no. 2, pp. 106–110, 2012.
- [23] Y. Tan, B. Yu, X. Hu, and X. Liu, "Prediction model for thermal conductivity of unsaturated soil," *Chinese Journal of Geotechnical Engineering*, vol. 35, no. S1, pp. 129–133, 2013.
- [24] J. Yang, *Similarity Theory and Structural Model Test*, Wuhan University of Technology Press, Wuhan, China, 2005.
- [25] L. Tan, W. E. I. Changfu, T. Huihui, Z. Jiazuo, and W. Houzhen, "Experimental study of unfrozen water content of frozen soils by low-field nuclear magnetic resonance," *Rock and Soil Mechanics*, vol. 35, no. 6, pp. 53–59, 2015.
- [26] E. J. A. Spans and J. M. Baker, "The soil freezing characteristic: its measurement and similarity to the soil moisture characteristic," *Soil Science Society of America Journal*, vol. 60, no. 1, pp. 13–19, 1996.
- [27] Y. A. N. Changgen, T. Wang, and H. Jia, "Influence of the unfrozen water content on the shear strength of unsaturated silt during freezing and thawing," *Chinese Journal of Rock Mechanics and Engineering*, vol. 38, no. 6, pp. 1252–1260, 2018.
- [28] M. Ma, Y. Huang, and G. Cao, "Study on mechanical behavior of jurassic frozen sandstone in western China based on NMR porosity," *Journal of Chemistry*, vol. 38, 2020.
- [29] K. Shen and R. Chuanxin, "Finite element analysis on freezing temperature field of Chagannaer Mine auxiliary shaft," *Journal of Guangxi University*, vol. 36, no. 4, pp. 647–652, 2011.

## Research Article

# A Computing Model for Quantifying the Value of Structural Health Monitoring Information in Bridge Engineering

**Baoquan Cheng** <sup>1,2</sup> **Lijie Wang**,<sup>3</sup> **Jianling Huang**,<sup>1</sup> **Xu Shi**,<sup>4</sup> **Xiaodong Hu**,<sup>1</sup>  
and **Huihua Chen** <sup>1</sup>

<sup>1</sup>School of Civil Engineering, Central South University, Changsha, Hunan 410083, China

<sup>2</sup>Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China

<sup>3</sup>Department of Architecture, Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei 050041, China

<sup>4</sup>College of Optoelectronic Engineering, Chongqing University, Chongqing 400044, China

Correspondence should be addressed to Huihua Chen; [chh24770@163.com](mailto:chh24770@163.com)

Received 2 August 2020; Revised 17 September 2020; Accepted 23 September 2020; Published 23 October 2020

Academic Editor: Shaohui Wang

Copyright © 2020 Baoquan Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Structural health monitoring system can provide valuable information for improving decision-making process in maintenance and management of bridges. However, managers usually lack understanding of value of structural health monitoring information. This paper developed a computing model for quantifying the value of structural health monitoring information based on Bayesian theory. Then, the model was demonstrated and validated using a simple case and the key factors (i.e., system accuracy, repair cost, prior probability of structural failure, and manager's behavior pattern) influencing the value of structural health monitoring information were identified and discussed. Findings from this study help to answer the question of whether a structural health monitoring system should be installed and run, thus enriching the knowledge body of structural health monitoring.

## 1. Introduction

With the sustained economic development, the traffic volume particularly the overloading vehicles increased rapidly. This seriously threatens the structural safety of bridges particularly those constructed as early as the 19th~the mid-20th century whose extended service lives have caused dangerous accumulation of structural damages in long-term use [1–3]. In fact, it is reported that structural accidents such as sudden bridge breakages occurred frequently in recent years. Once the structural failure occurred, there would be huge social-economic losses and human casualties. Therefore, how to ensure structural safety of bridges especially at the background of such a huge traffic volume has become the great challenge nowadays. In fact, if in-time judgements and alarms are made before structural failure, managers can take corresponding urgent measures to avoid future damages and accidents [4]. This requires managers regularly measure and

assess the degree of cumulative structural damage [5]. Structural health monitoring (SHM) therefore emerged as times need and caught attention from both scholars and industrial stakeholders.

SHM can be defined as the strategies and process for identifying and characterizing structural damages. A completed SHM system usually consists of 6 modules including sensory system, data acquisition and transmission system, data processing and control system, structural health evaluation system, structural health data management system, and inspection and maintenance system [6–8]. It can provide real-time information about structural conditions of bridges (environmental loads and status, operation loads, bridge features, and structural responses) by integrating various state-of-art technologies such as sensory technology, deep learning, big data, and machine vision [9, 10]. These valuable information can provide decision-making references for managers in maintenance and management of bridges.

Nowadays, SHM has been conducted on more and more famous bridges such as Jiangyin Bridge (Jiangsu, China), Tsing Ma Bridge (Hong Kong), and Faroe Bridge (Denmark) in the global [11]. However, it is reported that many bridge managers in the real life are still reluctant to invest on SHM. Sometimes, even though the structural health monitoring information is available, they still make decisions based on their common sense and prior experience instead of actions suggested by the SHM system. The embarrassing situation actually results from a principle which goes beyond the scope of the bridge managers. Firstly, there exist many uncertainties affecting the quality of structural health monitoring information such as system accuracy and stability, environmental noise, and errors in data process. Bridge managers therefore usually cannot completely believe in the SHM system. They will weigh the actions suggested by SHM combining with their prior experience and common sense. At the same time, bridge managers are very concerned about results caused by wrong actions. They will consider all potential impacts of possible actions before making a decision, which also drives their preference to experience-based instead of SHM-suggested action [12]. In summary, lack of understanding of value of structural health monitoring information decrease managers' interests in SHM in the real life. To promote and improve application of SHM in future, the question of whether a SHM system should be installed and run must be addressed. This means that a rational model to evaluate the value of structural health monitoring information is urgently needed.

Current studies on SHM mainly focus on optimization of hardware system development, sensory system placement, and damage identification algorithms [13–18]. However, limited attention was paid to the value of structural health monitoring information [19–21]. To address this research gap, this paper aims to develop a computing model for quantitatively assessing the value of structural monitoring information based on Bayesian theory. Through scenario simulation and sensitivity analysis based on the developed model, the main influencing factor of value of structural health monitoring information is identified and discussed. This study can deepen managers' understanding of value of SHM and thus eliminate their doubt on reliability of SHM.

## 2. Method and Materials

**2.1. Basic Assumptions.** Many uncertainties are usually involved in the decision-making process. To make the best decision, the manager should collect related information of these uncertainties to decrease and eliminate them. These information that helps to reduce decision-making risks is of certain value to some extent. However, the process of information collection has to consume resources including time, manpower, and money. This means managers often face a difficult question, that is, if it is worthwhile to collect information. To answer this question, a standard should be developed, which means managers should convert all investment on information collection into monetary value and compare it with the value of information (VOI). Therefore, the quantitative assessment of VOI is of great significance for decision making.

To develop the computing model for quantifying VOI of SHM, some basic assumptions were made in advance as follows:

*Assumption 1.* A bridge must be at one of  $n$  possible structural states (e.g., undamaged, slight damage, serious damage, and structural failure) labeled as  $S_1, S_2, \dots, S_n$ , of which possibilities are assumed to be mutually exclusive and exhaustive.

*Assumption 2.* The bridge manager has  $m$  action choices (e.g., do nothing, repair, and rebuild), and he has to choose one. We labeled these actions as  $a_1, a_2, \dots, a_m$ .

*Assumption 3.* The manager's decision will cause a cost that depends on the bridge's actual structural state.  $c_{ij}$  indicates the expected cost for action  $a_i$  when the bridge is at state  $S_j$ . It should be noted that this cost includes many aspects such as financial losses, delayed transportation time, environmental impacts, and accident casualty caused by structural failure. It is assumed that all costs can be expressed by a monetary value.

*Assumption 4.* The manager is assumed to be a rational agent. This means the manager will always choose the most economic action with the minimum cost.

Based on above assumptions, a general model for quantifying the value of structural health monitoring information can be developed, which consists of 3 steps, i.e., (1) cost-benefit analysis, (2) prior decision analysis, and (3) preposterior decision analysis.

**2.2. Cost-Benefit Analysis.** The expected cost for every action when the bridge is in every structural state  $c_{ij}$  including both direct cost and indirect cost is evaluated in this step. This is the basis of accurate quantification of value of structural health monitoring information. Indirect cost (e.g., delayed transportation time, environmental impacts, and accident casualty caused by structural failure) can be converted into monetary value through value transfer approach [22].

**2.3. Prior Decision Analysis.** If there is not an SHM system, information about structural conditions of the bridge is not available. Therefore, the manager can only estimate the expected cost  $C_i$  of action  $a_i$  based on his prior experience and common sense by the following equation:

$$C_i = \sum_{j=1}^n c_{ij}P(S_j). \quad (1)$$

In which,  $P(S_j)$  represents the prior probability that the bridge is at structural state  $S_j$ .

As a rational agent, the bridge manager will choose the most economic action  $a_{\text{opt}}$  that minimizes the expected cost as follows:

$$a_{\text{opt}} = \arg \min C_i = \arg \min \sum_{j=1}^n c_{ij}P(S_j). \quad (2)$$

The corresponding expected management cost without SHM can thus be calculated as follows:

$$C = \min C_i = \min \sum_{j=1}^n c_{ij} P(S_j). \quad (3)$$

**2.4. Preposterior Decision Analysis.** When an SHM system is installed and run, the manager can obtain a set of observation  $y$  such as deflection, stress, acceleration, and strain. The information can modify the bridge manager's knowledge of structural state  $S_j$  from  $P(S_j)$  to  $P(S_j | y)$ , reducing the uncertainties and risks in the decision-making process. Equation (1) can therefore be rewritten as

$$C'_i = \sum_{j=1}^n c_{ij} P(S_j | y). \quad (4)$$

In which, the mark  $'$  is a reminder that it is the expected cost posterior to the observations of  $y$ .

The posterior probability  $P(S_j | y)$  can be calculated using the Bayesian rule:

$$P(S_j | y) = \frac{p(y | s_j) P(s_j)}{p(y)}. \quad (5)$$

In which, the probability of observations  $p(y)$  can be estimated using the law of total probability:

$$p(y) = \sum_{j=1}^n p(y | S_j) P(S_j). \quad (6)$$

Thus, with structural health monitoring information  $y$ , equations (2) and (3) can be rewritten as equations (7) and (8):

$$a'_{\text{opt}} = \arg \min C'_i = \arg \min \sum_{j=1}^n c_{ij} P(S_j | y), \quad (7)$$

$$C'(y) = \min C'_i(y) = \min \sum_{j=1}^n c_{ij} P(S_j | y). \quad (8)$$

Equation (8) represents the minimum expected cost with an individual observation set  $y$ . If the monitoring continues, the corresponding cost can be calculated with the probability distribution of  $y$  using

$$C' = \int_{\Omega_y} C'(y) p(y) dy = \int_{\Omega_y} \min \left\{ \sum_{j=1}^n c_{ij} P(S_j | y) \right\} p(y) dy. \quad (9)$$

In which,  $\Omega_y$  represents the domain of  $y$ .

Substituting equation (5) into equation (9), equation (9) can be rewritten as

$$C' = \int_{\Omega_y} \min \left\{ \sum_{j=1}^n c_{ij} p(y | S_j) P(S_j) \right\} dy. \quad (10)$$

Then, the VOI of an independent SHM can be calculated using

$$\begin{aligned} \text{VOI} &= C - C' \\ &= \min \sum_{j=1}^n c_{ij} P(S_j) - \int_{\Omega_y} \min \left\{ \sum_{j=1}^n c_{ij} p(y | S_j) P(S_j) \right\} dy. \end{aligned} \quad (11)$$

### 3. Scenario Simulation

The developed computing model is demonstrated and validated by a simple example. For simplification, it is assumed that there are only two possible structural states for the bridge, i.e., undamaged ( $U$ ) and damaged ( $D$ ). The bridge manager has only two action choices: do nothing ( $DN$ ) and repair ( $R$ ). It will cost nothing if the manager does nothing when the bridge is undamaged whereas making this decision when structure fails will result in a huge cost  $C_F$ . In addition, choosing to repair the bridge will cause a cost  $C_R$ . We assumed that  $C_R$  is independent of the actual structural state of the bridge and  $C_R \ll C_F$ . The results of cost-benefit analysis are summarized in Table 1.

The bridge manager will estimate the expected cost of each action. The reparation cost is identically equal to  $C_R$  no matter what the actual structural state is. The cost of doing nothing then depends on the manager's knowledge of structural failure.

When there is not an SHM system, he can only estimate the cost of doing nothing  $C_{DN}$  based on his prior experience and common sense through the following equation:

$$C_{DN} = C_F \cdot P(D). \quad (12)$$

In which,  $P(D)$  represents the prior probability of the structural failure.

Based on rational agent assumption, the manager will only repair the bridge when  $C_{DN} > C_R$ ; otherwise, he will accept risks and do nothing. Thus, the expected cost  $C$  without SHM can be got through

$$C = \min(C_R, C_F \cdot P(D)). \quad (13)$$

Next, the impacts of SHM on the manager's decision are investigated. For simplification, it is assumed that the SHM has only two outcomes (i.e., Silence ( $S$ ) and Warning ( $W$ )). The outcome of SHM is probabilistically related to the actual structural state of bridges. If the observation exceeds the threshold, a warning will be given by the SHM system. Otherwise, it keeps silence. To model the system accuracy, two parameters are introduced: the probability of a false warning when a warning is given but the bridge is undamaged  $P_{FW} = P(W | U)$ , and the probability of false silence when the SHM system keeps silence but the structure fails  $P_{FS} = P(S | D)$ . Thus, the probabilities of SHM outcomes at each structural state are summarized in Table 2.

Generally,  $P_{FW} \neq P_{FS}$ . When a higher threshold is set,  $P_{FW}$  tends to go up while  $P_{FS}$  tends to decrease. The lower the  $P_{FW}$  and  $P_{FS}$  are, the more reliable the SHM is. More reliable structural health monitoring information has higher

TABLE 1: Temperature and wildlife count in the three areas covered by the study.

	U	D
DN	0	$C_F$
R	$C_R$	$C_R$

TABLE 2: Conditional probability of SHM outcomes and structural states.

	U	D
S	$1 - P_{FW} = P(S U)$	$P_{FS} = P(S D)$
W	$P_{FW} = P(W U)$	$1 - P_{FS} = P(W D)$

corresponding VOI. Only the situation in which  $P_{FW}$  and  $P_{FS}$  are both below 0.5 is taken into consideration in this study because SHM with too high rate has no values. Based on equation (6), the probability of each SHM outcome can be estimated through the following equations:

$$P_W = P(W|U) \cdot P(U) + P(W|D) \cdot P(D), \quad (14)$$

$$P_S = P(S|U) \cdot P(U) + P(S|D) \cdot P(D). \quad (15)$$

Then, equation (5) can be rewritten as equations (16) and (17) for assessing the probability of damage:

$$P(D|W) = \frac{P(W|D) \cdot P(D)}{P(W)}, \quad (16)$$

$$P(D|S) = \frac{P(S|D) \cdot P(D)}{P(S)}. \quad (17)$$

Thus, the structural health monitoring information enables the manager to update their knowledge of structural state and estimate the expected cost through the following equations again:

$$C'(W) = \min(C_R, P(D|W) \cdot C_F), \quad (18)$$

$$C'(S) = \min(C_R, P(D|S) \cdot C_F). \quad (19)$$

Therefore, the expected cost with SHM can be calculated through:

$$C' = C'(W) \cdot P(W) + C'(S) \cdot P(S). \quad (20)$$

Then, following the same criterion in equation (11), the VOI can therefore be evaluated through:

$$\begin{aligned} \text{VOI} = C - C' = \min(C_R, C_F \cdot P(D)) \\ - (C'(W) \cdot P(W) + C'(S) \cdot P(S)). \end{aligned} \quad (21)$$

It can be found that the value of structural health monitoring information depends on the specific values of cost of each action  $C_F$  and  $C_R$ , the prior probability of structural damage  $P(D)$ , and the error rates of the SHM system  $P_{FW}$  and  $P_{FS}$ .

## 4. Results and Discussion

**4.1. VOI and System Accuracy.** Figure 1 describes the contours of VOI (K\$) in which  $C_R = 10\text{K\$}$ ,  $C_F = 1000\text{K\$}$ , and  $P(D) = 0.5\%$ . When the SHM can always offer completely precise information defined as perfect information (i.e.,  $P_{FW} = P_{FS} = 0$ ), the manager can always judge the actual structural state of the bridge accurately. Thus, he will only repair the bridge when it is damaged; otherwise, he will do nothing. In fact, structural failure will be prevented in either cases when perfect information is available. In this case, the value of perfect information  $\text{VOPI} = C - C' = 5000 - 50 = \$4950$ . It is also the upper limit of the VOI. No matter how state-of-art SHM technologies are adopted, the corresponding VOI can never exceed this boundary. The ratio of actual VOI and VOPI can be used to present the efficiency of the SHM system. The other extreme condition is  $P_{FW} = P_{FS} = 0.5$ . This means the SHM outcome is absolutely independent of the actual structural state. Under this condition,  $C = C' = \$5\text{K}$  and corresponding  $\text{VOI} = 0$ . It proves that the SHM system with too high error rate has no values. It is also the lower limit of VOI. Between these two limit values, the VOI goes down with the decreasing system accuracy. In addition, the VOI is more sensitive to  $P_{FW}$  than  $P_{FS}$ . This can be validated by the fact that the VOI in S2 ( $P_{FW} = 0.1$ ,  $P_{FS} = 0.2$ ) is 5000 higher than that in S1 ( $P_{FW} = 0.2$ ,  $P_{FS} = 0.1$ ).

**4.2. VOI and Repairation Cost.** Figure 2 describes impacts of repairation cost  $CR$  on the VOI of the SHM system in which cost of structural failure  $C_F = 1000\text{K\$}$  and the prior probability of damage  $P(D) = 0.5\%$ . We assume that  $P_{FW} = P_{FS}$ . In each lime, the VOI increases with the increasing repairation cost  $CR$  until reaching the peak at  $CR = 5\text{K\$}$ . Then, the VOI turns to decrease until  $\text{VOI} = 0$  if repairation cost  $CR$  continues to increase. This is consistent with our common sense. When the repairation is very cheap, the manager will always repair the bridge to eliminate risks of structural failure no matter what structural health monitoring information is provided. When the repairation cost goes up, the expected cost with SHM  $C'$  goes down, leading to growth in VOI. At the point of  $CR = 5\text{K\$}$ , the expected cost with SHM  $C'$  decreases to the bottom and the VOI reaches the peak. After that, with the continuously incensing repairation cost, the bridge manager is likely to accept risks of structural failure, causing decreases in VOI. After the VOI reduces to 0, the manager will never repair the bridge due to the high cost. By comparing three limes, it can be found that the higher the SHM system accuracy is, the higher the VOI of SHM is. A higher SHM system accuracy also expands the effective range of SHM towards repairation cost.

**4.3. VOI and Prior Probability of Damage.** Figure 3 describes impacts of the prior probability of damage  $P(D)$  on the VOI of the SHM system in which cost of structural failure  $C_F = 1000\text{K\$}$  and repairation cost  $C_R = 5\text{K\$}$ . We still assume that  $P_{FW} = P_{FS}$ . In each lime, the VOI increases with the prior

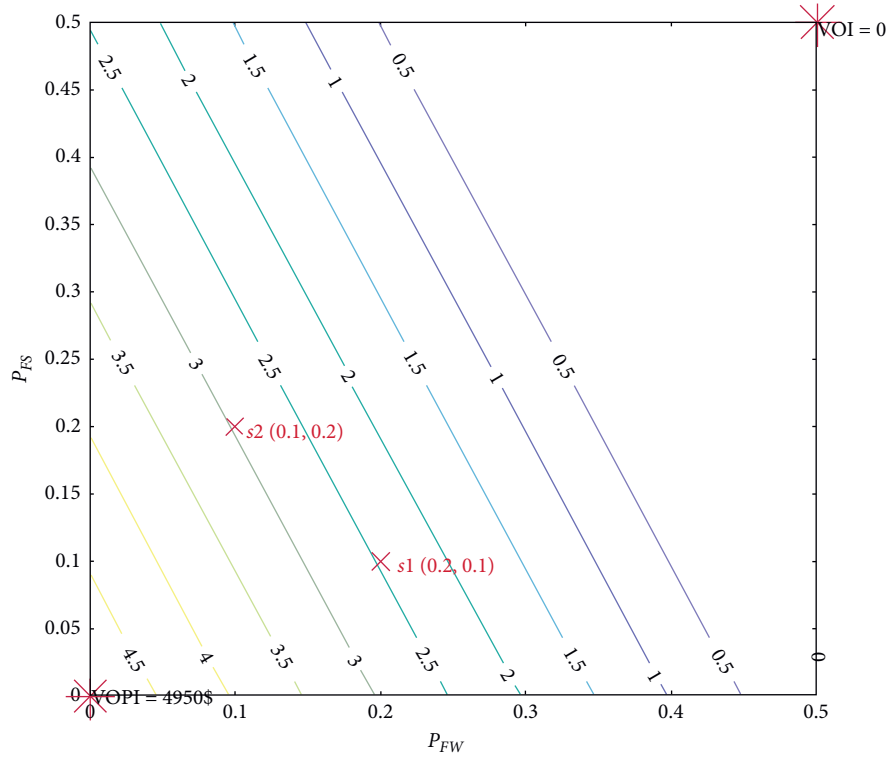


FIGURE 1: Contours of VOI (K\$) as a function of the failure probability of the SHM system  $P_{FW}$  and  $P_{FS}$  when  $C_R = 10K\$, C_F = 1000K\$, and  $P(D) = 0.5\%$ .$

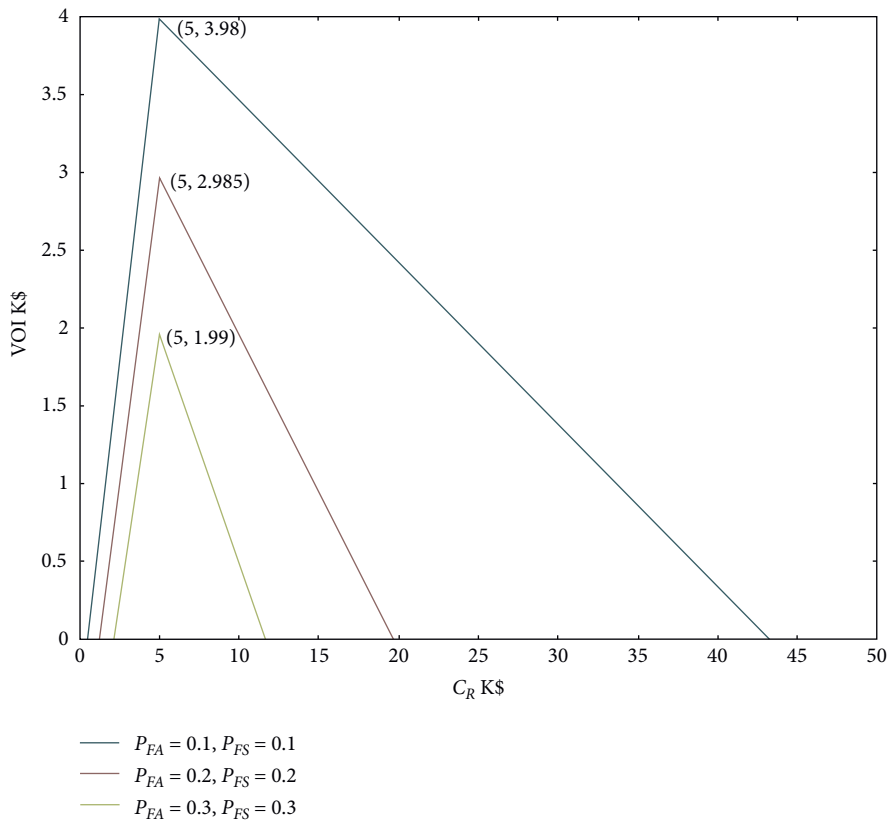


FIGURE 2: VOI as a function of the repair cost  $C_R$  when  $C_F = 1000K\$, P(D) = 0.5\%, and  $P_{FW} = P_{FS}$ .$



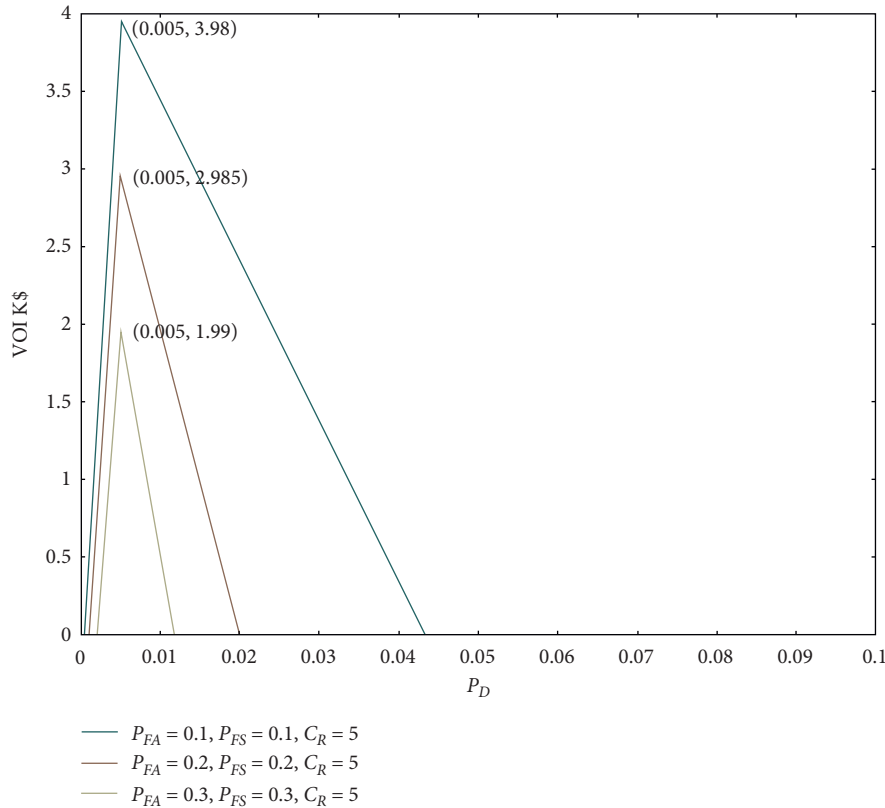


FIGURE 3: VOI as a function of the prior probability of damage  $P(D)$  when  $C_F = 1000K\$$ ,  $P(D) = 0.5\%$ , and  $P_{FW} = P_{FS}$ .

probability of damage  $P(D)$  until reaching the peak at  $P(D) = 0.005$ . Then, the VOI turns to decrease until  $VOI = 0$  if the prior probability of damage continues to go up. This can be explained by the fact that when the probability of structural failure is too low, the bridge manager is really confident in structural safety and he will accept risks of structural failure and do nothing. SHM information helps little. When the prior probability of damage increases, uncertainties and risks in decision-making process as well increase. VOI helps to decrease these uncertainties and risks, and the VOI grows correspondingly. At the point of  $P(D) = 0.005$ , the expected cost with SHM  $C'$  decreases to the bottom and the VOI reaches the peak. After that, with the continuously increasing prior probability of damage, the bridge manager is likely to repair the bridge. In fact, uncertainties of the bridge structural state decrease, causing decreases in VOI. After the VOI reduces to 0, the manager will always repair the bridge to avoid high loss due to structural failure. By comparing three lines, it can be found that the higher the SHM system accuracy is, the higher the VOI of SHM is. A higher SHM system accuracy also expands the effective range of SHM towards prior probability of damage.

**4.4. VOI and Manager's Behavior.** The bridge manager's behavior pattern also influences the actual value of structural health monitoring information. It is assumed that the manager is a completely rational agent. This means in the

developed model, the manager will always pursue the minimum expected economic cost in the decision-making process. This is not completely consistent with the reality. Two extreme situations are presented in this section. Firstly, if the manager is completely a risk-oriented agent, he will never accept the risk of structural failure. In this situation, he will select to repair the bridge. The other extreme situation occurs when the manager is a completely adventure agent. In this situation, he will do nothing for the largest interests. In above both extreme situations, the structural health monitoring information is not applied in the decision-making process and creates no value. Therefore,  $VOI = 0$ . This result proved that the manager's behavior is also an important influence factor of VOI of SHM.

## 5. Conclusions

This study was to quantify the value of structural health monitoring information. This aim has been achieved by developing a computing model based on Bayesian theory including three steps: (1) cost-benefit analysis, (2) prior decision analysis, and (3) preposterior decision analysis. The developed model was demonstrated and validated using a simple example. Simulation results show that VOI of SHM depends on system accuracy, the specific values of cost of each action, and the prior probability of damage. However, this study only considers a very ideal case with only two possible structural states and two feasible actions. However, the situation is much more complex in the real life. The value

of structural health monitoring information in a more complicated situation that is closer to the real world should be assessed in the further studies.

### Data Availability

All datasets generated for this study are included in the article.

### Disclosure

This paper is based on the master dissertation of Baoquan Cheng at HK PolyU.

### Conflicts of Interest

All authors declare no conflicts of interest.

### Authors' Contributions

All authors contributed equally to this paper.

### Acknowledgments

The authors wish to acknowledge the financial support from the 13th Five-Year National Key Research and Development Plan, China (no. 2017YFB1201204), China Innovation Funding, National Natural Science Foundation of China (no. 71942006), and the Natural Science Foundation of Hunan Province, China (no. 2019JJ40407).

### References

- [1] I. E. Harik, A. M. Shaaban, H. Gesund, G. Y. S. Valli, and S. T. Wang, "United States bridge failures, 1951–1988," *Journal of Performance of Constructed Facilities*, vol. 4, no. 4, pp. 272–277, 1990.
- [2] L. Deng, W. Wang, and Y. Yu, "State-of-the-art review on the causes and mechanisms of bridge collapse," *Journal of Performance of Constructed Facilities*, vol. 16, no. 2, Article ID 04015005, 2016.
- [3] B. H. Ji and Z. Q. Fu, "Analysis of Chinese bridge collapse accident causes in recent years," *China Civil Engineering Journal*, vol. 43, no. S1, pp. 495–498, 2010, in Chinese.
- [4] N.-J. Yau, M.-K. Tsai, H.-L. Wang, D.-M. Hung, C.-S. Chen, and W.-K. Hsu, "Improving bridge collapse detection and on-site emergency alarms: a case study in Taiwan," *Safety Science*, vol. 70, pp. 133–142, 2014.
- [5] S.-P. Chang, J. Yee, and J. Lee, "Necessity of the bridge health monitoring system to mitigate natural and man-made disasters," *Structure and Infrastructure Engineering*, vol. 5, no. 3, pp. 173–197, 2009.
- [6] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 303–315, 2006.
- [7] W.-X. Ren and G. De Roeck, "Structural damage identification using modal data. I: simulation verification," *Journal of Structural Engineering*, vol. 128, no. 1, pp. 87–95, 2002.
- [8] Y. Zhang, Zhu, W. Zhang, and H. Liu, "Analysis of deformation characteristics and stability mechanisms of typical landslide mass based on the field monitoring in the Three Gorges Reservoir, China," *Journal of Earth System Science*, vol. 128, no. 1, 2019.
- [9] L. Hui and O. Jinping, "Structural health monitoring: from sensing technology stepping to health diagnosis," *Procedia Engineering*, vol. 14, pp. 753–760, 2011.
- [10] D. Zhu and Y. Dong, "Experimental and 3D numerical investigation of solitary wave forces on coastal bridges," *Ocean Engineering*, vol. 209, 2020.
- [11] S. J. Beard, A. Kumar, X. Qing et al., "Practical issues in real-world implementation of structural health monitoring systems," in *Smart Structures and Materials 2005: Industrial and Commercial Applications of Smart Structures Technologies*, vol. 5762, pp. 196–204, International Society for Optics and Photonics, Bellingham, WA, USA, 2005.
- [12] D. Zonta, B. Glisic, and S. Adriaenssens, "Value of information: impact of monitoring on decision-making," *Structural Control and Health Monitoring*, vol. 21, no. 7, pp. 1043–1056, 2014.
- [13] Y. Li, Y. Dong, D. M. Frangopol, and D. Gautam, "Long-term resilience and loss assessment of highway bridges under multiple natural hazards," *Structure and Infrastructure Engineering*, vol. 16, no. 4, pp. 626–641, 2020.
- [14] J. Q. Cheng, "Optimal sensor placement for bridge structure based on improved effective independence," *Journal of Vibration, Measurement and Diagnosis*, vol. 32, pp. 812–816, 2012, In Chinese.
- [15] C. Papadimitriou, "Optimal sensor placement methodology for parametric identification of structural systems," *Journal of Sound and Vibration*, vol. 278, no. 4–5, pp. 923–947, 2004.
- [16] Y. Zhang, Z. Zhang, S. Xue, R. Wang, and M. Xiao, "Stability analysis of a typical landslide mass in the Three Gorges Reservoir under varying reservoir water levels," *Environmental Earth Sciences*, vol. 79, no. 1, 2020.
- [17] B. Cheng, J. Li, V. W. Tam, M. Yang, and D. Chen, "A BIM-LCA approach for estimating the greenhouse gas emissions of large-scale public buildings: a case study," *Sustainability*, vol. 12, no. 2, 2020.
- [18] Y. Liu, D. Bi, and Z. X. Li, "Optimal placement of accelerometers in long cable-stayed bridges based on genetic algorithm," *Journal of Southeast University (Natural Science Edition)*, vol. 39, no. 4, pp. 825–829, 2009, In Chinese.
- [19] H. Wang, Y. Wu, M. Wei, L. Wang, and B. Cheng, "Hysteretic behavior of geopolymer concrete with active confinement subjected to monotonic and cyclic axial compression: an experimental study," *Materials*, vol. 13, 2020.
- [20] Z. Wang, L. Li, and Y. Zhang, "Bond-slip model considering freeze-thaw damage effect of concrete and its application," *Engineering Structures*, vol. 201, 2019.
- [21] Z. Wang, L. Li, and Y. Zhang, "Reinforcement model considering slip effect," *Engineering Structures*, vol. 198, 2019.
- [22] B. Cheng, "Numerical investigation into the value of structural health monitoring information for bridge structure," Master thesis, The Hongkong Polytechnic University, Hung Hom, Hong Kong, 2019.

## Research Article

# Approximate Controllability for a Kind of Fractional Neutral Differential Equations with Damping

Jun Du , Dongling Cui, Yeguo Sun, and Jin Xu

Department of Applied Mathematics, Huainan Normal University, Huainan 232038, China

Correspondence should be addressed to Jun Du; [djwlm@163.com](mailto:djwlm@163.com)

Received 20 August 2020; Revised 5 September 2020; Accepted 28 September 2020; Published 21 October 2020

Academic Editor: M. Javaid

Copyright © 2020 Jun Du et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper gains several meaningful results on the mild solutions and approximate controllability for a kind of fractional neutral differential equations with damping (FNDED) and order belonging to  $[1, 2]$  in Banach spaces. At first, a new expression for the mild solutions of FNDED via the  $(p, q)$ -regularized operator family and the technique of Laplace transform is acquired. Then, we consider the approximate controllability of FNDED by means of the approximate sequence method, and simultaneously, some applicable sufficient conditions are obtained.

## 1. Introduction

The primary description of the fractional-order derivative was proposed by Riemann and Liouville toward the end of the nineteenth century, but the notion of the arbitrary derivative and integral which generalized the classical integer-order derivative and integral was presented by Leibniz and Liouville in 1695. However, until the late 1960s when many phenomena on physics, engineering technology, and economics were more accurately described by fractional differential equations (FDEs), scientists began to show great interest on fractional order calculus. For example, they are widely adopted for nonlinear oscillations of earthquakes and in the fluid-dynamic traffic model. In practice, FDEs are deemed to optimize the traditional differential equation model. About the elementary theory of fractional differential and evolution systems, one can refer to Podlubny [1], Kilbas et al. [2], Zhou [3, 4], and [5–10] and the references cited therein.

In the particles' realistic movement, the resistance to motion is unavoidable, so it is suitable to add the damping character in mathematical models and control systems. Recently, a great deal of meaningful conclusions for the mathematical models with damping influence have been presented by the researchers [11–17].

As we all know, the controllability exerts a momentous effect on control theory and engineering technology. It lies in the fact that it is bound up with quadratic optimal control, observer design, and pole assignment. For this reason, the controllability has been actively investigated by many investigators, and an impressive progress has been made in recent years [7–12, 14–20]. Controllability of the deterministic systems in infinite dimensional spaces has been broadly investigated. In some results of the controllability for systems described by fractional differential models, the fixed-point and the approximate sequence method are felicitously used. Nonetheless, as demonstrated by Triggiani [21], for many parabolic partial differential systems, the conditions of complete controllability are very finite. The research to approximate controllability is more proper for the practical systems than to complete controllability. For the past few years, as regards differential dynamical systems in Banach spaces, several results are achieved about the approximate controllability [8, 19]. However, as far as we know, the approximate controllability of the fractional neutral differential equations with damping and order belonging to  $[1, 2]$  is still relatively infrequent, so it is more interesting and necessary to study it.

## 2. Preliminaries and Notations

Inspired by the aforementioned analysis, the approximate controllability for a kind of fractional neutral differential equations with damping of order belonging to [1, 2] in the Banach space is studied in our work. We acquire several

$$\begin{cases} {}^C D_t^p x'(t) + q {}^C D_t^p [x(t) - h(t, x_t)] = Ax(t) + f(t, x_t) + Bu(t), & t \in (0, b], \\ x(t) = \varphi(t), & t \in [-\tau, 0], \\ x'(0) = x_1, \end{cases} \quad (1)$$

where  $p: 0 \leq p \leq 1$ ,  ${}^C D_t^p$  denotes the Caputo fractional derivative,  $q(q \geq 0)$ , and the linear densely unbounded closed operator  $A: D(A) \subseteq E \rightarrow E$  is the  $(p, q)$ -regularized family defined on the Banach space  $E$ . Here, the state  $x(\cdot)$  is evaluated in Banach space  $E$ . Let  $U$  be a Banach space of admissible control functions. The variable  $u(\cdot)$  takes a value in  $L_2([0, b]; U)$ ,  $B: U \rightarrow E$ , which is linear and bounded. In addition,  $f: [0, b] \times C([-\tau, 0]; E) \rightarrow E$  is nonlinear, which will be explained in detail later;  $x_t \in L^2([-\tau, 0], E)$  and is denoted by  $x_t(\vartheta) = \{x(t + \vartheta) | -\tau \leq \vartheta \leq 0\}$ ;  $\varphi = \{\varphi(\vartheta) | -\tau \leq \vartheta \leq 0\} \in L^2([-\tau, 0], E)$ . The relevant linear neutral system of FNDED (1) is

$$\begin{cases} {}^C D_t^p y'(t) + q {}^C D_t^p [y(t) - h(t, y_t)] = Ay(t) + Bv(t), & t \in (0, b], \\ x(t) = \varphi(t), & t \in [-\tau, 0], \\ x'(0) = x_1. \end{cases} \quad (2)$$

The structure of this paper is given as follows. In Section 2, we review several fundamental concepts and provide a new form of the mild solution for FNDED (1). Then, in Sections 3 and 4, we acquire several meaningful results for the existence and uniqueness of the mild solution and, furthermore, the approximate controllability for FNDED (1).

Let  $E$  be a Banach space with norm  $\|\cdot\|_E$  and  $L(E)$  be the space of all linear bounded operators on  $E$ . Let

$${}^C D_{0^+}^\alpha z(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\nu)^{-\alpha} z'(\nu) d\nu = I_{0^+}^{1-\alpha} z'(t), \quad t > 0, 0 < \alpha < 1. \quad (6)$$

**Lemma 1.** (see [2]). Let  $\alpha \in (n-1, n]$ . If  $z(\cdot) \in AC^n([a, b], E)$  or  $z(\cdot) \in C^n([a, b], E)$ , then the Caputo derivative

$$I_{0^+}^\alpha D_{0^+}^\alpha z(t) = z(t) - \sum_{i=0}^{n-1} \frac{z^{(i)}(0)}{i!} t^i, \quad t > 0. \quad (7)$$

**Definition 3.** (see [15, 16]). Let  $q \geq 0$  and  $0 \leq p \leq 1$ .  $A$  is a linear closed operator, and its domain  $D(A)$  is in Banach space  $E$ . We say that  $A$  is the generator of a  $(p, q)$ -regularized operator family  $\{S_p(t)\}_{t \geq 0} \subset L(E)$  if the three formulas are established in the following:

sufficient conditions to pledge the approximate controllability of the FNDED via the contraction mapping theory and approximate sequence method. The FNDED and order in [1, 2] is debated as follows:

$L_m([0, b]; E)$ ,  $1 \leq m < \infty$ , be the space of  $E$ -valued Bochner integrable function  $f: [0, b] \rightarrow E$ , and

$$\|f\|_{L_m} = \left( \int_0^b \|f\|_E^m dt \right)^{1/m}. \quad (3)$$

In the following, we recall some definitions to be adopted in the entire work.

**Definition 1.** (see [1, 2]). The Riemann–Liouville fractional integral of a function  $z$  and order  $\alpha$  and from lower limit 0 can be denoted by

$$I_{0^+}^\alpha z(t) = \frac{1}{\Gamma(\alpha)} \int_0^t \frac{z(s)}{(t-s)^{1-\alpha}} ds, \quad t > 0, \alpha > 0. \quad (4)$$

**Definition 2.** (see [1, 2]). The Caputo derivative of order  $\alpha$  for a function  $z \in C([0, \infty), R)$  can be defined as

$${}^C D_{0^+}^\alpha z(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-\nu)^{-\alpha} z'(\nu) d\nu, \quad t > 0, 0 < \alpha < 1. \quad (5)$$

**Remark 1** (see [1, 2]). If  $z \in C([0, \infty), R)$ , then

- (a)  $S_p(t)$  is strongly continuous on  $R_+$ , and  $S_p(0) = I$ .
- (b)  $S_p(t)D(A) \subset D(A)$  and  $AS_p(t)x = S_p(t)Ax$  for all  $x \in D(A)$ ,  $t \geq 0$ .
- (c) For every  $x \in D(A)$ , the following equation holds:

$$S_p(t)x = e^{-qt} + \frac{1}{\Gamma(p)} \int_0^t \int_0^\nu \mu^{(p-1)} e^{q(\mu-\nu)} AS_p(t-\nu)x d\mu d\nu, \quad t \geq 0, \quad (8)$$

where  $S_p(t)$  is the  $(p, q)$ -regularized operator family generated by  $A$ . On the basis of the operator  $S_p$ , the two operators  $P_p, Q_p: R_+ \rightarrow L(E)$  are reminded:

$$P_p(t) = \int_0^t S_p(\nu) d\nu, \quad t \geq 0, \tag{9}$$

$$Q_p(t) = \frac{1}{\Gamma(p)} \int_0^t (t-\nu)^{(p-1)} S_p(\nu) d\nu, \quad t \geq 0. \tag{10}$$

The following basic statement is deduced from [22].

**Lemma 2.** (see [15, 16]). Let  $A$  be a densely linear closed operator in Banach space  $E$ . Then,  $A$  is the generator of a  $(p, q)$ -regularized operator family if and only if there is a strong continuous operator  $S_p: R_+ \rightarrow L(E)$  and constant  $\alpha \geq 0$  satisfying  $\{\eta^p(\eta + q): \eta > \delta\} \subset \rho(A)$  and

$$H(\eta)x := \eta^p(\eta^p(\eta + q)I - A)^{-1}x = \int_0^\infty e^{-\eta t} S_p(t)x dt, \quad \eta > \delta, x \in E. \tag{11}$$

In view of Lemma 2, from (9) and (11), we have

$$\eta^{p-1}(\eta^p(\eta + q)I - A)^{-1}x = \int_0^\infty e^{-\eta t} P_p(t)x dt, \quad \eta > \delta, x \in E. \tag{12}$$

Combining (9) and (10), we obtain

$$(\eta^p(\eta + q)I - A)^{-1}x = \int_0^\infty e^{-\eta t} Q_p(t)x dt, \quad \eta > \delta, x \in E. \tag{13}$$

At first, we put forward a new form of the mild solution for the system in the following.

**Lemma 3.** Let  $0 \leq p \leq 1$  and  $g \in L^1(J, E)$ ; if  $x$  satisfies the equation

$$\begin{cases} {}^C D_t^p x'(t) + q {}^C D_t^p [x(t) - h(t, x_t)] = Ax(t) + g(t), & t \in (0, b], \\ x(t) = \varphi(t), & t \in [-\tau, 0], \\ x'(0) = x_1, \end{cases} \tag{14}$$

then  $x$  is denoted by the integral equation

$$\begin{aligned} x(t) &= S_p(t)\varphi(0) + P_p(t)[x_1 + q\varphi(0) - qh(0, x_0)] \\ &+ q \int_0^t S_p(t-\nu)h(\nu, x_\nu) d\nu + \int_0^t Q_p(t-\nu)g(\nu) d\nu. \end{aligned} \tag{15}$$

*Proof.* Taking the Riemann–Liouville integral to both sides of the first equality of (14), we have

$$I_t^p D_t^p x'(t) + q I_t^p D_t^p [x(t) - h(t, x_t)] = I_t^p [Ax(t) + g(t)]. \tag{16}$$

From Lemma 1, we can get that

$$x'(t) - x'(0) + q[x(t) - h(t, x_t) - x(0) + h(0, x_0)] = I_t^p [Ax(t) + g(t)]. \tag{17}$$

That is,

$$\begin{aligned} x'(t) + qx(t) &= x'(0) + q\varphi(0) - qh(0, x_0) + qh(t, x_t) \\ &+ I_t^p [Ax(t) + g(t)]. \end{aligned} \tag{18}$$

Let  $\eta > 0$ . Taking the Laplace transform for (18), we yield

$$\begin{aligned} -\varphi(0) + \eta X(\eta) + qX(\eta) &= \frac{1}{\eta} [x_1 + q\varphi(0) - qh(0, x_0)] \\ &+ qH(\eta) + \frac{1}{\eta^p} [AX(\eta) + G(\eta)]. \end{aligned} \tag{19}$$

Hence, we obtain

$$\begin{aligned} X(\eta) &= \eta^p [\eta^p(\eta + q)I - A]^{-1} [\varphi(0) + \eta^{p-1} [\eta^p(\eta + q)I - A]^{-1} [x_1 + q\varphi(0) - qh(0, x_0)]] \\ &+ \eta^p [\eta^p(\eta + q)I - A]^{-1} qH(\eta) + [\eta^p(\eta + q)I - A]^{-1} G(\eta). \end{aligned} \tag{20}$$

Then, applying Laplace inverse transform on (20) and combining with the result of Lemma 2 and the Laplace transform of the convolution, we achieve

$$\begin{aligned} x(t) &= S_p(t)\varphi(0) + P_p(t)[x_1 + q\varphi(0) - qh(0, x_0)] \\ &+ q \int_0^t S_p(t-\nu)h(\nu, x_\nu) d\nu + \int_0^t Q_p(t-\nu)g(\nu) d\nu. \end{aligned} \tag{21}$$

This completes the proof.

Because of Lemma 3, we naturally deduce a new representation of the mild solution for (14).  $\square$

**Definition 4.** A function  $x(\cdot) \in C([-\tau, b], E)$  is reputedly a mild solution for the fractional neutral linear system with damping (14) if there exists  $g \in L^1(J, E)$  satisfying

$$\begin{aligned} x(t) &= S_p(t)\varphi(0) + P_p(t)[x_1 + q\varphi(0) - qh(0, x_0)] \\ &+ q \int_0^t S_p(t-\nu)h(\nu, x_\nu) d\nu + \int_0^t Q_p(t-\nu)g(\nu) d\nu, \quad t \in (0, b]. \end{aligned} \tag{22}$$

To gain the global existence of the mild solution for FNDED (1), we give hypothesis S(0).

S(0):  $A$  is the infinitesimal generator of an exponentially bounded  $(p, q)$ -regularized operator family  $S_p(t)$  on  $E$ ,

namely, there exist two real numbers  $M \geq 1$  and  $\omega \geq 0$  satisfying

$$\|S_p(t)\| \leq Me^{\omega t}, \quad t \geq 0. \quad (23)$$

*Remark 2.* From hypothesis S(0) and formulas (9) and (10), we have

- (i)  $\|S_p(t)\| \leq Me^{\omega t} \triangleq M_S, t \in [0, b]$ .
- (ii)  $\|P_p(t)\| = \left\| \int_0^t S_p(s) ds \right\| \leq Me^{\omega t} b \triangleq M_P, t \in [0, b]$ .
- (iii)  $\|Q_p(t)\| = \left\| \int_0^t g_p(t-s) S_p(s) ds \right\| \leq Me^{\omega t} b^p / (\Gamma(p+1)) \triangleq M_Q, t \in [0, b]$ .

### 3. Existence and Uniqueness of Mild Solutions

In this segment, firstly, we consider the existence and uniqueness of mild solutions for FNDED (1). To this end, we impose the following conditions.

S(1):  $h: [0, b] \times E \rightarrow E$  is continuous and satisfies

$$\|h(t, x_t) - h(t, y_t)\|_E \leq n_h(\cdot) |x_t - y_t|, \quad (24)$$

for every  $t \in [0, b]$ , and  $x_t, y_t \in L^2([-\tau, 0]; E)$ ,  $n_h(\cdot) \in L_1(J, R^+)$ , and  $n_h(\cdot) \leq N_h \in R^+$ .

S(2): for almost every  $t \in [0, b]$ , the nonlinear item  $f(t, \cdot): E \rightarrow E$  is continuous, and there is  $n_f(\cdot) \in L_1(J, R^+)$  satisfying

$$\|f(t, x_t) - f(t, y_t)\|_E \leq n_f(\cdot) |x_t - y_t|, \quad (25)$$

for all  $x_t, y_t \in L^2([-\tau, 0]; E)$  and  $n_f(t) \leq N_f \in R^+$ .

**Theorem 1.** Under conditions S(0)-S(2), the fractional neutral differential equations with damping (1) has one and only one mild solution belonging to the space  $C([-\tau, b], E)$  with  $x(t) = \varphi(t)$ ,  $t \in [-\tau, b]$ , and  $x'(t) = x_1$  for every control item  $u(\cdot) \in L_2([0, b]; U)$ , provided that

$$b(b + \tau)(qM_S N_h + M_Q N_f) < 1. \quad (26)$$

*Proof.* Define the mapping  $\Omega: C([-\tau, b], E) \rightarrow C([-\tau, b], E)$  by

$$\left\{ \begin{array}{l} (\Omega x)(t) = S_p(t)\varphi(0) + P_p(t)[x_1 + q\varphi(0) - qh(0, x_0)] + q \int_0^t S_p(t-\nu)h(\nu, x_\nu) ds \\ + \int_0^t Q_p(t-\nu)[Bu(\nu) + f(\nu, x_\nu)] d\nu, \quad t \in [0, b], \\ x(t) = \varphi(t), \quad t \in [-\tau, 0], \\ x'(0) = x_1. \end{array} \right. \quad (27)$$

Let  $M_h = \max_{0 \leq t \leq b} \|h(t, 0)\|_E$ ,  $M_f = \max_{0 \leq t \leq b} \|f(t, 0)\|_E$ ,  $\|B\|_E \leq M_B$ ,  $r = [1 - (qM_S N_h + M_Q N_f)b(b + \tau)]^{-1} [(M_S + qM_P + qM_P N_h)\|\varphi\| + M_P \|x_1\| + q(M_S b + M_P)M_h + bM_Q M_f + b^{(1/2)}M_Q M_B \|u\|_{L_2}]$ , and  $B_r = \{x(\cdot) \in C([-\tau, b]; E): |x| \leq r, x(t) = \varphi(t), t \in [-\tau, 0]\}$ , which is a closed and

bounded subset of  $C([-\tau, b]; E)$ . For each  $y \in B_r$ , we prove  $\Omega$  has a fixed point in  $C([-\tau, b]; E)$ . Firstly, we prove that  $\Omega$  maps  $B_r$  into itself. Obviously,  $\Omega x_0 \in B_r$ . For  $t \in (0, b]$ , we achieve

$$\begin{aligned} \|(\Omega x)(t)\| &\leq \|S_p(t)\varphi\| + \|P_p(t)[x_1 + q\varphi - qh(0, x_0)]\| + q \int_0^t \|S_p(t-\nu)h(\nu, x_\nu)\| d\nu \\ &\quad + \int_0^t \|Q_p(t-\nu)[Bu(\nu) + f(\nu, x_\nu)]\| d\nu \\ &\leq M_S \|\varphi\| + M_P [\|x_1\| + q\|\varphi\| + q\|h(0, x_0) - h(0, 0) + h(0, 0)\|] \\ &\quad + qM_S \int_0^t \|h(\nu, x_\nu) - h(\nu, 0) + h(\nu, 0)\| d\nu \\ &\quad + M_Q \int_0^t \|Bu(\nu)\| d\nu + M_Q \int_0^t \|f(\nu, x_\nu) - f(\nu, 0) + f(\nu, 0)\| d\nu \\ &\leq (M_S + qM_P + qM_P N_h)\|\varphi\| + M_P \|x_1\| + q(M_S b + M_P)M_h \\ &\quad + bM_Q M_f + b^{(1/2)}M_Q M_B \|u\|_{L_2} + (qM_S N_h + M_Q N_f)b(b + \tau)r \\ &= r. \end{aligned} \quad (28)$$



This is because

$$\begin{aligned} \int_0^t \|x_\nu\| d\nu &\leq \int_0^b \int_{-\tau}^0 \|x(\nu + \vartheta)\| d\vartheta d\nu \\ &= \int_0^b \int_{\nu-\tau}^\nu \|x(\omega)\| d\omega d\nu \\ &\leq b(b + \tau) \sup_{t \in [-\tau, b]} \|x(\omega)\| \\ &= b(b + \tau)|x|, \end{aligned} \tag{29}$$

which means  $\|(\Omega x)(t)\| \leq r, \Omega x \subseteq B_r$ .

Next, we demonstrate that  $\Omega$  is a contraction mapping on  $B_r$ . In fact, for  $x(\cdot), y(\cdot) \in B_r$ ,

$$\begin{aligned} \|(\Omega x)(t) - (\Omega y)(t)\| &\leq q \int_0^t \|S_p(t - \nu)\| \|h(\nu, x_\nu) - h(\nu, y_\nu)\| d\nu \\ &\quad + \int_0^t \|Q_p(t - s)\| \|f(s, x_s) - f(s, y_s)\| ds \\ &\leq qM_S \int_0^t n_h(\nu) \|x_\nu - y_\nu\| ds + M_Q \int_0^t n_f(\nu) \|x_\nu - y_\nu\| d\nu \\ &\leq b(b + \tau)(qM_S N_h + M_Q N_f) |x - y|. \end{aligned} \tag{30}$$

Since

$$\begin{aligned} \int_0^t \|x_\nu - y_\nu\| d\nu &\leq \int_0^b \int_{-\tau}^0 \|x(\nu + \vartheta) - y(\nu + \vartheta)\| d\vartheta d\nu \\ &= \int_0^b \int_{\nu-\tau}^\nu \|x(\omega) - y(\omega)\| d\omega d\nu \\ &\leq \int_0^b \int_{-\tau}^b \|x(\omega) - y(\omega)\| d\omega d\nu \\ &\leq b(b + \tau) \sup_{t \in [-\tau, b]} \|x(\omega) - y(\omega)\| \\ &= b(b + \tau)|x - y|, \end{aligned} \tag{31}$$

due to  $b(b + \tau)(qM_S N_h + M_Q N_f) < 1$ ,  $\Omega$  is a contraction mapping. Consequently,  $\Omega$  has the unique fixed point belonging to the space  $C([-\tau, b]; E)$ . (27) is the unique mild solution of FNDED (1).  $\square$

#### 4. Approximate Controllability

In this segment, we acquire several appropriate sufficient conditions of the approximate controllability for FNDED (1) by virtue of the approximate technique and the iterative approach.

Define two continuous linear operators  $\mathcal{L}$  and  $\mathcal{F}$  from  $L_2([0, b], E)$  to  $E$  by

$$\begin{aligned} \mathcal{L}v &= \int_0^b P_p(b - \nu)v(\nu)d\nu, \quad v(\cdot) \in L_2([0, b], E), \\ \mathcal{F}v &= q \int_0^b S_p(b - \nu)v(\nu)d\nu, \quad v(\cdot) \in L_2([0, b], E). \end{aligned} \tag{32}$$

If we let the combination  $(x, u)$  be the mild solution for FNDED (1) with  $u \in L_2([0, b], U)$ , then we represent it as  $x(t) = x(t; x_0, x_1, u)$ , and the terminal item  $x(b)$  can be written as

$$\begin{aligned} x(b) &= x(b; x_0, x_1, u) = S_p(b)\varphi(0) + P_p(b)[x_1 + q\varphi(0) - qh(0, x_0)] \\ &\quad + q \int_0^b S_p(b - \nu) \cdot h(\nu, x_\nu(Bu))d\nu + \int_0^b Q_p(b - \nu)[Bu(\nu) + f(\nu, x_\nu(Bu))]d\nu \\ &\triangleq H(b, \varphi, x_1) + \mathcal{F}h(\nu, x_\nu(Bu)) + \mathcal{L}f(\nu, x_\nu(Bu)) + \mathcal{L}Bu(\nu), \end{aligned} \tag{33}$$

where  $H(b; \varphi, x_1) = S_p(b)\varphi(0) + P_p(b)[x_1 + q\varphi(0) - qh(0, x_0)]$ .

So, the reachable set  $K_b(f)$  which is composed of all possible final states at time  $b$  is

$$K_b(f) = \{x(b; x_0, x_1, u); u \in U\}. \quad (34)$$

Thus, the approximate controllability for FNDED (1) means the set  $K_b(f)$  is dense on space  $E$ . That is to say, the definition of approximate controllability is acquired.

*Definition 5.* Let  $x_0, x_1 \in E$ . We called the fractional neutral differential equations with damping (1) is approximately controllable on  $[0, b]$  if, for any  $\varepsilon > 0$  and  $x_b \in E$ , there is a control function  $u_{N_\varepsilon} \in U$  satisfying

$$\|x_b - H(b; \varphi, x_1) - \mathcal{F}h(\nu, x_\nu(Bu_{N_\varepsilon})) - \mathcal{L}f(\nu, x_\nu(Bu_{N_\varepsilon})) - \mathcal{L}Bu_{N_\varepsilon}\| < \varepsilon. \quad (35)$$

And we add some of the postulated conditions.

S(3): for each given  $\varepsilon > 0$  and  $\nu(\cdot), w(\cdot) \in L_2([0, b]; E)$ , there is  $u(\cdot) \in U$  satisfying

$$\|\mathcal{F}w - \mathcal{L}v - \mathcal{L}Bu\|_E < \varepsilon, \quad (36)$$

where  $\|Bu(\cdot)\|_{L_2(J,U)} \leq M_u \|\nu(\cdot)\|_{L_2(J,U)}$ , and  $M_u$  is a positive real number irrelevant of  $\nu(\cdot)$  and holds

$$(qM_S N_h + M_Q N_f + M_u M_Q N_f) b(b + \tau) < 1. \quad (37)$$

Because condition (37) implies (26), the existence is still met when inequality (26) is changed into (37) in Theorem 1. Next, to demonstrate the conclusion of the approximate controllability remains true for (1), we give the following two lemmas.

**Lemma 4.** Let  $(x_1, u_1), (x_2, u_2)$  be two pairs relevant to FNDED (1). Then, in terms of conditions S(0)–S(3), the following result holds:

$$|x_t(Bu_1) - x_t(Bu_2)| \leq \frac{M_Q b^{(1/2)}}{1 - (qM_S N_h + M_Q N_f) b(b + \tau)} \|Bu_1 - Bu_2\|_{L_2}. \quad (38)$$

*Proof.* The mild solution  $x(t) = x(t; x_0, x_1, u)$  of FNDED (1) in  $E$  satisfies

$$x_t(Bu) = \begin{cases} \varphi(t), & t \in [-\tau, 0], \\ S_p(t)\varphi(0) + P_p(t)[x_1 + q\varphi(0) - qh(0, x_0)] \\ \quad + q \int_0^t S_p(t-\nu)h(\nu, x_\nu) d\nu + \int_0^t Q_p(t-\nu)g(\nu) d\nu, & t \in [0, b]. \end{cases} \quad (39)$$

We define  $y(\cdot, \varphi): [-\tau, b] \rightarrow E$  as

$$y(t, \varphi) = \begin{cases} \varphi(t), & t \in [-\tau, 0], \\ S_p(t)\varphi(0) + P_p(t)[x_1 + q\varphi(0) - qh(0, x_0)], & t \in [0, b]. \end{cases} \quad (40)$$

Let  $x_t = y_t + z_t, t \in [-\tau, b]$ . Apparently,  $x(\cdot)$  meets (1) and only if  $z_0 = 0$  and  $t \in [-\tau, 0]$ , and for each  $t \in [0, b]$ , we acquire

$$\begin{aligned} z(t) &= q \int_0^t S_p(t-\nu)h(\nu, x_\nu) d\nu + \int_0^t Q_p(t-\nu)[Bu(\nu) + f(\nu, x_\nu)] d\nu \\ &= q \int_0^t S_p(t-\nu)h(\nu, z_\nu(Bu) + y_\nu) d\nu + \int_0^t Q_p(t-\nu)[Bu(\nu) + f(\nu, z_\nu(Bu) + y_\nu)] d\nu. \end{aligned} \quad (41)$$

Now, let us take  $x_t(Bu_1), x_t(Bu_2) \in E$  and  $u_1(\cdot), u_2(\cdot) \in U$ ; then,

$$\begin{aligned} &|x_t(Bu_1) - x_t(Bu_2)| \\ &= |[z_t(Bu_1) + y_t] - [z_t(Bu_2) + y_t]| \\ &= |z_t(Bu_1) - z_t(Bu_2)| \\ &\leq q \int_0^t \|S_p(t-\nu)\| \cdot \|h(\nu, x_\nu(Bu_1)) - h(\nu, x_\nu(Bu_2))\| d\nu \\ &\quad + \int_0^t \|Q_p(t-\nu)\| \cdot \|Bu_1(\nu) - Bu_2(\nu)\| d\nu \\ &\quad + \int_0^t \|Q_p(t-\nu)\| \cdot \|f(\nu, x_\nu(Bu_1)) - f(\nu, x_\nu(Bu_2))\| d\nu \\ &\leq qM_S N_h \int_0^t \|x_\nu(Bu_1) - x_\nu(Bu_2)\| d\nu + b^{(1/2)} M_Q \|Bu_1(\nu) - Bu_2(\nu)\|_{L_2} \\ &\quad + M_Q N_f \int_0^t \|x_\nu(Bu_1) - x_\nu(Bu_2)\| d\nu \\ &= M_Q b^{(1/2)} \|Bu_1(\nu) - Bu_2(\nu)\|_{L_2} + (qM_S N_h + M_Q N_f) \int_0^t \|z_\nu(Bu_1) - z_\nu(Bu_2)\| d\nu \\ &\leq M_Q b^{(1/2)} \|Bu_1(\nu) - Bu_2(\nu)\|_{L_2} + (qM_S N_h + M_Q N_f) b(b + \tau) \|z_\nu(Bu_1) - z_\nu(Bu_2)\|. \end{aligned} \quad (42)$$

Recombining the formula, we get

$$\begin{aligned} & |x_t(Bu_1) - x_t(Bu_2)| \\ &= |z_t(Bu_1) - z_t(Bu_2)| \\ &\leq \frac{M_Q b^{(1/2)}}{1 - (qM_S N_h + M_Q N_f)b(b + \tau)} \|Bu_1 - Bu_2\|_{L_2}. \end{aligned} \tag{43}$$

This completes the proof. □

**Theorem 2.** *If hypotheses S(0)–S(3) hold, the fractional neutral differential equations with damping (1) is approximately controllable on J.*

*Proof.* A is the infinitesimal generator of a  $(p, q)$ -regularized operator family in a Banach space, and its domain is dense in Banach space E. So, in accordance with the definition of approximate controllability, it is only needed to demonstrate that  $D(A) \subset K_b(\varphi, x_1)$ . Next, we should prove that, for every  $\varepsilon > 0$  and  $x_b \in D(A)$ , there is  $u_{N_\varepsilon} \in U$  with

$$\|x_b - H(b, \varphi, x_1) - \mathcal{F}h(\nu, x_\nu(Bu_{N_\varepsilon}(\nu))) - \mathcal{L}f(\nu, x_\nu(Bu_{N_\varepsilon}(\nu))) - \mathcal{L}Bu_{N_\varepsilon}(\nu)\| < \varepsilon, \tag{44}$$

where  $x_\nu(t) = x(t; \varphi, x_1, u_{N_\varepsilon})$  satisfies

$$\begin{aligned} x_\nu(t) &= H(t; \varphi, x_1) + q \int_0^t S_p(t - \nu)h(\nu, x_\nu(Bu_{N_\varepsilon}(\nu)))d\nu \\ &+ \int_0^t Q_p(t - \nu)[Bu_{N_\varepsilon}(\nu) + f(\nu, x_\nu(Bu(\nu)))]d\nu, \quad t \in [0, b]. \end{aligned} \tag{45}$$

By the theory of the fractional resolvent, we deduce that  $H(b; \varphi, x_1) \in D(A)$  for  $\varphi(t), x_1 \in E$ , which implies that  $x_b - H(b; \varphi, x_1) \in D(A)$  for  $x_b \in D(A)$ . It is evident that there is one  $\nu \in L_1([0, b], E)$  such that

$$\mathcal{L}\nu = x_b - H(b; \varphi, x_1). \tag{46}$$

Condition (S3) suggests for each  $\varepsilon > 0$  and  $u_1 \in U$ , there is  $u_2 \in U$  such that

$$\|x_b - H(b; \varphi, x_1) - \mathcal{F}h(\nu, x_\nu(Bu_1)) - \mathcal{L}f(\nu, x_\nu(Bu_1)) - \mathcal{L}Bu_2\| < \frac{\varepsilon}{3^2}. \tag{47}$$

Furthermore, for  $u_2 \in U$ , we determine  $w_2 \in U$  again by condition S(3) and (36) with the following two properties:

$$\begin{aligned} & \|\mathcal{F}[h(\nu, x_\nu(Bu_2)) - h(\nu, x_\nu(Bu_1))] + \mathcal{L}[f(\nu, x_\nu(Bu_2)) - f(\nu, x_\nu(Bu_1))] - \mathcal{L}Bw_2\| < \frac{\varepsilon}{3^3}, \\ & \|Bw_2\|_{L_2} \leq M_u \|f(\nu, x_\nu(Bu_2)) - f(\nu, x_\nu(Bu_1))\|_{L_2} \\ & \leq M_u N_f \left( \int_0^b \|x_\nu(Bu_2) - x_\nu(Bu_1)\|_{L_1}^2 d\nu \right)^{(1/2)} \\ & \leq M_u N_f b^{(1/2)} (b + \tau) |x_\nu(Bu_2) - x_\nu(Bu_1)| \\ & \leq \frac{M_Q M_u N_f b (b + \tau)}{1 - (qM_S N_h + M_Q N_f)b(b + \tau)} \|Bu_1 - Bu_2\|_{L_2}, \end{aligned} \tag{48}$$

where  $x_2(t) = x(t; x_0, x_1, u_2), t \in (0, b)$ .

Thus, we may define  $u_3 = u_2 - w_2$  in U and derive the following inequality:

$$\begin{aligned} & \|x(b) - H(b; \varphi, x_1) - \mathcal{F}h(\nu, x_\nu(Bu_2)) - \mathcal{L}f(\nu, x_\nu(Bu_2)) - \mathcal{L}Bu_3\| \\ & \leq \|x(b) - H(b; \varphi, x_1) - \mathcal{F}h(\nu, x_\nu(Bu_1)) - \mathcal{L}f(\nu, x_\nu(Bu_1)) - \mathcal{L}Bu_2\| \\ & \quad + \|\mathcal{L}Bw_2 - \mathcal{F}[h(\nu, x_\nu(Bu_2)) - h(\nu, x_\nu(Bu_1))] - \mathcal{L}[f(\nu, x_\nu(Bu_2)) - f(\nu, x_\nu(Bu_1))]\| \\ & \leq \frac{\varepsilon}{3^2} + \frac{\varepsilon}{3^3}. \end{aligned} \tag{49}$$

By the same way, we obtain the sequence  $\{u_n; n \geq 1\} \subset U$  satisfying

$$\begin{aligned} & \|x(b) - H(b; \varphi, x_1) - \mathcal{H}h(\nu, x_\nu(Bu_n)) - \mathcal{L}f(\nu, x_\nu(Bu_n)) - \mathcal{L}Bu_{n+1}\| \\ & \leq \left(\frac{1}{3^2} + \frac{1}{3^3} + \dots + \frac{1}{3^{n+1}}\right)\varepsilon, \\ & \|Bu_{n+1} - Bu_n\|_{L_2} \leq \frac{M_Q M_u N_f b(b + \tau)}{1 - (qM_S N_h + M_Q N_f)b(b + \tau)} \|Bu_n - Bu_{n-1}\|_{L_2}, \end{aligned} \tag{50}$$

where  $x_n(t) = x(t; x_0, x_1, u_n)$ ,  $t \in (0, b]$  and  $n = 1, 2, \dots$ . Because hypothesis S(3) holds, we can acquire that the sequence  $\{Bu_n; n \geq 1\}$  is a Cauchy sequence in  $L_2([0, b], E)$ , and in this way, we can gain a function item  $\nu \in L_2([0, b], E)$  satisfying

$$\lim_{n \rightarrow \infty} Bu_n = \nu. \tag{51}$$

On account of the mapping  $\mathcal{L}: L_2([0, b], E) \rightarrow E$  is linear-continuous, for any given  $\varepsilon > 0$ , we can find a real integer number  $N_\varepsilon$  satisfying

$$\|L(Bu_{N_\varepsilon+1}) - L(Bu_{N_\varepsilon})\| < \frac{\varepsilon}{3}. \tag{52}$$

Consequently, by inequalities (50) and (52), we derive

$$\begin{aligned} & \|x(b) - H(b; \varphi, x_1) - \mathcal{H}h(\nu, x_\nu(Bu_{N_\varepsilon})) - \mathcal{L}f(\nu, x_\nu(Bu_{N_\varepsilon})) - \mathcal{L}Bu_{N_\varepsilon}\| \\ & \leq \|x(b) - H(b; \varphi, x_1) - \mathcal{H}h(\nu, x_\nu(Bu_{N_\varepsilon})) - \mathcal{L}f(\nu, x_\nu(Bu_{N_\varepsilon})) - \mathcal{L}Bu_{N_\varepsilon+1}\| \\ & \quad + \|\mathcal{L}Bu_{N_\varepsilon+1} - \mathcal{L}Bu_{N_\varepsilon}\| \\ & \leq \left(\frac{1}{3^2} + \frac{1}{3^3} + \dots + \frac{1}{3^{n+1}}\right)\varepsilon + \frac{\varepsilon}{3} \leq \varepsilon, \end{aligned} \tag{53}$$

where  $x_{n_\varepsilon}(t) = x(t; x_0, x_1, u_{n_\varepsilon})$ ,  $t \in (0, b]$ .

This means that  $\xi \in \overline{K_b(f)}$ . Thus,  $\overline{K_b(0)} \subset K_b(f)$ ; therefore, the fractional neutral differential equations with damping (1) is approximately controllable on  $[0, b]$ .  $\square$

**Theorem 3.** Postulate the range of operator  $B$  is denoted by  $R(B)$  and it is dense in  $L_2([0, b]; E)$ . If conditions S(0)–S(3) are true, the fractional neutral differential equations with damping (1) is approximately controllable on  $[0, b]$ .

*Proof.* Because  $R(B)$  is dense in  $L_2([0, b]; E)$ , for every function  $j(\cdot) \in L_2([0, b]; E)$  and  $\iota > 0$ , there is  $Bu(\cdot) \in R(B)$ , where  $u(\cdot) \in U$ , satisfying

$$\|Bu(\cdot) - j(\cdot)\|_{L_2([0, b], E)} < \iota \|j(\cdot)\|_{L_2([0, b], E)}. \tag{54}$$

Now, we have

$$\begin{aligned} \|\mathcal{L}p - \mathcal{L}Bu\| & \leq M_Q \int_0^b \|j(s) - Bu(s)\| ds \\ & \leq M_Q b^{(1/2)} \|j(s) - Bu(s)\|_{L_2([0, b], E)} \\ & \leq M_Q b^{(1/2)} \iota \|j(\cdot)\|_{L_2([0, b], E)} \\ & < \varepsilon. \end{aligned} \tag{55}$$

Thus, from (50), we have

$$\begin{aligned} \|Bu(\cdot)\|_{L_2([0, b], E)} & = \|Bu(\cdot) - j(\cdot) + j(\cdot)\|_{L_2([0, b], E)} \\ & \leq \|Bu(\cdot) - j(\cdot)\|_{L_2([0, b], E)} + \|j(\cdot)\|_{L_2([0, b], E)} \\ & \leq \iota \|j(\cdot)\|_{L_2([0, b], E)} + \|j(\cdot)\|_{L_2([0, b], E)} \\ & = (\iota + 1) \|j(\cdot)\|_{L_2([0, b], E)}. \end{aligned} \tag{56}$$

It indicates that condition S(3) holds if we select  $\iota > 0$ . Therefore, FNDED (1) is approximately controllable on  $[0, b]$  by using Theorem 2.  $\square$

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research was funded by Program for Innovative Research Team in Huainan Normal University (No.

XJTD202008); Natural Science Foundation of Anhui Province (Nos. 2008085MF200 and 2008085QA19); Natural Science Fund of Colleges and Universities in Anhui Province (Nos. KJ2018A0470 and KJ2019A0696); Humanities and Social Sciences Fund of Colleges and Universities in Anhui Province (No. SK2018A0522).

## References

- [1] I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, CA, USA, 1999.
- [2] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, "Theory and applications of fractional differential equations," in *North-Holland Mathematics Studies* Vol. 204, Elsevier, Amsterdam, Netherlands, 2006.
- [3] Y. Zhou, *Basic Theory of Fractional Differential Equations*, World Scientific, Singapore, 2014.
- [4] Y. Zhou, *Fractional Evolution Equations and Inclusions: Analysis and Control*, Elsevier, Amsterdam, Netherlands, 2016.
- [5] H. Liu, Y. Pan, J. Cao, Y. Zhou, and H. Wang, "Positivity and stability analysis for fractional-order delayed systems: A T-S fuzzy model approach," *IEEE Transactions on Fuzzy Systems*, p. 1, 2020.
- [6] H. Liu, Y. Pan, J. Cao, H. Wang, and Y. Zhou, "Adaptive neural network backstepping control of fractional-order nonlinear systems with actuator faults," *IEEE Transactions on Neural Networks and Learning Systems*, 2020.
- [7] J. Du, W. Jiang, D. Pang, and A. U. K. Niazi, "Exact controllability for Hilfer fractional differential inclusions involving nonlocal initial conditions," *Complexity*, vol. 2018, p. 1, Article ID 9472847, 2018.
- [8] S. Kumar and N. Sukavanam, "Approximate controllability of fractional order semilinear systems with bounded delay," *Journal of Differential Equations*, vol. 252, no. 11, pp. 6163–6174, 2012.
- [9] K. X. Li and J. G. Peng, "Controllability of fractional neutral stochastic functional differential systems," *Zeitschrift für Angewandte Mathematik und Physik*, vol. 65, no. 5, pp. 941–959, 2014.
- [10] K. X. Li, J. G. Peng, and J. H. Gao, "Controllability of nonlocal fractional differential systems of order  $\alpha \in (1, 2]$  in Banach spaces," *Reports on Mathematical Physics*, vol. 71, no. 1, pp. 33–43, 2013.
- [11] K. Balachandran, V. Govindaraj, M. Rivero, and J. J. Trujillo, "Controllability of fractional damped dynamical systems," *Applied Mathematics and Computation*, vol. 257, pp. 66–73, 2015.
- [12] B.-B. He, H.-C. Zhou, and C.-H. Kou, "The controllability of fractional damped dynamical systems with control delay," *Communications in Nonlinear Science and Numerical Simulation*, vol. 32, pp. 190–198, 2016.
- [13] Z.-D. Mei and J.-G. Peng, "Riemann-Liouville abstract fractional Cauchy problem with damping," *Indagationes Mathematicae*, vol. 25, no. 1, pp. 145–161, 2014.
- [14] X. Ding and J. J. Nieto, "Controllability and optimality of linear time-invariant neutral control systems with different fractional orders," *Acta Mathematica Scientia*, vol. 35, no. 5, pp. 1003–1013, 2015.
- [15] X. Li, Z. Liu, and C. C. Tisdell, "Existence and exact controllability of fractional evolution inclusions with damping," *Mathematical Methods in the Applied Sciences*, vol. 40, no. 12, pp. 4548–4559, 2017.
- [16] X. Li, Z. Liu, J. Li, and C. Tisdell, "Existence and controllability for nonlinear fractional control systems with damping in Hilbert spaces," *Acta Mathematica Scientia*, vol. 39, no. 1, pp. 229–242, 2019.
- [17] G. Arthi, J. H. Park, and K. Suganya, "Controllability of fractional order damped dynamical systems with distributed delays," *Mathematics and Computers in Simulation*, vol. 165, pp. 74–91, 2019.
- [18] J. Wang, A. G. Ibrahim, M. Fečkan, and Y. Zhou, "Controllability of fractional non-instantaneous impulsive differential inclusions without compactness<sup>†</sup>," *IMA Journal of Mathematical Control and Information*, vol. 36, no. 2, pp. 443–460, 2019.
- [19] Y. K. Chang, A. Pereira, and R. Ponce, "Approximate controllability for fractional differential equations of Sobolev type via properties on resolvent operators," *Fractional Calculus & Applied Analysis*, vol. 20, no. 4, pp. 963–987, 2017.
- [20] V. Vijayakumar, R. Murugesu, R. Poongodi, and S. Dhanalakshmi, "Controllability of second-order impulsive nonlocal Cauchy problem via measure of noncompactness," *Mediterranean Journal of Mathematics*, vol. 14, no. 1, pp. 1–23, 2017.
- [21] R. Triggiani, "A note on the lack of exact controllability for mild solutions in Banach spaces," *SIAM Journal on Control and Optimization*, vol. 15, no. 3, pp. 407–411, 1977.
- [22] C. Lizama, "An operator theoretical approach to a class of fractional order differential equations," *Applied Mathematics Letters*, vol. 24, no. 2, pp. 184–190, 2011.

## Research Article

# Locating-Total Domination Number of Cacti Graphs

Jianxin Wei <sup>1</sup>, Uzma Ahmad,<sup>2</sup> Saira Hameed,<sup>2</sup> and Javaria Hanif<sup>2</sup>

<sup>1</sup>School of Mathematics and Statistics Science, Ludong University, Yantai, Shandong 264025, China

<sup>2</sup>Department of Mathematics, University of the Punjab, Lahore, Pakistan

Correspondence should be addressed to Jianxin Wei; [wjx0426@ldu.edu.cn](mailto:wjx0426@ldu.edu.cn)

Received 13 August 2020; Revised 7 September 2020; Accepted 30 September 2020; Published 19 October 2020

Academic Editor: M. Javaid

Copyright © 2020 Jianxin Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For a connected graph  $J$ , a subset  $W \subseteq V(J)$  is termed as a locating-total dominating set if for  $a \in V(J)$ ,  $N(a) \cap W \neq \emptyset$ , and for  $a, b \in V(J) - W$ ,  $N(a) \cap W \neq N(b) \cap W$ . The number of elements in a smallest such subset is termed as the locating-total domination number of  $J$ . In this paper, the locating-total domination number of unicyclic graphs and bicyclic graphs are studied and their bounds are presented. Then, by using these bounds, an upper bound for cacti graphs in terms of their order and number of cycles is estimated. Moreover, the exact values of this domination variant for some families of cacti graphs including tadpole graphs and rooted products are also determined.

## 1. Introduction

Let  $J = (V, E)$  be a simple undirected and connected graph with vertex set  $V = V(J)$  and edge set  $E = E(J)$ . The number of vertices in  $J$  is called the order of  $J$ . The set of vertices in  $V$  adjacent to any vertex  $v$  is called the open neighbourhood of  $v$  in  $V$ , denoted by  $N_J(v)$ , and the set  $N_J(v) \cup \{v\}$  is called the closed neighbourhood of  $v$  in  $V$ , denoted by  $N_J[v]$ . A vertex of degree one is called leaf (pendant) and its unique neighbouring vertex is called support vertex. A set  $W \subseteq V(J)$  is said to be a dominating set of  $J$  if every vertex in  $V - W$  has nonempty open neighbourhood in  $W$  and furthermore if  $N(a) \cap W \neq \emptyset$  for every vertex  $a \in V(J)$ , then this set  $W$  is referred to as a total dominating set in  $J$ . A total dominating set  $W$  with the additional condition that every pair of distinct vertices outside  $W$  have distinct open neighbourhood in  $W$  is known as a locating-total dominating set, abbreviated as LTD-set. The smallest dominating, total dominating, and locating-total dominating sets of  $J$  are denoted by  $\gamma(J)$ -set,  $\gamma_t(J)$ -set, and  $\gamma_t^l(J)$ -set, respectively, and their cardinalities are known as domination, total domination, and locating-total domination numbers, respectively.

A path and cycle having  $m$  vertices are denoted by  $P_m$  and  $C_m$ . A simple connected graph having no cycle is a tree.

A graph is a unicyclic or bicyclic graph if it contains exactly one cycle or two cycles, respectively. A graph in which any two distinct cycles have maximum one vertex in common is termed as a cactus. A tadpole graph is obtained by joining a leaf (pendant) vertex of path to a vertex of a cycle through an edge. For connected graphs  $K$  and  $L$ , the rooted product of  $K$  by  $L$  ( $K \odot L$ ) is obtained by considering  $|K|$  copies of  $L$  and identifying the  $i$ -th vertex of  $L$  with the root vertex of  $i$ -th copy of  $L$ . The idea of locating-total domination and differentiating-total domination was presented first by Haynes et al. [1]. They studied these variants of domination for trees and estimated a bound on this class. The differentiating-total and locating-total domination numbers of path graphs are also determined. The improved new bound for locating-total dominating number of trees is presented by Chen and Sohn [2]. They also gave the sharpness of these bounds and identified the trees achieving these bounds. The different versions of locating-total domination number for different classes of graphs are investigated in [3–5]. The bounds on locating-total domination number for different types of products are studied in [6, 7]. A new bound of locating-total domination number using annihilation number of a tree was presented by Ning et al. [8]. They also identified the trees attaining this bound. In this paper, locating-total domination number of unicyclic and bicyclic graphs is investigated



and their upper bounds are given. Consequently, using these bounds, an upper bound for cacti graphs is presented. Moreover, the exact values of this domination variant for some classes of cacti graphs are computed. In particular tadpole graphs, rooted products of path by cycle graph and cycle by path graph are computed.

## 2. Bounds of Locating-Total Domination Number of Cacti Graphs

**Theorem 1** (see [2]). *For a tree  $\mathfrak{F}$  having  $r \geq 3$  vertices and  $l$  pendant vertices, one has*

$$\gamma_t^L(\mathfrak{F}) \leq \frac{r+l}{2}. \quad (1)$$

**Proposition 1** (see [9]). *The  $\gamma_t^L$ -set of a graph  $J$  must include all the support vertices of  $J$ .*

*A unicycle graph  $U$  having  $r \geq 3$  vertices can be obtained by connecting any two nonadjacent vertices of a tree  $\mathfrak{F}$  having  $r \geq 3$  vertices by an edge.*

**Lemma 1.** *Let  $W$  be a  $\gamma_t^L$ -set of a tree  $\mathfrak{F}$  with order  $r \geq 4$  and  $l'$  pendant vertices. Let  $U$  be the unicyclic graph of order  $r \geq 4$  with  $l$  pendant vertices constructed by joining any two nonadjacent vertices  $u, v \in V(\mathfrak{F})$  through an edge. If either both belong to  $W$  or none of them belongs to  $W$ , then*

$$\gamma_t^L(U) \leq \frac{r+l+2}{2}. \quad (2)$$

*Proof.* It is easy to see that the  $\gamma_t^L$ -set of a graph  $G'$  obtained by adding edges between two vertices of  $\gamma_t^L$ -set or between two vertices not in  $\gamma_t^L$ -set of a graph  $G$  is the same as that of  $\gamma_t^L$ -set of  $G$ . This implies that  $W$  also becomes a  $\gamma_t^L$ -set for  $U$ . Hence, by Theorem 1, we have

$$\gamma_t^L(U) = |W| = \gamma_t^L(\mathfrak{F}) \leq \frac{r+l'}{2}. \quad (3)$$

Two cases arise:

Case 1: Assume that one of them is a pendant vertex. If exactly one is a pendant vertex, then  $l = l' - 1$  and by (3),

$$\gamma_t^L(U) \leq \frac{r+l+1}{2} < \frac{r+l+2}{2}. \quad (4)$$

But if both are leaves, then  $l = l' - 2$  and by (3),

$$\gamma_t^L(U) \leq \frac{r+l+2}{2}. \quad (5)$$

Case 2: Now assume none of them is a leaf; then,  $l' = l$  and by (3),

$$\gamma_t^L(U) \leq \frac{r+l}{2} < \frac{r+l+2}{2}. \quad (6) \quad \square$$

**Theorem 2.** *Let  $U$  be a unicycle graph having  $r \geq 3$  vertices and  $l$  pendant vertices. Then,*

$$\gamma_t^L(U) \leq \frac{r+l+3}{2}. \quad (7)$$

*Proof.* Let  $\mathfrak{F}$  be a tree of order  $r$  with  $l'$  leaves and  $u, v \in V(\mathfrak{F})$  be its two nonadjacent vertices such that unicycle  $U$  is obtained from  $\mathfrak{F}$  by joining  $u$  and  $v$ . Let  $W$  be a  $\gamma_t^L$ -set of  $\mathfrak{F}$ . Now the vertices  $u$  and  $v$  may belong to  $W$  or not. If both vertices  $u$  and  $v$  belong to  $W$  or both vertices  $u$  and  $v$  do not belong to  $W$ , then by Lemma 1, it is easy to see that  $W$  becomes  $\gamma_t^L$ -set for  $U$  and

$$\gamma_t^L(U) = |W| = \gamma_t^L(\mathfrak{F}) \leq \frac{r+l+2}{2}. \quad (8)$$

Now suppose  $u \notin W$ ,  $v \in W$ , and  $v$  is a pendant vertex. Then there exists support vertex  $s$  such that  $v \sim s$ . By Proposition 1,  $s \in W$ . Now as  $v$  is a pendant,  $N(v) \cap W = \{s\}$  in  $\mathfrak{F}$ , and  $v$  is not adjacent to any vertex of  $V(T) - W$ . Also,  $u$  is adjacent to some vertex  $w \neq u \in W$  so that in  $U$ ,  $\{w, v\} \subseteq N(u) \cap W$ . As in  $U$ ,  $v$  is not adjacent to any other vertex of  $V(U) - W$  except  $u$ , so  $N(u) \cap W \neq N(y) \cap W$  for all  $y \in V(U) - W$ . Moreover, for all  $u \neq y, z \in V(U) - W$ ,  $N(y) \cap W \neq N(z) \cap W$  because for all  $u \neq y, z \in V(T) - W$ ,  $N(y) \cap W$  and  $N(z) \cap W$  remain the same in both  $\mathfrak{F}$  and  $U$ . This further implies that, in this case,  $W$  again becomes  $\gamma_t^L$ -set for  $U$ . Therefore, by Theorem 1 and the fact that either  $l = l' - 1$  or  $l = l' - 2$ , we have

$$\gamma_t^L(U) = |W| = \gamma_t^L(\mathfrak{F}) \leq \frac{r+l'}{2} \leq \frac{r+l+2}{2}. \quad (9)$$

Finally, suppose that all  $u \in W$ ,  $v \notin W$ , and  $v$  is not a leaf. Then,  $W \cup \{v\}$  is  $LTD$ -set for  $U$ . Therefore, by Theorem 1, we have

$$\gamma_t^L(U) \leq |W| + 1 = \gamma_t^L(\mathfrak{F}) + 1 \leq \frac{r+l'}{2} + 1. \quad (10)$$

If  $u$  is a pendant vertex, then  $l = l' - 1$  and by (10), we have

$$\gamma_t^L(U) \leq \frac{r+l+1}{2} + 1 = \frac{r+l+3}{2}. \quad (11)$$

But if  $u$  is not a pendant vertex, then  $l' = l$  and by (10)

$$\gamma_t^L(U) \leq \frac{r+l}{2} + 1 = \frac{r+l+2}{2}. \quad (12)$$

The result follows from (8), (9), (11), and (12).

If in a unicycle graph  $U$  with cycle length  $m$ , all the edges of its unique cycle are removed, then resulting graph is a disconnected graph which is union of  $m$  disjoint subtrees  $T_i$ , where  $i = 1, \dots, m$ . Each subtree  $T_i$  is a maximal subtree of  $U$  containing only one cycle vertex. A bicyclic graph  $B$  can

be constructed from a noncorona unicyclic graph  $U$  by choosing any two nonadjacent vertices  $u, v \in V(T_i) \subseteq V(U)$  for some  $i$  (i.e.,  $u$  and  $v$  must be in one subtree  $T_i$  of  $U$ ) and joining the vertices must contain at least one cycle edge,  $u$  and  $v$  through an edge. It is easy to see that any path between vertices of two distinct subtrees.  $\square$

**Lemma 2.** Consider a unicyclic graph  $U$  with  $r > 5$  vertices and  $l'$  leaves and  $\gamma_t^l$ -set  $W'$ . Choose any two nonadjacent vertices  $u$  and  $v$  from maximal subtree  $T_i$  of  $U$  containing exactly one cycle vertex such that either both belong to  $W'$  or none of them belongs to  $W'$ . Let  $B$  be a bicyclic graph having  $r > 5$  vertices and  $l$  pendant vertices obtained by connecting  $u, v \in V(U)$  through an edge. Then,

$$\gamma_t^l(B) \leq \frac{r+l+5}{2}. \tag{13}$$

*Proof.* The proof is the same as that of Lemma 1.  $\square$

**Theorem 3.** For bicyclic graph  $B$  with order  $r > 4$  and  $l$  pendant vertices, one has

$$\gamma_t^l(B) \leq \frac{r+l+6}{2}. \tag{14}$$

*Proof.* Let  $U$  be the unicyclic graph of order  $r > 4$  with  $l'$  pendant vertices and  $u, v \in V(T_i) \subseteq V(U)$  be its two nonadjacent vertices, where  $T_i$  is maximal subtree of  $U$  containing exactly one cycle vertex. Then, bicyclic graph  $B$  can be obtained from  $U$  by joining  $u$  and  $v$ . Let  $W'$  be a  $\gamma_t^l$ -set of  $U$ . Now the vertices  $u$  and  $v$  may belong to  $W'$  or not.

The result follows by using Lemma 2 and the similar arguments as used in Theorem 2.

Now, as an extension of Theorems 2 and 3, an upper bound for locating-total domination number of cacti graphs is appraised as follows.  $\square$

$$\gamma_t^l(\zeta) \leq \gamma_t^l(\zeta') + 1 \leq \frac{r+l'+3(q-1)}{2} + 1 = \frac{r+l+2+3q-3+2}{2} = \frac{r+l+3q}{2}. \tag{17}$$

Select any two nonpendant vertices  $u$  and  $v$  of  $\zeta'$  such that the path between  $u$  and  $v$  does not include any cycle edge. Now if the vertices  $u$  and  $v$  are connected by an edge, we get a cactus  $\zeta$ . In the case, when exactly one of the  $u$  or  $v$ , say  $u$ , belongs to  $W$ , then  $W \cup \{v\}$  is  $LTD$ -

**Theorem 4.** Let  $\zeta$  be cacti with  $r$  vertices,  $l$  pendant vertices, and  $q$  cycles. Then, one has

$$\gamma_t^l(\zeta) \leq \frac{r+l+3q}{2}. \tag{15}$$

*Proof.* We prove this by making use of mathematical induction on  $q$ . If  $q = 1, 2$ , then the desired result follows immediately by Theorems 2 and 3. Hence, we have a base for induction. Now assume that the bound is satisfied for any cacti graph having order  $r$  and containing  $q - 1$  cycles. Now a cactus  $\zeta$  with order  $r$ , leaves  $l$ , and cycles  $q$  can be obtained from a cactus  $\zeta'$  of order  $r$  with  $l'$  leaves,  $q - 1$  cycles, and  $\gamma_t^l$ -set  $W$  in one of the following three ways:

Select any two pendant vertices  $u$  and  $v$  of  $\zeta'$  such that the shortest path between  $u$  and  $v$  must not have any cycle edge. Now if the vertices  $u$  and  $v$  are connected by an edge, we get a cactus  $\zeta$ . In this case, we can see that the set  $W$  is also  $LTD$ -set of  $\zeta$  by using similar arguments as used in Theorems 2 and 3. Now here  $l' = l + 2$  and using induction hypothesis, we have

$$\begin{aligned} \gamma_t^l(\zeta) \leq \gamma_t^l(\zeta') &\leq \frac{r+l'+3(q-1)}{2} = \frac{r+l+2+3q-3}{2} \\ &< \frac{r+l+3q}{2}. \end{aligned} \tag{16}$$

Select a pendant vertex  $u$  and a nonpendant vertex  $v$  of  $\zeta'$  such that the shortest path between  $u$  and  $v$  does not include any cycle edge. Now if the vertices  $u$  and  $v$  are connected by an edge, we get a cactus  $\zeta$ . In the case, when  $u \notin W$ ,  $W \cup \{u\}$  is  $LTD$ -set of  $\zeta$ ; otherwise,  $W$  is  $LTD$ -set for  $\zeta$  by using similar arguments as used in Theorems 2 and 3. Now here  $l' = l + 2$  and using induction hypothesis, we have

set; otherwise,  $W$  is  $LTD$ -set for  $\zeta$  by using similar arguments as used in Theorems 2 and 3. Now here  $l' = l$  and using induction hypothesis, we have

$$\gamma_t^l(\zeta) \leq \gamma_t^l(\zeta') + 1 \leq \frac{r+l'+3(q-1)}{2} + 1 = \frac{r+l+2+3q-3+2}{2} < \frac{r+l+3q}{2}. \tag{18}$$

Hence, the result follows.  $\square$

### 3. Tadpole and Rooted Product Graphs

In this section, locating-total domination number for few cacti graphs is determined. In particular, tadpole graphs (unicyclic graphs), rooted product of cycle by path (unicyclic graphs), and rooted product of  $P_q$  by  $C_m$  (cacti graph with  $q$  cycles) are investigated.

**Theorem 5** (see [1]). For  $q \geq 3$ ,  $\gamma_t^L(P_q) = \lfloor q/2 \rfloor + \lceil q/4 \rceil - \lfloor q/4 \rfloor$ .

**Theorem 6** (see [4]). For  $q \geq 3$ ,  $\gamma_t^L(C_q) = \lfloor q/2 \rfloor + \lceil q/4 \rceil - \lfloor q/4 \rfloor$ .

*Remark 1.* The locating-total domination number of  $C_q$  and  $P_q$  for different values of  $q$  can be simplified as

- (1)  $\gamma_t^L(C_q) = \gamma_t^L(P_q) = q/2$  if  $q \equiv 0 \pmod{4}$ .
- (2)  $\gamma_t^L(C_q) = \gamma_t^L(P_q) = q/2 + 1$  if  $q \equiv 2 \pmod{4}$ .
- (3)  $\gamma_t^L(C_q) = \gamma_t^L(P_q) = q + 1/2$  if  $q$  is odd.

Let  $T_{q_1, q_2}^r$  ( $q \geq 3$ ) be a tadpole having  $r = q_1 + q_2$  vertices and is formed by connecting a vertex of cycle  $C_{q_1}$  to pendant vertex of path by an edge.

**Theorem 7.** Let  $q_1$  be odd positive integer. Then, one has

$$\gamma_t^L(T_{q_1, q_2}^r) = \begin{cases} \frac{r}{2}, & \text{if } q_2 \equiv 1 \pmod{4}, \\ \frac{r+1}{2}, & \text{if } q_2 \equiv 0, 2 \pmod{4}, \\ \frac{r}{2} + 1, & \text{if } q_2 \equiv 3 \pmod{4}, q_1 \equiv 3 \pmod{4}, \\ \frac{r}{2}, & \text{if } q_2 \equiv 3 \pmod{4}, q_1 \equiv 1 \pmod{4}. \end{cases} \quad (19)$$

*Proof.* Suppose that  $P_{q_2}: a_1 \rightarrow a_2 \rightarrow a_3 \rightarrow \dots \rightarrow a_{q_2}$  is connected through  $a_{q_2}$  to a cycle vertex  $u$  of  $T_{q_1, q_2}^r$ . Let  $W$  be a  $\gamma_t^L$ -set for  $T_{q_1, q_2}^r$ . As  $|W|$  is minimum, therefore  $a_2$  being support vertex by Proposition 1 and  $a_3$  being its unique neighbouring vertex by definition of locating-total dominating set must belong to  $W$ . Consequently,  $a_4$  and  $a_5$  can be excluded from  $W$  and then the next two consecutive vertices belong to  $W$ . By continuing this process, we can see that for  $j = 2, 3 \pmod{4}$ ,  $a \in W$  and for  $j = 0, 1 \pmod{4}$ ,  $a_j \notin W$ .

Case 1: Suppose  $q_2 \equiv 1 \pmod{4}$ . If  $q_2 = 1$ , then  $v_{q_2} \in T_{q_1, q_2}^r$  is a pendant vertex and  $u$  is a support vertex. Thus,  $v_{q_2}, u \in W$  and

$$\gamma_t^L(T_{q_1, 1}^r) = \gamma_t^L(C_{q_1}) = \frac{r}{2}. \quad (20)$$

But if  $q_2 > 1$ , then as  $q_2 \equiv 1 \pmod{4}$ ,  $a_{q_2}, a_{q_2-1} \notin W$  and  $a_{q_2-2}, a_{q_2-3} \in W$ .

Consequently, total  $(q_2 - 1)/2$  vertices of  $P_{q_2}$  belong to  $W$ . Since  $a_{q_2}, a_{q_2-1} \notin W$ , therefore  $u \in W$ . Thus, we have

$$\gamma_t^L(T_{q_1, q_2}^r) = \gamma_t^L(C_{q_2}) + \frac{q_2 - 1}{2} = \frac{r}{2}. \quad (21)$$

The  $\gamma_t^L$ -set of tadpole graph  $T_{q_1, q_2}^r$  for  $q \equiv 1 \pmod{4}$  is represented by black vertices in Figure 1.

Case 2: Suppose  $q_2 \equiv 2 \pmod{4}$ . If  $q_2 = 2$ , then  $a_{q_2}$  is a support vertex. Therefore,  $a_{q_2} \in W$ . Also, using the definition of LTD-set,  $u \in W$ . This implies that  $(q_1 + 1)/2$  cycle vertices must belong to  $W$  and hence we have

$$\gamma_t^L(T_{q_1, 2}^r) = \frac{q_1 + 1}{2} + 1 = \frac{r + 1}{2}. \quad (22)$$

But if  $q_2 > 2$ , then as  $q_2 \equiv 2 \pmod{4}$ , therefore  $a_{q_2} \in D$  and  $a_{q_2-1}, a_{q_2-2} \notin W$ . Consequently,  $q_2/2$  path vertices belong to  $W$ . Also, since  $a_{q_2} \in W$  and  $a_{q_2-1}, a_{q_2-2} \notin W$ , therefore  $u \in W$  which implies that  $(q_1 + 1)/2$  cycle vertices must be in  $W$ . Hence,

$$\gamma_t^L(T_{q_1, q_2}^r) = \frac{q_1 + 1}{2} + \frac{q_2}{2} = \frac{r + 1}{2}. \quad (23)$$

The  $\gamma_t^L$ -set of tadpole graph  $T_{q_1, q_2}^r$  for  $q \equiv 2 \pmod{4}$  is represented by black vertices as shown in Figure 2.

Case 3: Now assume  $q_2 \equiv 0 \pmod{4}$ . If  $q_2 = 0$ , then we have

$$\gamma_t^L(T_{q_1, 0}^r) = \gamma_t^L(C_{q_1}) = \frac{r + 1}{2}. \quad (24)$$

But if  $q_2 > 0$ , then as  $q \equiv 0 \pmod{4}$  and  $|W|$  is minimum, so  $a_{q_1} \notin W$  and  $a_{q_2-1}, a_{q_2-2} \in W$ . Consequently,  $q_2/2$  path vertices lie in  $W$ . Also, since  $a_{q_2} \notin W$  and  $a_{q_2-1}, a_{q_2-2} \in W$  and  $u$  can be excluded from  $W$ . This implies that  $(q_1 + 1)/2$  cycle vertices must lie in  $W$  and hence

$$\gamma_t^L(T_{q_1, q_2}^r) = \frac{q_1 + 1}{2} + \frac{q_2}{2} = \frac{r + 1}{2}. \quad (25)$$

The  $\gamma_t^L$ -set of tadpole  $T_{q_1, q_2}^r$  for  $q_2 \equiv 0 \pmod{4}$  is shown in Figure 3 with black vertices.

Case 4: Finally, assume that  $q_2 \equiv 3 \pmod{4}$ . If  $q_2 = 3$ , then  $a_2$  being support vertex and  $a_3$  being its neighbouring vertex lie in  $W$ . But then  $u$  can be taken out of  $W$ . Moreover, for  $q_1 \equiv 3 \pmod{4}$ ,  $(q_1 + 1)/2$  cycle vertices lie in  $W$  and hence



FIGURE 1: Tadpole  $T_{q_1, q_2}^r$  for  $q_2 \equiv 1 \pmod{4}$ .



FIGURE 4: Tadpole  $T_{q_1, q_2}^r$  for  $q_2 \equiv 3 \pmod{4}$ .



FIGURE 2: Tadpole  $T_{q_1, q_2}^r$  for  $q \equiv 2 \pmod{4}$ .



FIGURE 3: Tadpole  $T_{q_1, q_2}^r$  for  $q_2 \equiv 0 \pmod{4}$ .

$$\gamma_t^L(T_{q_1, 3}^r) = \frac{q_1 + 1}{2} + 2 = \frac{r}{2} + 1. \quad (26)$$

On the other hand, if  $q_1 \equiv 1 \pmod{4}$ , then  $(q_1 - 1)/2$  cycle vertices belong to  $W$  and we have

$$\gamma_t^L(T_{q_1, 3}^r) = \frac{q_1 - 1}{2} + 2 = \frac{r}{2}. \quad (27)$$

Now if  $q_2 > 3$ , then as  $q_2 \equiv 3 \pmod{4}$ ; therefore,  $a_{q_2}, a_{q_2-1} \in W$  and  $a_{q_2-2}, a_{q_2-3} \notin W$ . Consequently,  $(q_2 + 1)/2$  path vertices lie in  $W$ . Also, as  $a_{q_2}, a_{q_2-1} \in W$ , and  $u$  can be excluded from  $W$ . But then the vertices which lie in  $\gamma_t^L(C_{q_1})$  must lie in  $W$ . Moreover, if  $q_1 \equiv 3 \pmod{4}$ , then  $(q_1 + 1)/2$  cycle vertices are in  $W$ . Thus, we have

$$\gamma_t^L(T_{q_1, q_2}^r) = \frac{q_1 + 1}{2} + \frac{q_2 + 1}{2} = \frac{r}{2} + 1, \quad (28)$$

whereas if  $q_1 \equiv 1 \pmod{4}$ , then  $(q_1 - 1)/2$  cycle vertices belong to  $W$  and we have

$$\gamma_t^L(T_{q_1, q_2}^r) = \frac{q_1 - 1}{2} + \frac{q_2 + 1}{2} = \frac{r}{2}. \quad (29)$$

The  $\gamma_t^L$ -set of tadpole  $T_{q_1, q_2}^r$  for  $q_2 \equiv 3 \pmod{4}$  is shown in Figure 4 with black vertices.  $\square$

**Theorem 8.** Let  $q_1$  be an even integer. Then one has

$$\gamma_t^L(T_{q_1, q_2}^r) = \begin{cases} \frac{r+1}{2}, & \text{if } q_2 \equiv 1, 3 \pmod{4}, \\ \frac{r}{2}, & \text{if } q_2 \equiv 2 \pmod{4}, \\ \frac{r}{2}, & \text{if } q_2 \equiv 0 \pmod{4}, q_1 \equiv 0 \pmod{4}, \\ \frac{r}{2} + 1, & \text{if } q_2 \equiv 0 \pmod{4}, q_1 \equiv 2 \pmod{4}. \end{cases} \quad (30)$$

*Proof.* Let  $W$  be a  $\gamma_t^L$ -set for  $T_{q_1, q_2}^r$ . As  $|W|$  is minimum, therefore  $a_2$  being support vertex by Proposition 1 and  $a_3$  being its unique neighbouring vertex by definition of locating-total dominating set must belong to  $W$ . Consequently,  $a_4$  and  $a_5$  can be excluded from  $W$  and then the next two consecutive vertices belong to  $W$ . By continuing this process, we can see that for  $j \equiv 2, 3 \pmod{4}$ ,  $a_j \in W$  and for  $j \equiv 0, 1 \pmod{4}$ ,  $a_j \notin W$ .

Case 1: Suppose  $q_2 \equiv 1 \pmod{4}$ . If  $q_2 = 1$ , then  $a_{q_2} \in T_{q_1, q_2}^r$  is a pendant vertex and  $u$  is a support vertex. Thus,  $a_{q_2}, u \in W$  and

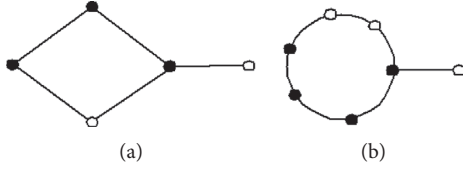
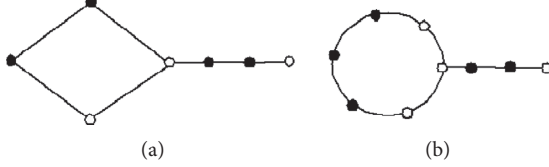
$$\gamma_t^L(T_{q_1, 1}^r) = \frac{q_1}{2} + 1 = \frac{r+1}{2}. \quad (31)$$

The  $\gamma_t^L$ -sets of tadpoles  $T_{4,1}^5$  and  $T_{6,1}^7$  are shown in Figure 5 with black vertices. But if  $q_2 > 1$ .

Then as  $q_2 \equiv 1 \pmod{4}$ ,  $a_{q_2}, a_{q_2-1} \notin W$  and  $a_{q_2-2}, a_{q_2-3} \in W$ . Consequently, total  $(q_2 - 1)/2$  path vertices belong to  $W$ . Further, since  $a_{q_2}, a_{q_2-1} \notin W$ , therefore  $u \in W$ . Thus,  $(q_1/2) + 1$  cycle vertices lie in  $D$  and consequently,

$$\gamma_t^L(T_{q_1, q_2}^r) = \frac{q_1}{2} + 1 + \frac{q_2 - 1}{2} = \frac{r+1}{2}. \quad (32)$$

Case 2: Suppose  $q_2 \equiv 3 \pmod{4}$ . If  $q_2 = 3$ , then  $a_2$  being support vertex and  $a_3$  being its neighbouring vertex lie in  $W$ . And  $u$  can be excluded from  $W$ . This implies that

FIGURE 5: The tadpoles  $T_{4,1}^5$  and  $T_{6,1}^7$ .FIGURE 6: The tadpoles  $T_{4,3}^7$  and  $T_{6,3}^9$ .

$q_1/2$  cycle vertices must lie in  $W$ , and hence, we have

$$\gamma_t^L(T_{q_1,3}^r) = \frac{q_1}{2} + 2 = \frac{r+1}{2}. \quad (33)$$

The  $\gamma_t^L$ -set of tadpoles  $T_{4,3}^7$  and  $T_{6,3}^9$  are shown in Figure 6 with black vertices.

But if  $q_2 > 3$ , then as  $q_2 \equiv 3 \pmod{4}$ , so  $a_{q_2}, a_{q_2-1}$  lie in  $W$ , whereas  $a_{q_2-2}, a_{q_2-3} \notin W$ . Consequently,  $(q_2+1)/2$  path vertices belong to  $W$ . Further, as  $a_{q_2}, a_{q_2-1}$  lie in  $W$  so  $u$  can be taken out of  $W$ . Thus,  $q_1/2$  cycle vertices lie in  $W$  so that

$$\gamma_t^L(T_{q_1,q_2}^r) = \frac{q_1}{2} + \frac{q_2+1}{2} = \frac{r+1}{2}. \quad (34)$$

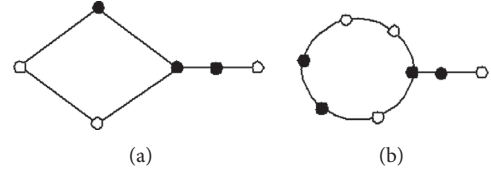
Case 3: Now suppose  $q_2 \equiv 2 \pmod{4}$ . If  $q_2 = 2$ , then  $a_{q_2}$  being a support vertex belongs to  $W$ . This further implies that  $u \in W$ . Hence,  $q_1/2$  cycle vertices lie in  $W$ , and consequently,

$$\gamma_t^L(T_{q_1,2}^r) = \frac{q_1}{2} + 1 = \frac{r}{2}. \quad (35)$$

The  $\gamma_t^L$ -set of tadpoles  $T_{4,2}^6$  and  $T_{6,2}^8$  are shown in Figure 7 with black vertices.

On the other hand, if  $q_2 > 2$ , then  $a_{q_2} \in W$  but  $a_{q_2-1}, a_{q_2-2} \notin W$ . This implies that  $q_2/2$  path vertices are included in  $W$  along with cycle vertex  $u$ . Consequently,  $q_1/2$  cycle vertices lie in  $W$  and we have

$$\gamma_t^L(T_{q_1,q_2}^r) = \frac{q_1}{2} + \frac{q_2}{2} = \frac{r}{2}. \quad (36)$$

FIGURE 7: The tadpoles  $T_{4,2}^6$  and  $T_{6,2}^8$ .

Case 4: Finally assume  $q_2 \equiv 0 \pmod{4}$ . It is easy to see that if  $q_2 = 0$ , then  $T_{q_1,0}^r \cong C_{q_1}$  and  $r = q_1$ .

Hence,

$$\gamma_t^L(T_{q_1,0}^r) = \gamma_t^L(C_{q_1}). \quad (37)$$

Using Remark 3, we have

$$\gamma_t^L(T_{q_1,0}^r) = \begin{cases} \frac{q_1}{2}, & \text{if } q_1 \equiv 0 \pmod{4}, \\ \frac{q_1}{2} + 1, & \text{if } q_1 \equiv 2 \pmod{4}. \end{cases} \quad (38)$$

As  $r = q_1$ , so we have

$$\gamma_t^L(T_{q_1,0}^r) = \begin{cases} \frac{r}{2}, & \text{if } q_1 \equiv 0 \pmod{4}, \\ \frac{r}{2} + 1, & \text{if } q_1 \equiv 2 \pmod{4}. \end{cases} \quad (39)$$

But if  $q_2 > 0$ , then  $a_{q_2} \notin W$  and  $a_{q_2-1}, a_{q_2-2} \in W$ . This shows that  $q_2/2$  path vertices belong to  $W$ . Also, since  $a_{q_2} \notin W$  and  $a_{q_2-1}, a_{q_2-2} \in W$ ,  $u$  must be in  $W$ . Hence, the vertices that belong to  $\gamma_t^L(C_{q_1})$  must lie in  $W$ . Hence, we have

$$\gamma_t^L(T_{q_1,q_2}^r) = \frac{q_2}{2} + \gamma_t^L(C_{q_1}). \quad (40)$$

Again, using Remark 3, we have

$$\gamma_t^L(T_{q_1,q_2}^r) = \begin{cases} \frac{q_2}{2} + \frac{q_1}{2}, & \text{if } q_1 \equiv 0 \pmod{4}, \\ \frac{q_2}{2} + \frac{q_1}{2} + 1, & \text{if } q_1 \equiv 2 \pmod{4}. \end{cases} \quad (41)$$

Thus,

$$\gamma_t^L(T_{q_1,q_2}^r) = \begin{cases} \frac{r}{2}, & \text{if } q_1 \equiv 0 \pmod{4}, \\ \frac{r}{2} + 1, & \text{if } q_1 \equiv 2 \pmod{4}. \end{cases} \quad (42)$$

In Theorem 9, the rooted product  $P_{q_2} \odot C_{q_1}$  is computed. Any cycle vertex can be considered as root vertex.  $\square$

**Theorem 9.** For positive integers  $q_1 \geq 3$  and  $q_2 \geq 1$ , one has



$$\gamma_t^L(P_{q_2} \odot C_{q_1}) = \begin{cases} \gamma_t^L(C_{q_1}), & \text{if } q_2 = 1, \\ q_2 \gamma_t^L(C_{q_1}), & \text{if } q_1 \equiv 0, 3 \pmod{4}, \\ q_2 \gamma_t^L(C_{q_1-1}) + \frac{q_2}{2}, & \text{if } q_1 \equiv 1 \pmod{4}, \\ q_2 \gamma_t^L(C_{q_1-2}) + q_2, & \text{if } q_1 \equiv 2 \pmod{4}. \end{cases} \quad (43)$$

*Proof.* If  $q_2 = 1$ , then  $P_{q_2} \odot C_{q_1} \cong C_{q_1}$  and  $\gamma_t^L(P_{q_2} \odot C_{q_1}) = \gamma_t^L(C_{q_1})$ . Now we may suppose  $q_2 \geq 2$ . Then four cases arise:

Case 1: Suppose  $q_1 \equiv 0 \pmod{4}$ , that is,  $q_1 = 4k$  for integer  $k > 0$ . The graph  $P_{q_2} \odot C_{q_1}$  contains  $q_2$  mutually disjoint cycles  $C_{q_1}^i$ , where  $i = 1, \dots, q_2$ . For each  $i$ , cycle  $C_{q_1}^i$  consists of  $4k$  vertices, namely,  $v_j^i$  where  $j = 1, \dots, 4k$ . Take  $v_1^i$  as root vertex of each cycle. The  $4k$  vertices in each cycle are divided into  $k$  sets such that each set contains four consecutive vertices of cycle. From each set, the middle two vertices can be taken as elements of minimum LTD-set for each  $C_{q_1}^i$ . Let this minimum LTD-set of  $C_{q_1}^i$  be  $W_i = \{v_2^i, v_3^i, v_6^i, v_7^i, \dots, v_{4k-2}^i, v_{4k-1}^i\}$  with  $|W_i| = 2k = \gamma_t^L(C_{q_1}^i)$ . These  $2k$  vertices in each  $W_i$  are shown in Figure 8 as black vertices.

Now the root vertices  $v_1^i$  from each  $C_{q_1}^i$  are actually the path vertices and are locating totally dominated by the elements of  $W_i$  for each  $i$ . Hence, all the vertices of  $P_{q_2} \odot C_{q_1}$  are locally dominated by  $\cup_{i=1}^{q_2} W_i$ . Also, this set is minimum by construction. Thus,

$$\gamma_t^L(P_{q_2} \odot C_{q_1}) = q_2 \gamma_t^L(C_{q_1}). \quad (44)$$

Case 2: Suppose  $q_1 \equiv 1 \pmod{4}$ , that is,  $q_1 = 4k + 1$ , for integer  $k > 0$ . Let  $v_1^i$  be the root vertex of cycle  $C_{q_1}^i$  for each  $i$ . The graph  $C_{q_1}^i - \{v_1^i\}$  contains  $4k$  vertices and these  $4k$  vertices are divided into  $k$  sets. Now we select two middle vertices from each set. Then, all  $4k$  vertices of  $C_{q_1}^i - \{v_1^i\}$  are locating totally dominated by  $2k$  chosen vertices. Since  $v_1^i$  does not have a neighbour in the selected LTD-set of  $C_{q_1}^i$ , therefore,  $v_1^i$  must be a LTD-vertex. The graph  $C_{q_1}^i - \{v_1^i\}$  also contains  $4k$  vertices and these  $4k$  vertices are again divided into  $k$  sets and all are locating totally dominated by  $2k$  vertices by selecting the first and last vertices from each set. Now we are interested in minimum LTD-set;  $v_1^i$  may not be an LTD-vertex. By continuing in this way, we can see that if  $i$  is odd then  $v_1^i$  is not an LTD-vertex otherwise  $v_1^i$  is an LTD-vertex. This along with the fact that path vertices are root vertices implies that all the vertices of graph are locating totally dominated by a set  $W$  of smallest cardinality. We note that if the path (root) vertex is not an LTD-vertex, then associated cycle

has  $2k$  vertices in  $W$  and otherwise each cycle has  $2k + 1$  vertices in  $W$ . Hence,

$$\gamma_t^L(P_{q_2} \odot C_{q_1}) = q_2 \gamma_t^L(C_{q_1-1}) + \frac{q_2}{2}. \quad (45)$$

The  $\gamma_t^L$ -set of  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 1 \pmod{4}$  is shown in Figure 9 with black vertices.

Case 3: Suppose  $q_1 \equiv 2 \pmod{4}$  such that  $q_1 = 4k + 2$  and  $k > 0$  is an integer. In a similar manner as in Cases 1 and 2, we divide the vertices of each  $C_{q_1}^i - \{v_1^i, a_i\}$ , where  $a_i$  is neighbouring vertex of root vertex into  $k$  sets. For each  $i$ , all the vertices of  $C_{q_1}^i - \{v_1^i, a_i\}$  are locating totally dominated by  $2k$  chosen vertices (middle two vertices in each set). The vertices  $v_1^i$  and  $a_i$  for all  $i$  can be locating totally dominated if we include all the root vertices. This implies that

$$\gamma_t^L(P_{q_2} \odot C_{q_1}) = q_2 \gamma_t^L(C_{q_1-2}) + q_2. \quad (46)$$

The  $\gamma_t^L$ -set of  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 2 \pmod{4}$  is shown in Figure 10 with black vertices.

Case 4: Finally, suppose  $q_1 = 4k + 3$ , where  $k > 0$  is an integer. Here, the vertices of  $C_{q_1}^i - \{v_1^i, a_i, b_i\}$ , where  $a_i$  and  $b_i$  are consecutive two vertices of root vertex  $v_1^i$  on cycle, are divided into  $k$  sets. Again for each  $i$ , all the vertices  $C_{q_1}^i - \{v_1^i, a_i, b_i\}$  are locating totally dominated by  $2k$  chosen vertices (first two vertices in each set). The vertices  $\{v_1^i, a_i, b_i\}$  for all  $i$  can be locating-total dominated by selecting two of them except root vertex. Therefore, in this case, we have

$$\gamma_t^L(P_{q_2} \odot C_{q_1}) = q_2 \gamma_t^L(C_{q_1}). \quad (47)$$

The  $\gamma_t^L$ -set of  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 3 \pmod{4}$  is shown in Figure 11 with black vertices.

In the following Theorem 10, the rooted product  $C_{q_1} \odot P_{q_2}$  is computed. Here, we select pendant vertex of path as root vertex.  $\square$

**Theorem 10.** For  $q_1 \geq 3$  and  $q_2 \geq 2$  and by taking pendant vertex of  $P_{q_2}$  as root vertex, one has

$$\gamma_t^L(C_{q_1} \odot P_{q_2}) = \begin{cases} q_1 \gamma_t^L(P_{q_2}), & \text{if } q_2 \equiv 0, 2 \pmod{4}, \\ q_1 \gamma_t^L(P_{q_2-1}) + \gamma_t^L(C_{q_1}), & \text{if } q_2 \equiv 1 \pmod{4}, \\ q_1 \gamma_t^L(P_{q_2-3}) + 2q_1, & \text{if } q_2 \equiv 3 \pmod{4}. \end{cases} \quad (48)$$

*Proof*

Case 1: Suppose  $q_2 \equiv 0 \pmod{4}$ , such that  $q_2 = 4k$  and  $k > 0$ . The rooted product  $C_{q_1} \odot P_{q_2}$  consists of mutually disjoint  $q_1$  paths  $P_{q_2}^i$  for  $i = 1, \dots, q_1$ . Suppose  $v_1^i$  is the



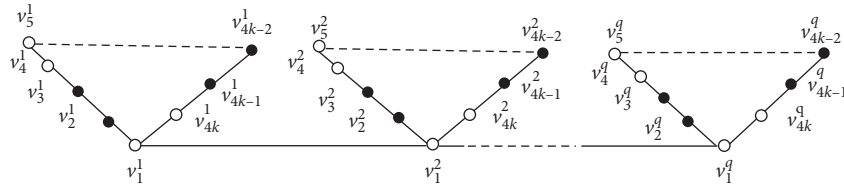


FIGURE 8: The rooted product  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 0 \pmod{4}$ .

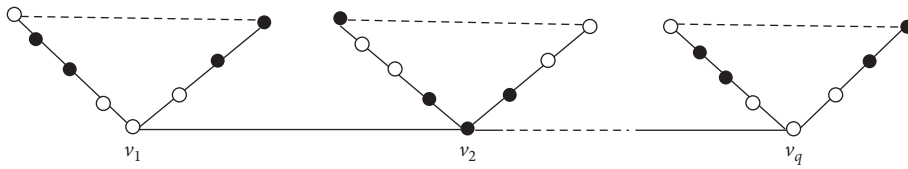


FIGURE 9: The rooted graph  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 1 \pmod{4}$ .

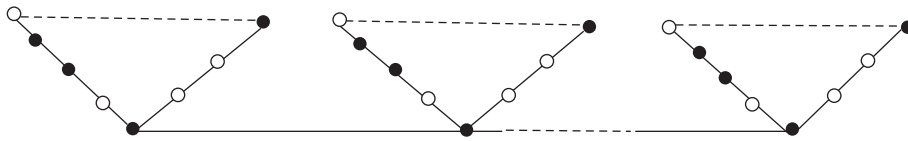


FIGURE 10: The graph  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 2 \pmod{4}$ .

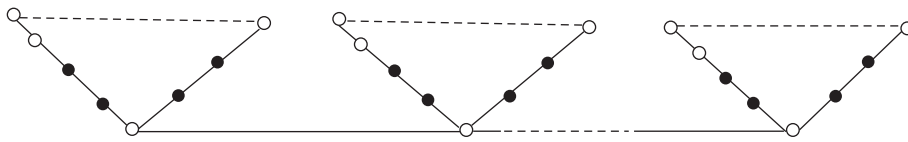


FIGURE 11: The graph  $P_{q_2} \odot C_{q_1}$  for  $q_1 \equiv 3 \pmod{4}$ .

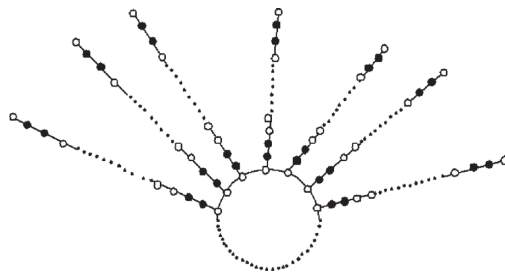


FIGURE 12: The product  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 0 \pmod{4}$ .

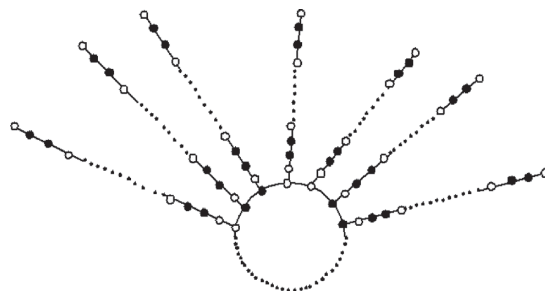


FIGURE 13: The graph  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 1 \pmod{4}$ .

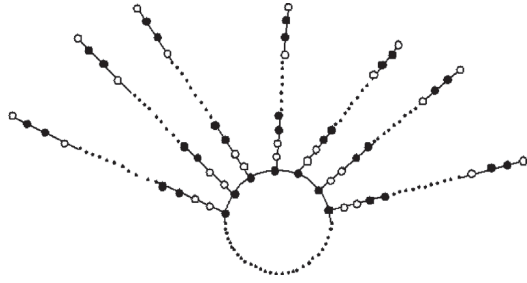


FIGURE 14: The product  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 2 \pmod{4}$ .

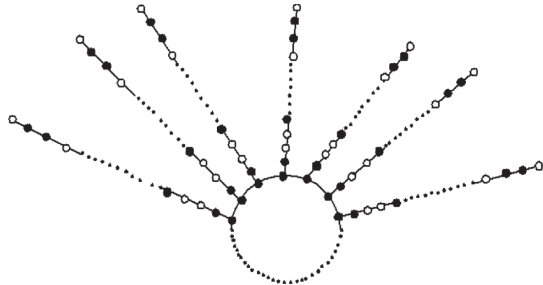


FIGURE 15: The graph  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 3 \pmod{4}$ .

root vertex of  $P_{q_2}^i$  for  $i = 1, \dots, q_1$ . The  $4k$  vertices of  $P_{q_2}^i$  are split into  $k$  sets, each containing four consecutive vertices. Then the collection of middle two vertices from each set in every path gives the  $LTD$ -set for graph  $C_{q_1} \odot P_{q_2}$  which is also minimum one:

$$\gamma_t^L(C_{q_1} \odot P_{q_2}) = q_1 \gamma_t^L(P_{q_2}). \tag{49}$$

The  $\gamma_t^L$ -set of  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 0 \pmod{4}$  is represented by black vertices in Figure 12.

Case 2: Assume  $q_2 \equiv 1 \pmod{4}$  such that  $q_2 = 4k + 1$  and  $k > 0$ . Let  $v_o^i$  be the root vertex in each path  $P_{q_2}^i$  of  $C_{q_1} \odot P_{q_2}$ . Consider  $P_{q_2}^i - \{v_o^i\}$  consisting of  $4k$  vertices and all the vertices in  $P_{q_2}^i - \{v_o^i\}$  can be locating totally dominated by  $2k$  chosen vertices (two middle vertices from each set). Since all the paths  $P_{q_2}^i$  for  $i = 1, \dots, q_1$  are mutually disjoint, therefore  $\cup_i^{q_1} \{P_{q_2}^i - \{v_o^i\}\}$  is minimum locating-total dominated set consisting of  $2k \cdot q_1$  vertices. Therefore,

$$\begin{aligned} \gamma_t^L(C_{q_1} \odot P_{q_2}) &= 2kq_1 + \gamma_t^L(C_{q_1}) \\ &= q_1 \gamma_t^L(P_{q_2-1}) + \gamma_t^L(C_{q_1}). \end{aligned} \tag{50}$$

The  $\gamma_t^L$ -set of  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 1 \pmod{4}$  is represented by black vertices as shown in Figure 13.

Case 3: Now suppose that  $q_2 \equiv 2 \pmod{4}$  such that  $q_2 = 4k + 2$ . Here, each  $P_{q_2}^i$  for  $i = 1, \dots, q_1$  consists of  $4k + 2$  vertices. All of them except the root and its neighbouring vertex are locating totally dominated by  $2k$  chosen vertices from each path as in the previous

cases. These chosen vertices must belong to  $\gamma_t^L$ -set. To obtain, minimum  $LTD$ -se for rooted product, all the root vertices are also included. Hence,

$$\gamma_t^L(C_{q_1} \odot P_{q_2}) = 2kq_1 + q_1 = q_1 \gamma_t^L(P_{q_2-2}) + q_1. \tag{51}$$

The  $\gamma_t^L$ -set of  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 2 \pmod{4}$  is represented by black vertices in Figure 14.

Case 4: The case  $q_2 \equiv 3 \pmod{4}$  is dealt in a similar manner as in previous cases. Thus, we have

$$\gamma_t^L(C_{q_1} \odot P_{q_2}) = q_1 \gamma_t^L(P_{q_2-3}) + 2q_1. \tag{52}$$

The  $\gamma_t^L$ -set of  $C_{q_1} \odot P_{q_2}$  for  $q_2 \equiv 3 \pmod{4}$  is shown in Figure 15 with black vertices.

### 4. Conclusion

In this paper, the upper bounds for locating-total domination number of unicyclic and bicyclic graphs are obtained. It is shown that, for unicyclic graph  $U$  of order  $r$  and leaves  $l$ ,  $\gamma_t^L(U) \leq r + l + 3/2$  and for bicyclic graph  $B$  of order  $r$  and leaves  $l$ ,  $\gamma_t^L(B) \leq r + l + 6/2$ . These bounds are then generalized for cacti graphs of order  $n$  having  $q$  cycles as  $r + l + 3q/2$ . Also, the exact value of locating-total domination numbers for some cacti graphs is determined. For tadpole graph  $T_{q_1, q_2}^r$ , it is found that locating-total dominating number is  $r/2$ ,  $r + 1/2$ , or  $r/2 + 1$  depending upon the parity of  $r$  and values of  $q_1$  and  $q_2$ . The locating-total dominating number of rooted product of  $C_{q_1}$  by  $P_{q_2}$  in terms of  $q_1$  and  $\gamma_t^L(P_{q_2})$  is computed. Similarly, locating-total dominating number of rooted products of  $P_{q_2}$  by  $C_{q_1}$  in terms of  $q_2$  and  $\gamma_t^L(C_{q_1})$  is also determined. It is still an open problem to find the sharpness of estimated bounds of locating-total dominating number for cacti graphs.

### Data Availability

All the required data are included within the paper. However, the reader may contact the corresponding author for more details of the data.

### Conflicts of Interest

The authors do not have any conflicts of interest.

### Acknowledgments

This work was sponsored by Shandong Provincial Natural Science Foundation, China (ZR2018MA010).

### References

[1] T. W. Haynes, M. A. Henning, and J. Howard, "Locating and total dominating sets in trees," *Discrete Applied Mathematics*, vol. 154, no. 8, pp. 1293–1300, 2006.

- [2] X.-G. Chen and M. Y. Sohn, "Bounds on the locating-total domination number of a tree," *Discrete Applied Mathematics*, vol. 159, no. 8, pp. 769–773, 2011.
- [3] M. Chellali and N. Jafari Rad, "Locating-total domination critical graphs," *Australasian Journal of Combinatorics*, vol. 45, pp. 227–234, 2009.
- [4] M. A. Henning and C. Löwenstein, "Locating-total domination in claw-free cubic graphs," *Discrete Mathematics*, vol. 312, no. 21, pp. 3107–3116, 2012.
- [5] J. McCoy and M. A. Henning, "Locating and paired-dominating sets in graphs," *Discrete Applied Mathematics*, vol. 157, no. 15, pp. 3268–3280, 2009.
- [6] S. B. Rashmi and I. P. Kelkar, "Total domination number of rooted product graph," *Journal of Computer and Mathematical Sciences*, vol. 8, no. 6, pp. 0976–5697, 2017.
- [7] H. Xing and M. Y. Sohn, "Bounds on locating total domination number of the Cartesian product of cycles and paths," *Information Processing Letters*, vol. 115, no. 12, pp. 950–956, 2015.
- [8] W. Ning, M. Lu, and K. Wang, "Bounding the location-total domination number of a tree in terms of its annihilation number," *Discussiones Mathematicae Graph Theory*, vol. 39, no. 1, 2019.
- [9] U. Ahmad and J. Ibrahim, "Lower bound of locating-total domination of unicyclic graphs," preprint.

## Research Article

# Control of a Tilting Hexacopter under Wind Disturbance

Hao Xu <sup>1,2</sup>, Zhong Yang <sup>1</sup>, Kaiwen Lu <sup>1</sup>, Changliang Xu <sup>1</sup> and Qiuyan Zhang <sup>3</sup>

<sup>1</sup>College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

<sup>2</sup>School of Mathematics & Physics, Anhui University of Technology, Ma'anshan 243032, China

<sup>3</sup>Guizhou Electric Power Research Institute, Guizhou Power Grid Co., Ltd., Guiyang, Guizhou 550000, China

Correspondence should be addressed to Zhong Yang; yangzhong@nuaa.edu.cn

Received 20 August 2020; Revised 23 September 2020; Accepted 28 September 2020; Published 13 October 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Hao Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multicopters are well suited for executing fast maneuvers with high velocity, but they are still affected by the external atmospheric environment because attitude and position cannot be independently controlled. In this paper, we present a novel hexacopter which improves the wind resistance and strong coupling between attitude control and position control. The copter is designed such that the rotor sections can tilt around their respective arm axes. We present the aerodynamic methods to analyze the system dynamics model in windy environments. The entire system is decomposed into six loops based on the model, and the presented flight controller uses the ADRC method to consider both the attitude and the position. The controller introduces the extended state observer to estimate the white noise and wind disturbance. We use the nonlinear state error feedback law to control the output with disturbance compensation. Finally, we linearize the control allocation matrix that the controllers output directly mapped to the rotor velocities and tilting angles. The new theoretical results are thoroughly validated through comparative experiments.

## 1. Introduction

In recent years, multicopters had proven to execute aggressive maneuvers [1]. The copter is an underactuated system that can move independently in six directions [2]. Because the control inputs are less than the motion directions, the generated driving force is in the same direction as the body frame and the strong coupling between the torque and the force, so the six degree-of-freedom omnidirectional motion cannot be realized [3]. However, to the multicopter controller, only the height and attitude are usually independently controllable, and it is challenging to meet the applications that require independent control of the position and attitude [4] such as a fixed-point hover while doing attitude tracking or maintaining attitude while doing position tracking [5, 6].

In [7, 8], a reconstruction scheme of a six-rotor aircraft was shown; changing the mounting angle of the rotor, the rotors were arranged in pairs in three different planes, constituting a Euclidean space. The rotor installation angle was changed in the two schemes that the rotor thrust had

components in the Z-axis and the X-Y plane of body frames. Thus, the copter can generate control and control torque in any direction for the independent position and attitude control [9]. However, since the thrust direction is fixed in the body frame, there is still horizontal thrust cancellation in equilibrium gravity, which reduces flight efficiency [10, 11].

The addition of a tiltable rotor to a multicopter also enabled independent control of the position and attitude, solving the fixed thrust in the body frame and improving the flight performance [12]. In [13], a design scheme for tilting quadrotors was proposed; simulation and actual experiments verify the design effectiveness of the copter [14]. A model-based tilting multirotor nonlinear inverse dynamics control scheme was reported in [15]. In [12], the proportional-integral-derivative (PID) control method implements a linearized control allocation method.

The tilting multicopter is an aircraft with a rotor that can tilt around its arm axis, featuring strong coupling, uncertainty, and nonlinearity [16]. The tilting multicopter has internal disturbances and external disturbances that cannot be ignored, and its controller design is more difficult than

conventional multicopters. Control algorithms commonly used in flight control research include model-independent methods and model-based methods. PID is a model-independent method, which is widely used and easy to implement. However, the range of parameters adapting to the object is small, and the set parameters cannot be adapted to the disturbance factors outside the system [17]. For example, when the aircraft is subjected to wind disturbance, the quality of control will deteriorate. Model-based methods include robust control [18], backstepping control [19], and sliding mode control [20]. Meanwhile, graph theory can be used to estimate the observations of multicopters [21–23]. All these kinds of model-based methods can solve some internal and external disturbance problems, but the control performance has a high dependence on the accuracy of the model and is limited by environmental factors. It is difficult to obtain accurate aerodynamic performance parameters of the aircraft. And the model-based control algorithm is difficult to implement and has a poor real-time performance.

Active disturbance rejection control (ADRC) [24] is a method that does not depend on the precise plant model. It can estimate internal disturbances and external disturbances in real time and compensate them [21–23]. ADRC has the characteristics of small calculation, easy implementation, and wide range of controller parameters [25]. Researchers have studied ADRC on conventional quadrotors [17, 26, 27], verifying the immunity and robustness of the controller. In order to improve the adaptability of the copter to external environmental disturbances and uncertain internal factors, we design a flight controller based on the ADRC method. The contributions are summarized as follows.

Firstly, a six degree-of-freedom dynamic model under wind disturbance is established based on the Newton–Eulerian method and rotor slip theory. A six-channel single-loop omnidirectional ADRC controller is designed for the position and attitude control decoupling of the tilting hexacopter. Then, a method for implementing variable substitution and the linearization control distribution matrix by constructing virtual control quantity is given. Finally, the simulation experiments verify the effectiveness of the designed omnidirectional ADRC controller and verify the immunity and robustness of the whole system.

## 2. Symbols

The symbols used in this article are collected in Table 1. The table includes notation, reference frames, system-wide parameters, and symbols related to the modeling.

## 3. System Overview

**3.1. Kinematic Modeling.** We consider a tilting six-rotor copter that is modeled as a rigid body in Figure 1 and making use of a world frame  $\mathcal{F}_E$  with an orthonormal basis  $\mathcal{F}_E: \{O_E, X_E, Y_E, Z_E\}$  represented in world frames. A body frame  $\mathcal{F}_B$  with an orthonormal basis  $\mathcal{F}_B: \{O_B, X_B, Y_B, Z_B\}$  is also represented in world frames. Besides six frames  $\mathcal{F}_{r,i}: \{O_{r,i}, X_{r,i}, Y_{r,i}, Z_{r,i}\}, i = 1, \dots, 6$ , fixed to the rotors defined, the origin of the frame  $O_{r,i}$  is fixed to the motor,  $X_{r,i}$  is the tilting axis of the rotor, and  $Z_{r,i}$  is opposite to the thrust of the rotor.

${}^E R_B \in SO(3)$  is a rotation matrix representing the body frame to the world frame, and it can be written as

$${}^E R_B = \begin{bmatrix} \cos\theta\cos\psi & \sin\phi\sin\theta\cos\psi - \cos\phi\sin\psi & \cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi \\ \cos\theta\sin\psi & \sin\phi\sin\theta\sin\psi + \cos\phi\cos\psi & \cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi \\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix}, \quad (1)$$

and  ${}^B R_{r,i}$  is a rotation matrix representing the  $i$ -th rotor frame to the body frame and can be written as

$${}^B R_{r,i} = R_z\left(\frac{\pi}{6} - \frac{i\pi}{3}\right)R_x(-\alpha_i), \quad (2)$$

where  $\alpha_i$  is the tilt angle of the  $i$ -th rotor around  $X_{r,i}$ .

$$R_z(\cdot) = \begin{pmatrix} \cos(\cdot) & \sin(\cdot) & 0 \\ -\sin(\cdot) & \cos(\cdot) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (3)$$

$$R_x(\cdot) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\cdot) & \sin(\cdot) \\ 0 & -\sin(\cdot) & \cos(\cdot) \end{pmatrix}.$$

$O_{r,i}^B$  is the frame of  $O_{r,i}$  in the body frame:

$$O_{r,i}^B = \begin{bmatrix} l \cos\left(-\frac{\pi}{6} + \frac{i\pi}{3}\right) \\ l \sin\left(-\frac{\pi}{6} + \frac{i\pi}{3}\right) \\ 0 \end{bmatrix}, \quad i = 1, \dots, 6, \quad (4)$$

where  $l$  is the distance from the rotor center to the center of mass (COM) of the copter.

In this work, we adopt the rigid body kinematics model presented in the following:

TABLE 1: Symbols mentioned in the article.

Symbols	Meaning	Units
$\theta, \phi, \psi$	Roll, pitch, and yaw Euler angles	deg or rad
$\eta$	Vector of Euler angles	deg or rad
$P$	Position of the hexacopter	m
$V$	Velocity of the hexacopter	$\text{m}\cdot\text{s}^{-1}$
$\omega$	Rotation rate of the hexacopter	$\text{deg}\cdot\text{s}^{-1}$ or $\text{rad}\cdot\text{s}^{-1}$
$n_i$	Rotation velocity of the $i$ -th propellers	$\text{deg}\cdot\text{s}^{-1}$ or $\text{rad}\cdot\text{s}^{-1}$
$F$	Force acting on the hexacopter	N
$M$	Torque acting on the hexacopter	N·m
$\mathcal{F}_E$	Inertial world-fixed frame	—
$\mathcal{F}_B$	Body-fixed frame	—
$\mathcal{F}_{R,i}$	$i$ -th rotor frame	—
${}^{(\cdot)}R_{(\cdot)}$	The rotation matrix between different frames	—
$M$	Total mass of the hexacopter	kg
$J$	Total inertia matrix of the hexacopter	N·m
$g$	Gravitational acceleration	$\text{m}\cdot\text{s}^{-2}$

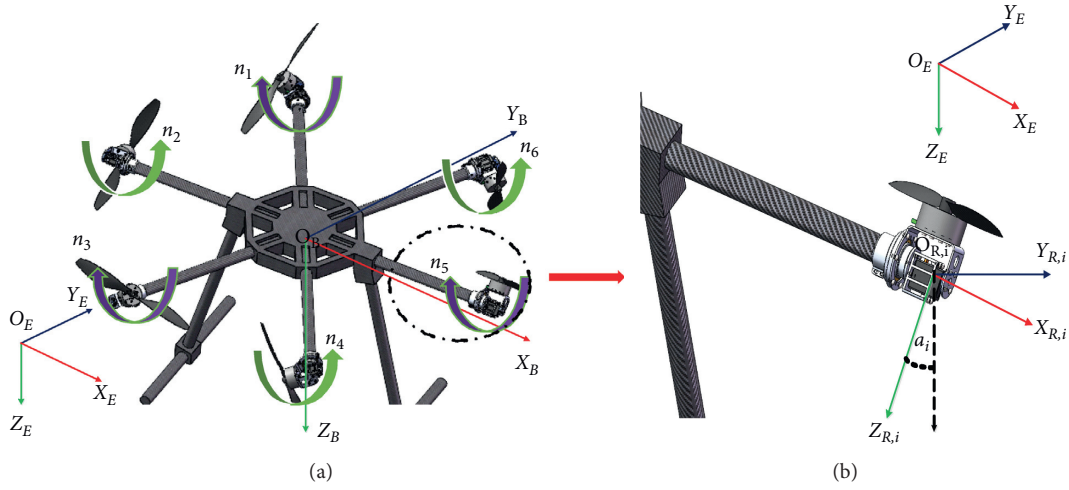


FIGURE 1: Frames of the tilting hexacopter.

$$\begin{aligned} \dot{P}^E &= V^E, \\ \dot{\eta} &= W\dot{\omega}^B, \end{aligned} \quad (5)$$

where  $P^E = [x \ y \ z]^T$  and  $V^E = [u \ v \ w]^T$  represent the position and the velocity of the aircraft in the world frame,  $\eta = [\phi \ \theta \ \psi]^T$  is a vector of Euler angles,  $\omega^B = [p \ q \ r]^T$  is the rotation rate under the body frame, and

$$W = \begin{bmatrix} 1 & \tan \theta \sin \phi & \tan \theta \cos \phi \\ 0 & \cos \phi & -\sin \phi \\ 0 & \frac{\sin \phi}{\cos \theta} & \frac{\cos \phi}{\cos \theta} \end{bmatrix}. \quad (6)$$

**3.2. Dynamical Modeling.** This aircraft is a six degree-of-freedom (DOF) rigid body, and the dynamics model can be derived from the Newton–Eulerian equation [13]. The translation equation is described in the world frame, and the rotational equation is described in the body frame.

$$\begin{bmatrix} mI_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & J_b \end{bmatrix} \begin{bmatrix} \dot{V}^E \\ \dot{\omega}^B \end{bmatrix} + \begin{bmatrix} 0 \\ \omega^B \times (J_b \omega^B) \end{bmatrix} = \begin{bmatrix} F^E \\ M^B \end{bmatrix}, \quad (7)$$

where  $m$  is the total mass,  $J_b = \text{diag}(I_{xx}, I_{yy}, I_{zz})$  is the inertia matrix,  $O_{3 \times 3}$  is the zero matrix, and  $I_{3 \times 3}$  is the unit matrix.

The force  $F^E$  analysis of the aircraft is composed of gravity  $F_m^E$ , the thrust  $F_t^E$  generated by the rotor, wind disturbance  $F_{\text{wind}}^E$ , and air resistance  $F_{\text{aero}}^E$  that can be written as



$$F^E = F_m^E + F_t^E + F_{\text{wind}}^E + F_{\text{aero}}^E, \quad (8) \quad \text{where } F_m^E = [0 \ 0 \ mg]^T \text{ and}$$

$$F_t^E = {}^E R_B \sum_{i=1}^6 ({}^B R_{r,i} F_{r,i})$$

$$= {}^E R_B \begin{bmatrix} \frac{F_{r,1}}{2} \sin \alpha_1 + F_{r,2} \sin \alpha_2 + \frac{F_{r,3}}{2} \sin \alpha_3 - \frac{F_{r,4}}{2} \sin \alpha_4 - F_{r,5} \sin \alpha_5 - \frac{F_{r,6}}{2} \sin \alpha_6 \\ \frac{\sqrt{3} F_{r,1}}{2} \sin \alpha_1 + \frac{\sqrt{3} F_{r,3}}{2} \sin \alpha_3 + \frac{\sqrt{3} F_{r,4}}{2} \sin \alpha_4 - \frac{\sqrt{3} F_{r,5}}{2} \sin \alpha_6 \\ F_{r,1} \cos \alpha_1 + F_{r,2} \cos \alpha_2 + F_{r,3} \cos \alpha_3 + F_{r,4} \cos \alpha_4 + F_{r,5} \cos \alpha_5 + F_{r,6} \cos \alpha_6 \end{bmatrix}, \quad (9)$$

where  $F_{r,i}$  is the lift generated by the  $i$ -th rotor, which is under the  $i$ -th rotor frame.

The external torque  $M^B$  includes the moment  $M_t^B$  generated by the rotor thrust, the rotor rotation counter-torque  $M_{\text{anti}}^B$ , the rotor tilting reverse torque  $M_{\alpha}^B$ , air resistance  $M_{\text{wind}}^B$ , and the gyro effect  $M_{\text{gyro}}^B$ . So, we have

$$M^B = M_t^B + M_{\text{anti}}^B + M_{\alpha}^B + M_{\text{wind}}^B + M_{\text{gyro}}^B. \quad (10)$$

The torque generated by the rotor thrust under the body frame is

$$M_t^B = \sum_{i=1}^6 (O_{r,i}^B \times {}^B R_{r,i} F_{r,i})$$

$$= l \begin{bmatrix} \frac{F_{r,1}}{2} \cos \alpha_1 + F_{r,2} \cos \alpha_2 + \frac{F_{r,3}}{2} \cos \alpha_3 - \frac{F_{r,4}}{2} \cos \alpha_4 - F_{r,5} \cos \alpha_5 - \frac{F_{r,6}}{2} \cos \alpha_6 \\ -\frac{\sqrt{3} F_{r,1}}{2} \cos \alpha_1 + \frac{\sqrt{3} F_{r,3}}{2} \cos \alpha_3 + \frac{\sqrt{3} F_{r,4}}{2} \cos \alpha_4 - \frac{\sqrt{3} F_{r,6}}{2} \cos \alpha_6 \\ -F_{r,1} \sin \alpha_1 - F_{r,2} \sin \alpha_2 - F_{r,3} \sin \alpha_3 - F_{r,4} \sin \alpha_4 - F_{r,5} \sin \alpha_5 - F_{r,6} \sin \alpha_6 \end{bmatrix}, \quad (11)$$

$$M_{\text{anti}}^B = \sum_{i=1}^6 ({}^B R_{r,i} M_{r,i})$$

$$= \begin{bmatrix} \frac{M_{r,1}}{2} \sin \alpha_1 + M_{r,2} \sin \alpha_2 + \frac{M_{r,3}}{2} \sin \alpha_3 - \frac{M_{r,4}}{2} \sin \alpha_4 - M_{r,5} \sin \alpha_5 - \frac{M_{r,6}}{2} \sin \alpha_6 \\ -\frac{\sqrt{3} M_{r,1}}{2} \sin \alpha_1 + \frac{\sqrt{3} M_{r,3}}{2} \sin \alpha_3 + \frac{\sqrt{3} M_{r,4}}{2} \sin \alpha_4 - \frac{\sqrt{3} M_{r,6}}{2} \sin \alpha_6 \\ M_{r,1} \cos \alpha_1 + M_{r,2} \cos \alpha_2 + M_{r,3} \cos \alpha_3 + M_{r,4} \cos \alpha_4 + M_{r,5} \cos \alpha_5 + M_{r,6} \cos \alpha_6 \end{bmatrix}, \quad (12)$$

$$M_{\alpha}^B = J_{\alpha} \begin{bmatrix} \frac{\sqrt{3} \ddot{\alpha}_1}{2} - \frac{\sqrt{3} \ddot{\alpha}_3}{2} - \frac{\sqrt{3} \ddot{\alpha}_4}{2} + \frac{\sqrt{3} \ddot{\alpha}_6}{2}, \frac{\ddot{\alpha}_1}{2} + \ddot{\alpha}_2 + \frac{\ddot{\alpha}_3}{2} - \frac{\ddot{\alpha}_4}{2} - \ddot{\alpha}_5 - \frac{\ddot{\alpha}_6}{2}, 0 \end{bmatrix}^T. \quad (13)$$

Gyro effect terms include two parts caused by attitude changes and motor tilt.

$$\begin{aligned}
M_{gyro}^B &= J_p \sum_{i=1}^6 \left( \left( {}^B R_{r,i} \begin{bmatrix} \dot{\alpha}_i \\ 0 \\ 0 \end{bmatrix} + \Omega^B \right) \times (-1)^i \cdot {}^B R_{r,i} \begin{bmatrix} 0 \\ 0 \\ n_i \end{bmatrix} \right) \\
&= J_p \begin{bmatrix} -n_1 \cos \alpha_1 \left( \frac{\dot{\alpha}_1}{2} + q \right) + n_2 \cos \alpha_2 (\dot{\alpha}_2 + q) - n_3 \cos \alpha_3 \left( \frac{\dot{\alpha}_3}{2} + q \right) - n_4 \cos \alpha_4 \left( \frac{\dot{\alpha}_4}{2} - q \right) \\ + n_5 \cos \alpha_5 (\dot{\alpha}_5 - q) - n_6 \cos \alpha_6 \left( \frac{\dot{\alpha}_6}{2} - q \right) - \frac{\sqrt{3}}{2} n_1 r \sin \alpha_1 + \frac{\sqrt{3}}{2} n_3 r \sin \alpha_3 \\ - \frac{\sqrt{3}}{2} n_4 r \sin \alpha_4 + \frac{\sqrt{3}}{2} n_6 r \sin \alpha_6 - \frac{r n_1 \sin \alpha_1}{2} + r n_2 \sin \alpha_2 - \frac{r n_3 \sin \alpha_3}{2} \\ \frac{r n_4 \sin \alpha_4}{2} + r n_5 \sin \alpha_5 - \frac{r n_6 \sin \alpha_6}{2} + n_1 \cos \alpha_1 \left( \frac{\sqrt{3}}{2} \dot{\alpha}_1 + p \right) - p n_2 \cos \alpha_2 \\ + n_3 \cos \alpha_3 \left( -\frac{\sqrt{3}}{2} \dot{\alpha}_3 + p \right) - n_4 \cos \alpha_4 \left( -\frac{\sqrt{3}}{2} \dot{\alpha}_4 + p \right) + p n_5 \cos \alpha_5 \\ - n_6 \cos \alpha_6 \left( \frac{\sqrt{3}}{2} \dot{\alpha}_6 + p \right) \frac{1}{2} n_1 \sin \alpha_1 \left( \frac{\dot{\alpha}_1}{2} + q \right) - n_2 \sin \alpha_2 (\dot{\alpha}_2 + q) \\ + \frac{1}{2} n_3 \sin \alpha_3 \left( \frac{\dot{\alpha}_3}{2} + q \right) - \frac{1}{2} n_4 \sin \alpha_4 \left( \frac{\dot{\alpha}_4}{2} - q \right) + n_5 \sin \alpha_5 (\dot{\alpha}_5 - q) \\ + \frac{1}{2} n_6 \sin \alpha_6 \left( \frac{\dot{\alpha}_6}{2} - q \right) + \frac{\sqrt{3}}{2} n_1 \sin \alpha_1 \left( \frac{\sqrt{3}}{2} \dot{\alpha}_1 + p \right) - \frac{\sqrt{3}}{2} n_3 \\ \sin \alpha_3 \left( -\frac{\sqrt{3}}{2} \dot{\alpha}_3 + p \right) + \frac{\sqrt{3}}{2} n_4 \sin \alpha_4 \left( -\frac{\sqrt{3}}{2} \dot{\alpha}_4 + p \right) - \frac{\sqrt{3}}{2} n_6 \sin \alpha_6 \left( \frac{\sqrt{3}}{2} \dot{\alpha}_6 + p \right) \end{bmatrix} \quad (14)
\end{aligned}$$

In this paper, as shown in Figure 1, the 2, 4, and 6 rotors rotate clockwise, and the 1, 3, and 5 rotors rotate counterclockwise. Therefore, the vector form of the thrust vector  $\widehat{F}_{r,i}$  and the antitorque vector  $\widehat{M}_{r,i}$  generated by the  $i$ -th motor in the corresponding rotor frame system is

$$\begin{aligned}
\widehat{F}_{r,i} &= [0 \ 0 \ F_{r,i}]^T, \\
\widehat{M}_{r,i} &= [0 \ 0 \ M_{r,i}]^T, \\
F_{r,i} &= -k_t n_i^2, \\
M_{r,i} &= \begin{cases} -k_q n_i^2, & i = 2, 4, 6, \\ k_q n_i^2, & i = 1, 3, 5, \end{cases} \quad k_t, k_q > 0, \quad (15)
\end{aligned}$$

where  $k_t$  denotes the rotor thrust coefficient and  $k_q$  is the antitorque coefficient.

**3.3. Wind Disturbance Modeling.** The wind field has two main effects on the aircraft. First, it affects the aerodynamic

effect of the rotor and then affects the air resistance of each windward surface.

The aerodynamic analysis of the rotor under windy conditions is shown in Figure 2. In the figure,  $V_w$  represents the wind velocity,  $V_{t,i}$  represents the rotor induction velocity, and  $\widehat{V}$  is the vector sum of the wind velocity and the induced velocity. According to the theory of rotor slip flow [25], the velocity calculation formula is as follows:

$$\|V_{t,i}\| = \sqrt{\frac{k_t \cdot n_i^2}{2\rho\pi r^2}}, \quad (16)$$

where  $\rho$  is the air density and  $r$  is the radius of the rotor paddle. The total lift of the rotor [25] can be expressed as

$$\|F_{r,i}\| + \|F_{w,i}\| = 2\pi\rho r^2 \|V_{t,i}\| \|V_w + V_{t,i}\|. \quad (17)$$

The wind disturbance  $F_{w,i}$  and torque of  $i$ -th rotor  $M_{w,i}$  are as follows:

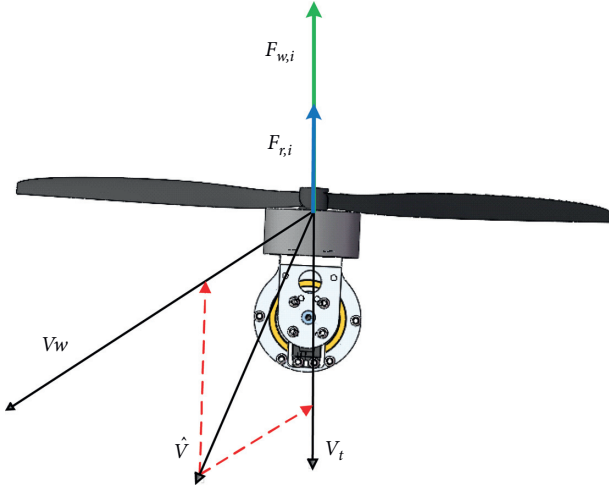


FIGURE 2: Aerodynamic analysis of the rotor under wind.

$$F_{w,i} = k_t n_i^2 - 2\rho\pi r^2 \left\| \begin{bmatrix} 0 \\ 0 \\ \sqrt{\frac{k_f n_i^2}{2\rho\pi r^2}} \end{bmatrix} + ({}^B R_{r,i})^T ({}^E R_B)^T V_w^E \right\|, \quad (18)$$

$$M_{w,i} = \begin{cases} \frac{k_m F_{w,i}}{k_t}, & i = 1, 3, 5, \\ -\frac{k_m F_{w,i}}{k_t}, & i = 2, 4, 6. \end{cases}$$

So,

$$F_{\text{wind}}^E = {}^E R_B \sum_{i=1}^6 {}^B R_{r,i} \begin{bmatrix} 0 \\ 0 \\ F_{w,i} \end{bmatrix}, \quad (19)$$

$$M_{\text{wind}}^B = \sum_{i=1}^6 {}^B R_{r,i} \begin{bmatrix} 0 \\ 0 \\ M_{r,i} \end{bmatrix}.$$

The formula for air resistance is

$$D = \frac{1}{2} c \rho S V_{\text{air}}^2, \quad (20)$$

where  $c$  is the air resistance coefficient and  $V_{\text{air}}$  is the relative velocity of the aircraft and air. For the aircraft, there is  $V_{\text{air}} = V_w^E - V^E$ . When calculating the air resistance of the copter, consider it as a cylinder and take the average windward area  $\bar{S}$  as

$$\bar{S} = \sigma 2\pi l h + (1 - \sigma)\pi l^2, \quad (21)$$

where  $\sigma \in (0, 1)$  is the crosswind coefficient and  $h$  is the height of the body, so that

$$F_{\text{aero}}^E = \frac{1}{2} c \rho \bar{S} V_{\text{air}} \|V_{\text{air}}\|. \quad (22)$$

## 4. Control

The ADRC [24] does not depend on the accurate model of the controlled object, has strong resistance to internal and external disturbances, and has good robustness. The complete algorithm includes the tracking differentiator (TD) [28], extended state observer (ESO) [27], and nonlinear state error feedback (NLSEF) [29]. The standard ADRC structure is shown in Figure 3.

**4.1. Controller Structure.** The copter is divided into two subsystems, one of which is a full-drive subsystem including a height position  $z$  and a yaw angle  $\psi$ , and the other is an underactuation subsystem consisting of a horizontal position  $(x, y)$ , a roll angle  $\phi$ , and a pitch angle  $\theta$ . Conventional multicopters can only achieve six degree-of-freedom motion in a nonholonomic sense, and they need to change the roll and pitch angles while performing horizontal motion. When designing a controller based on a model-free control technology (such as PID), it is usually only possible to design a controller for the height channel and the yaw channel separately and then design a cascade controller for the underdrive subsystem to output the horizontal channel controller. The attitude angle is expected to be the input of the roll and pitch channel controllers to achieve horizontal position control. The tilting hexacopter in this paper is an overdrive system that can achieve six-degree-of-freedom independent motion, and its ADRC controller can achieve position and attitude control through a six-channel single-loop structure.

From equation (1), coupling and tilting angles exist in the three attitude channels of roll, pitch, and yaw, respectively. The ADRC can treat the interaction between different channels and the tilt angle effect as the internal disturbance of the system. The ESO of each channel independently estimates the real-time internal disturbance of the system and the external environmental disturbance such as wind disturbance and air resistance as the control compensation items. Therefore, decoupling independent control of each channel can be realized. And through the control allocation, the desired force and the desired torque in any direction can be directly mapped to the rotor velocities  $n_i$  and  $\alpha_i$  so that the controller can be designed as a decoupled six-channel single-loop structure.

Rewriting the dynamic model in equation (7) to the form corresponding to the ADRC theory,

$$\begin{cases} \ddot{x} = f_1(x, \dot{x}, \mu_1(t)) + b_1 U_1, \\ \ddot{y} = f_2(y, \dot{y}, \mu_2(t)) + b_2 U_2, \\ \ddot{z} = f_3(z, \dot{z}, \mu_3(t)) + b_3 U_3 - g, \\ \ddot{\phi} = f_4(\phi, \dot{\phi}, \theta, \dot{\theta}, \psi, \dot{\psi}, \alpha, \dot{\alpha} + \mu_4(t)) + b_4 U_4, \\ \ddot{\theta} = f_5(\phi, \dot{\phi}, \theta, \dot{\theta}, \psi, \dot{\psi}, \alpha, \dot{\alpha} + \mu_5(t)) + b_5 U_5, \\ \ddot{\psi} = f_6(\phi, \dot{\phi}, \theta, \dot{\theta}, \psi, \dot{\psi}, \alpha, \dot{\alpha} + \mu_6(t)) + b_6 U_6. \end{cases} \quad (23)$$

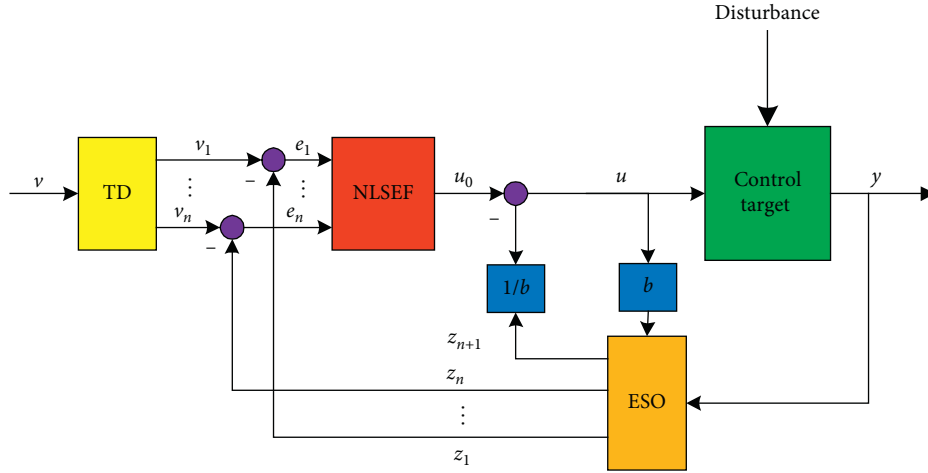


FIGURE 3: Standard ADRC structure.

$f_i(\cdot)$  and  $\mu_i(t)$  are uncertain parts, and

$$\begin{aligned} b_1 = b_2 = b_3 &= \frac{1}{m}, \\ b_4 &= \frac{1}{I_{xx}}, \\ b_5 &= \frac{1}{I_{yy}}, \\ b_6 &= \frac{1}{I_{zz}}, \\ U_F^E &= [U_1 \ U_2 \ U_3]^T, \\ U_M^B &= [U_4 \ U_5 \ U_6]^T. \end{aligned} \quad (24)$$

**4.2. Controller Design.** As shown in Figure 4, the controller designed in this paper has six single-loop structures and is decoupled from each other.

We choose the yaw channel as an example. The details of TD, ESO, and NLSEF in the ADRC controller are given, respectively.

- (1) In the TD, we use the given signal  $\psi_{\text{des}}$  as a desired input to arrange the transition process. Figure 4 shows the ADRC structure.

$$\begin{cases} e = v_1 - \psi_{\text{des}}, \\ fh = \text{fhan}(e, v_2, r, h), \\ v_1 = v_1 + hv_2, \\ v_2 = v_2 + hfh. \end{cases} \quad (25)$$

Fast factor  $r$  and filter factor  $h$  are adjustable parameters.

- (2) In the ESO, the system's output  $\psi$  and input  $U_6$  are tracked in real time to estimate the internal state of the system and internal and external disturbances.

$$\begin{cases} e = z_1 - \psi, \\ fe = \text{fal}(e, 0.5, \delta), \\ fe_1 = \text{fal}(e, 0.25, \delta), \\ z_1 = z_1 + h(z_2 - \beta_{01}e) \\ z_2 = z_2 + h(z_3 - \beta_{02}fe + b_6U_6), \\ z_3 = z_3 + h(-\beta_{03}fe_1), \end{cases} \quad (26)$$

where  $\beta_{01}$ ,  $\beta_{02}$ , and  $\beta_{03}$  are parameters.

- (3) And in the NLSEF,

$$\begin{cases} e_1 = v_1 - z_1, \\ e_2 = v_2 - z_2 \\ u_0 = k_1 \text{fal}(e_1, \alpha_1, \delta_0) + k_2 \text{fal}(e_2, \alpha_2, \delta_0), \\ u = u_0 - \frac{z_3}{b_6}. \end{cases} \quad (27)$$

There are 6 tunable parameters in the NLSEF, namely, compensation coefficient  $b_6$ , controller gains  $k_1$  and  $k_2$ , and nonlinear parameters  $\alpha_1$ ,  $\alpha_2$ , and  $\delta_0$ .

The fastest tracking control synthesis function  $\text{fhan}(x_1, x_2, r, h)$  and nonlinear function  $\text{fal}(e, \alpha, \delta)$  are expressed as

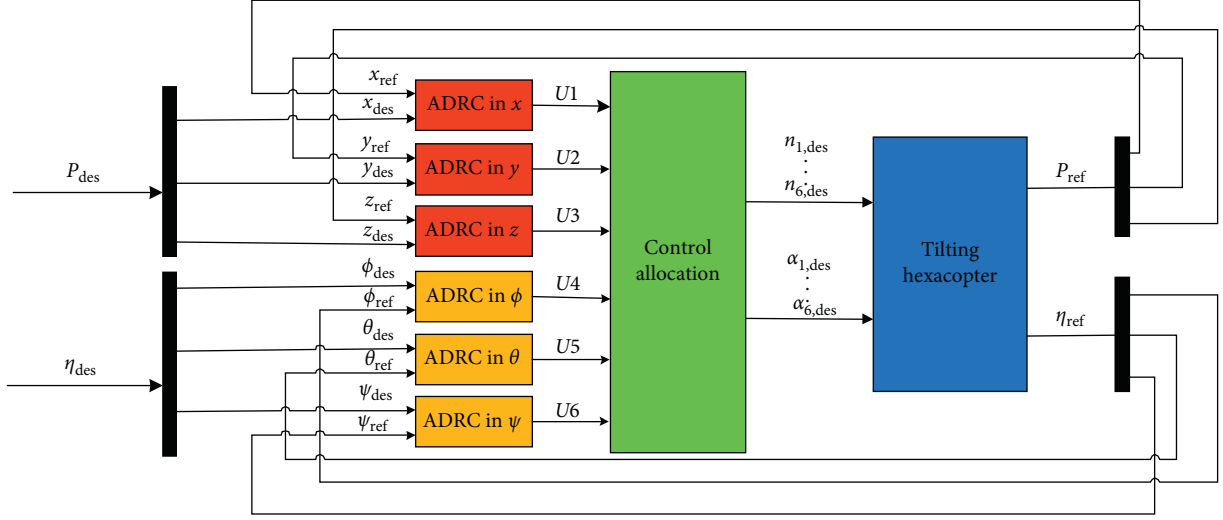


FIGURE 4: Controller diagram via ADRC.

$$\text{fhan}(x_1, x_2, r, h) = - \begin{cases} r \text{sign}(a), & |a| > d, \\ r \frac{a}{d}, & |a| \leq d, \end{cases}$$

$$d_0 = rh^2,$$

$$y = x_1 + hx_2,$$

$$a_0 = \sqrt{d^2 + 8r|y|},$$

$$a = \begin{cases} x_2 + \frac{(a_0 - d)}{2} \text{sign}(y), & |y| > d_0, \\ x_2 + \frac{y}{h}, & |y| \leq d_0. \end{cases}$$
(28)

In the actual system, the state of the feedback back-end will contain high-frequency noise, so low-pass filtering is generally done to remove noise. In general, the multirotor dual-loop control strategy requires two low-pass filters to increase the system phase angle lag and reduce the system bandwidth. In our work, the tracking differentiator maintains a phase angle of  $90^\circ$  ahead of the cutoff frequency, reducing the phase angle lag of the controller, and the system bandwidth is strongly correlated with parameter  $r$ . By adjusting the parameters, the closed-loop bandwidth is greater than the double-loop.

**4.3. Control Allocation.** The forces and torques that control the movement of the aircraft are mainly the thrust and antitorison moment generated by the rotor. Combined with equations (2)–(15), the six degree-of-freedom rigid body dynamics model in equation (7) is rewritten as

$$\begin{bmatrix} mI_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & J_b \end{bmatrix} \begin{bmatrix} \dot{V}^E \\ \dot{\omega}^B \end{bmatrix} + \begin{bmatrix} 0 \\ \omega^B \times (J_b \omega^B) \end{bmatrix} = \begin{bmatrix} F_m^E + F_t^E + F_{\text{wind}}^E + F_{\text{aero}}^E \\ M_t^B + M_{\text{anti}}^B + M_\alpha^B + M_{\text{wind}}^B + M_{\text{gyro}}^B \end{bmatrix}. \quad (29)$$

Define  $U_F^E = F_t^E$  and  $U_M^B = M_t^B + M_{\text{anti}}^B$ . Therefore, the relationship between  $U_F^E, U_M^B$ , the rotor velocity  $n$ , and the rotor tilt angle  $\alpha$  is defined as  $U_F^E = F_t^E, U_M^B = M_t^B + M_{\text{anti}}^B$ .

$$\begin{bmatrix} \left( {}^E R_B \right)^T U_F^E \\ U_M^B \end{bmatrix} = A \begin{bmatrix} n_1^2 \\ n_2^2 \\ n_3^2 \\ n_4^2 \\ n_5^2 \\ n_6^2 \end{bmatrix}, \quad (30)$$

where

$$A = \begin{bmatrix} \frac{-k_f}{2} & 0 & -k_f & 0 & \frac{-lk_f}{2} & 0 & \frac{-k_f}{2} & 0 & k_f & 0 & \frac{-k_f}{2} & 0 \\ -\frac{\sqrt{3}}{2}k_f & 0 & 0 & 0 & -\frac{\sqrt{3}}{2}k_f & 0 & -\frac{\sqrt{3}}{2}k_f & 0 & 0 & 0 & \frac{\sqrt{3}}{2}k_f & 0 \\ 0 & -k_f & 0 & -k_f & 0 & -k_f & 0 & -k_f & 0 & -k_f & 0 & -k_f \\ \frac{k_m}{2} & \frac{-lk_f}{2} & -k_m & -lk_f & \frac{k_m}{2} & \frac{-lk_f}{2} & \frac{k_m}{2} & \frac{-lk_f}{2} & -k_m & lk_f & \frac{k_m}{2} & \frac{-lk_f}{2} \\ -\sqrt{3}k_m & \frac{\sqrt{3}}{2}lk_f & 0 & 0 & \frac{\sqrt{3}}{2}k_m & -\frac{\sqrt{3}}{2}lk_f & -\frac{\sqrt{3}}{2}k_m & -\frac{\sqrt{3}}{2}lk_f & 0 & 0 & \frac{\sqrt{3}}{2}k_m & \frac{\sqrt{3}}{2}lk_f \\ lk_f & k_m & lk_f & -k_m & lk_f & k_m & lk_f & -k_m & lk_f & k_m & lk_f & -k_m \end{bmatrix}, \quad (31)$$

and  $A$  is the control allocation matrix and is constant, independent of the tilt angle  $\alpha_i$ . Define  $N$  as the control quantity, so we have

$$N = A^+ \begin{bmatrix} ({}^E R_B)^T U_F^E \\ U_M^B \end{bmatrix}. \quad (32)$$

By calculating the generalized inverse of equation (32) to calculate the control value, the Moore–Penrose method [30, 31] is selected in this paper to solve. Furthermore, the real control variable rotor speed  $n_i$  and tilt angle  $\alpha_i$  can be directly obtained, namely,

$$\begin{aligned} n_i^2 &= \sqrt{(N_{v,i}^2 + N_{l,i}^2)}, \\ \alpha_i &= a \tan 2(N_{l,i}, N_{v,i}). \end{aligned} \quad (33)$$

## 5. Experiments

The simulation parameters of the tilting hexacopter are shown in Table 2.

The tunable parameters of the TD are  $r$  and  $h$ . The larger the fast factor  $r$  is, the shorter the transition process is. The larger  $h$  is, the better the filtering effect is, but the phase loss is also increased. The parameters of the ESO affect the effect of disturbance estimation and compensation. They can be adjusted according to the empirical formula. When the sampling frequency is the same, the ESO of the six channels of the omnidirectional controller can adopt the same set of parameters. In the NLSEF, the nonlinear parameters  $\alpha_1$ ,  $\alpha_2$ , and  $\delta_0$  have a common satisfaction value, and the compensation coefficient  $b_0$  is related to the object model

and can also be used as a parameter setting when the model is unknown.  $k_1$  and  $k_2$  affect the system response velocity; the larger  $k_1$ , the faster the system response velocity. However, if  $k_1$  is too large, it will easily cause overshoot and oscillation. Increasing  $k_2$  can suppress oscillation and overshoot, but if  $k_2$  is too large, it is easy to brake in advance. This causes the adjustment time to become longer. In the simulation, the control period  $T = 0.001$  s. After multiple-parameter tuning, each channel parameter of the ADRC controller is shown in Table 3. The parameters for ADRC are adjusted by a learning algorithm [32].

**5.1. Attitude and Position Experiment.** Setting the initial position  $P_{\text{init}}^E = [0 \ 0 \ 0]^T$  (m) of the hexacopter, the desired position  $P_{\text{des}}^E = [8 \ 7 \ 11]^T$  (m) and the desired attitude  $\eta_{\text{des}} = [10 \ 15 \ 30]^T$  (deg). After reaching the desired position, keep hovering in the desired attitude. Figure 5 shows the flight trajectory, and Figure 6 shows the synchronized attitude curve.

Compared with the traditional multirotor, Figures 5 and 6 show that the tilting hexacopter with ADRC can fly to the target quickly and smoothly and hover at the target point, and its flight path is shorter. At the same time, the aircraft can maintain the attitude change within a very small range. The figures also show that the attitude of the aircraft can reach the expected value quickly. There are no overshoot and oscillation during the adjustment process. The stability control effect is good, and it can maintain the fixed-point hover.

**5.2. Trajectory Tracking Experiment.** In this simulation, the desired position reference involves a circular trajectory. The desired trajectory can be expressed by the following equation:



TABLE 2: System parameters.

Parameters	Value	Unit
Hexacopter's mass, $m$	2.274	kg
Gravitational acceleration, $g$	9.81	$\text{m} \cdot \text{s}^{-2}$
Inertia tensor in the $x$ -axis, $I_{xx}$	0.516	$\text{kg} \cdot \text{m}^2$
Inertia tensor in the $y$ -axis, $I_{yy}$	0.516	$\text{kg} \cdot \text{m}^2$
Inertia tensor in the $z$ -axis, $I_{zz}$	0.982	$\text{kg} \cdot \text{m}^2$
Lift coefficient, $k_f$	$3.10e-5$	—
Resistance coefficient, $k_m$	$5.61e-7$	—
Moment of inertia in tilting axis, $J_\alpha$	$1.50e-5$	$\text{kg} \cdot \text{m}^2$
Gyro moment coefficient, $J_p$	$6.00e-5$	—
Air density, $\rho$	1.293	$\text{kg} \cdot \text{m}^{-3}$
Average windward area, $\bar{S}$	0.107	$\text{m}^2$
Coefficient of air resistance, $c$	0.08	—

TABLE 3: ADRC controller parameters.

	Parameters	$x$	$y$	$z$	$\phi$	$\theta$	$\psi$
TD	$r$	2	2	2	0.533	0.533	0.533
	$h$	0.001	0.001	0.001	0.001	0.001	0.001
ESO	$\delta$	0.02	0.02	0.02	0.02	0.02	0.02
	$\beta_{01}$	200	200	200	200	200	200
	$\beta_{02}$	500	500	500	500	500	500
	$\beta_{03}$	2500	2500	2500	2500	2500	2500
NLSEF	$b_0$	0.44	0.44	0.44	2.008	2.008	1.0299
	$k_1$	155	155	90	140	140	100
	$k_2$	115	115	150	90	90	100
	$\alpha_1$	0.75	0.75	0.75	0.75	0.75	0.75
	$\alpha_2$	1.25	1.25	1.25	1.25	1.25	1.25
	$\delta_0$	0.2	0.2	0.2	0.2	0.2	0.2

$$\text{trajectory}_{\text{-des}} = \begin{bmatrix} r \sin\left(\frac{\pi}{10}(t-5)\right) \\ r \cos\left(\frac{\pi}{10}(t-5)\right) - 5 \\ -3 \end{bmatrix}, \quad (34)$$

where  $r$  is the radius of the circle formed by the trajectory. The aircraft follows a circular trajectory and maintains a height of 3 m on the  $z$ -axis.

From the results (see Figure 7), the controller can track a circular trajectory. At the same time, the steady-state error is very small during the whole tracking process.

**5.3. Anti-Interference Experiment.** In the actual flight process, the aircraft is often subjected to various disturbances in the external environment, such as wind field disturbances. The wind field disturbance models are mainly Dryden model [33]. In this paper, the Dryden turbulence model is used to simulate the atmospheric turbulence by shaping the standard Gaussian white noise to obtain colored noise. According to Jafar et al. [34], the transfer function of the shaping filter is

$$G_u(s) = \frac{K_u}{T_u s + 1},$$

$$K_u = \sigma_u \sqrt{\frac{L_u}{\pi v}},$$

$$T_u = \frac{L_u}{v},$$

$$G_v(s) = \frac{K_v}{T_v s + 1},$$

$$K_v = \sigma_v \sqrt{\frac{L_v}{\pi v}}, \quad (35)$$

$$T_v = \frac{2L_v}{\sqrt{3}v},$$

$$G_w(s) = \frac{K_w}{T_w s + 1},$$

$$K_w = \sigma_w \sqrt{\frac{L_w}{\pi v}},$$

$$T_w = \frac{2L_w}{\sqrt{3}v},$$

where  $L_u, L_v, L_w$  and  $\sigma_u, \sigma_v, \sigma_w$  represent turbulence intensities and turbulence scales. Multicopters are mainly low-altitude flights, and the turbulence intensity and scale calculation formula under the flight conditions are shown as

$$L_u = 2L_v = \frac{h}{(0.177 + 0.000823h)^{12}},$$

$$L_w = \frac{h}{2},$$

$$\sigma_u = \frac{\sigma_w}{(0.177 + 0.000833h)^{0.4}}, \quad (36)$$

$$\sigma_v = \frac{\sigma_u}{(0.177 + 0.000823h)^{0.4}},$$

$$\sigma_w = 0.1V_h,$$

where  $V_h$  is the wind velocity at height  $h$ . The turbulent wind simulation module is built in the MATLAB/Simulink environment. The turbulent flow parameters are shown in Table 4. When the continuous wind is added, the turbulent wind field formed is shown in Figure 8.

The immunity of the ADRC controller under the fixed-point hover is tested under three different wind disturbance conditions and then compared with the PID controller. The

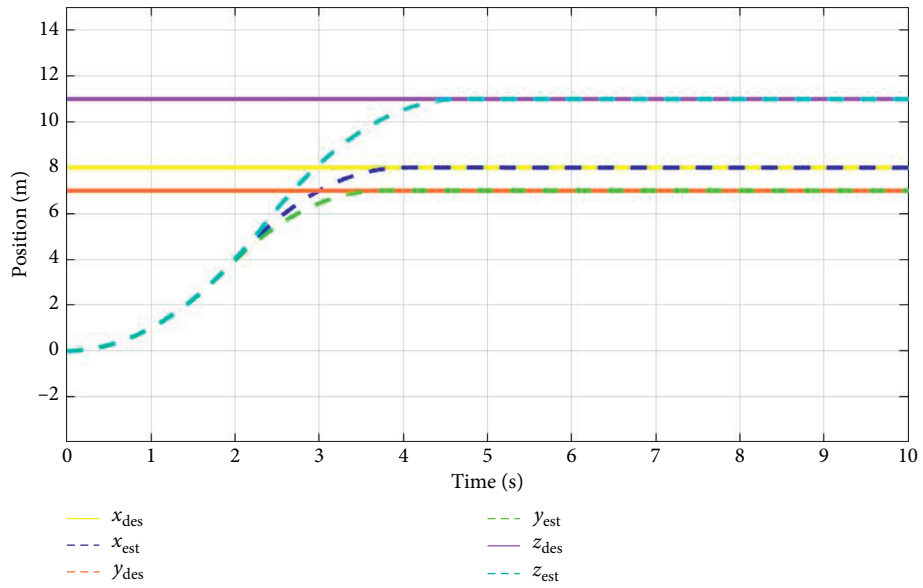


FIGURE 5: Path curve of position-independent control.

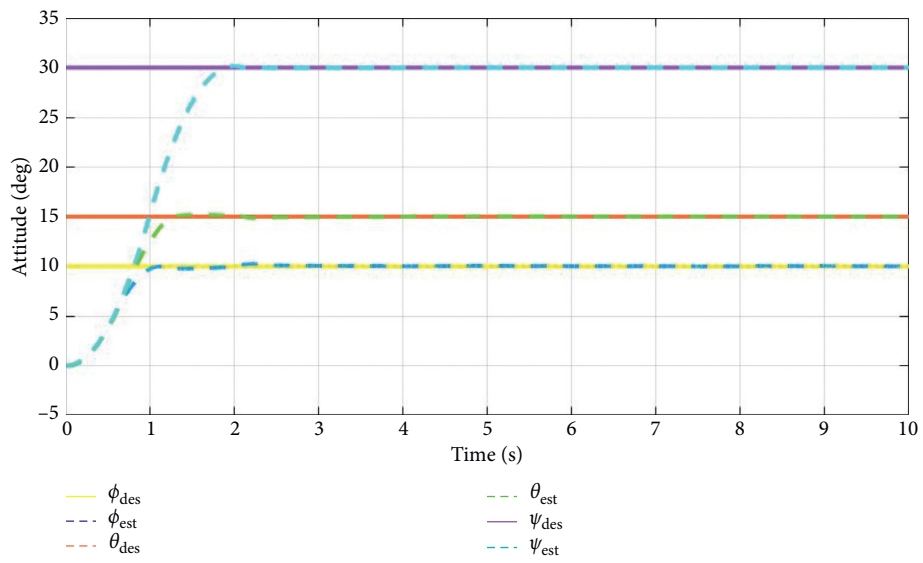


FIGURE 6: Orientation curve of attitude-independent control.

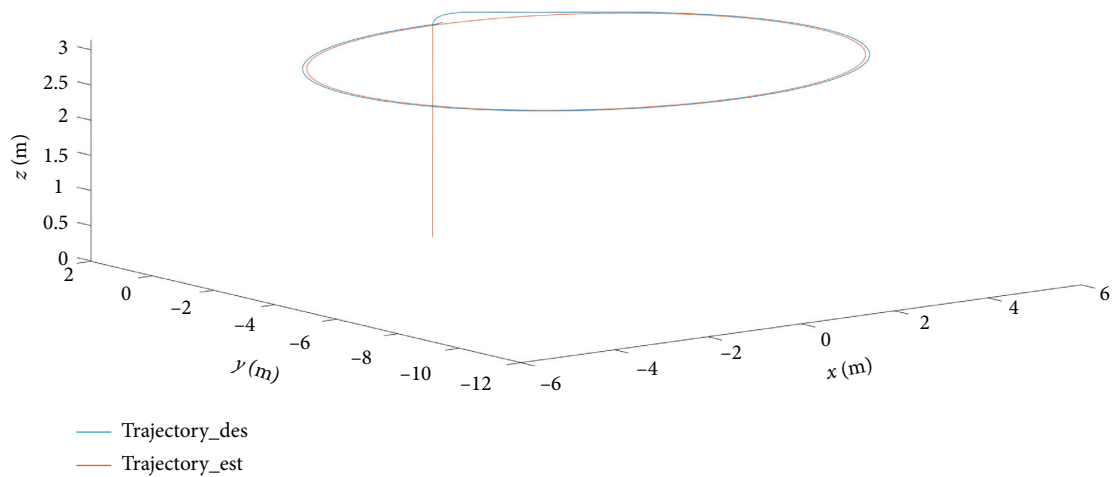


FIGURE 7: Trajectory tracking output.

TABLE 4: Turbulent flow parameters.

Parameters	Value
$x$ -axis scale factor, $L_u$	110.5
$y$ -axis scale factor, $L_v$	55.3
$z$ -axis scale factor, $L_w$	7.5
$x$ -axis intensity factor, $\sigma_u$	0.5
$y$ -axis intensity factor, $\sigma_v$	0.81
$z$ -axis intensity factor, $\sigma_w$	1.31

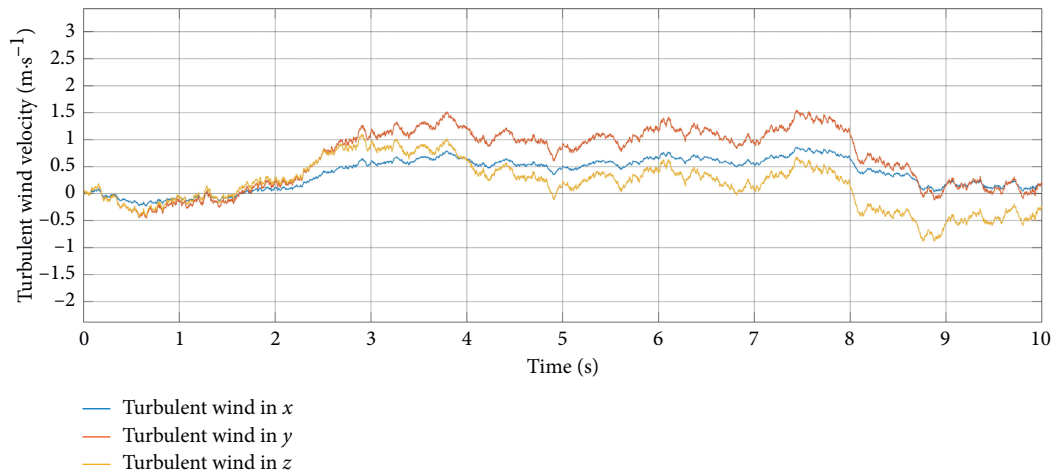
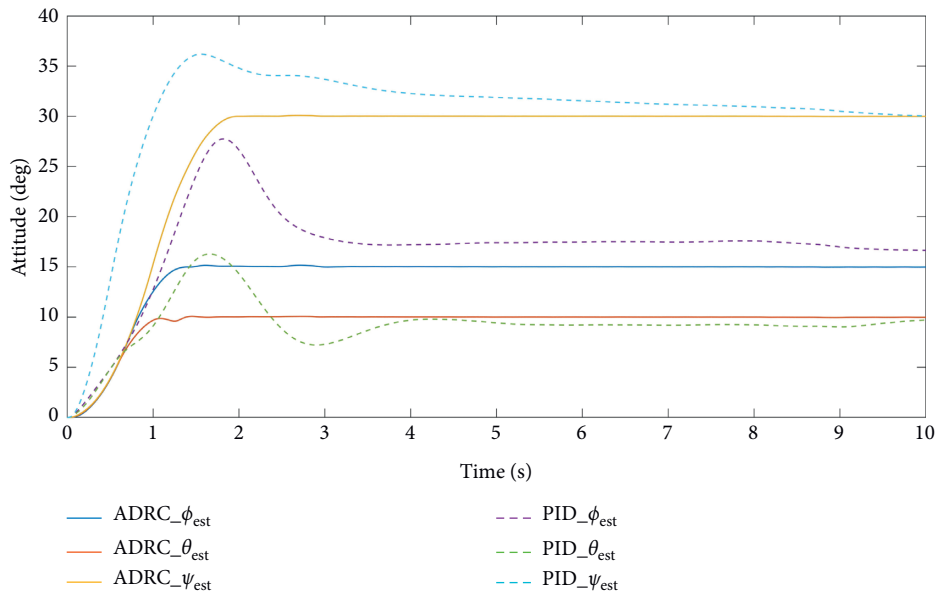


FIGURE 8: Turbulent wind field.



(a)

FIGURE 9: Continued.

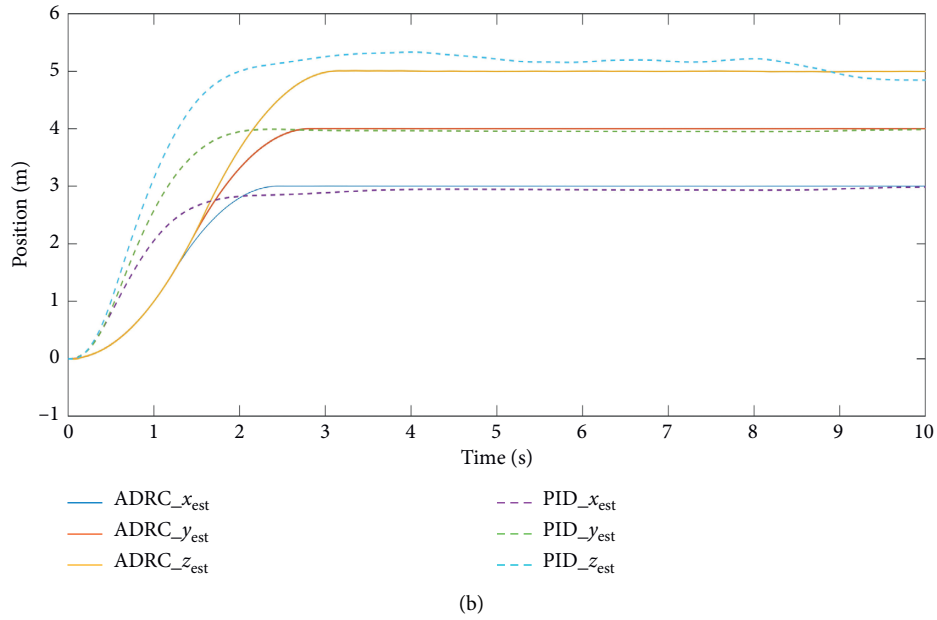


FIGURE 9: Attitude and position without wind.

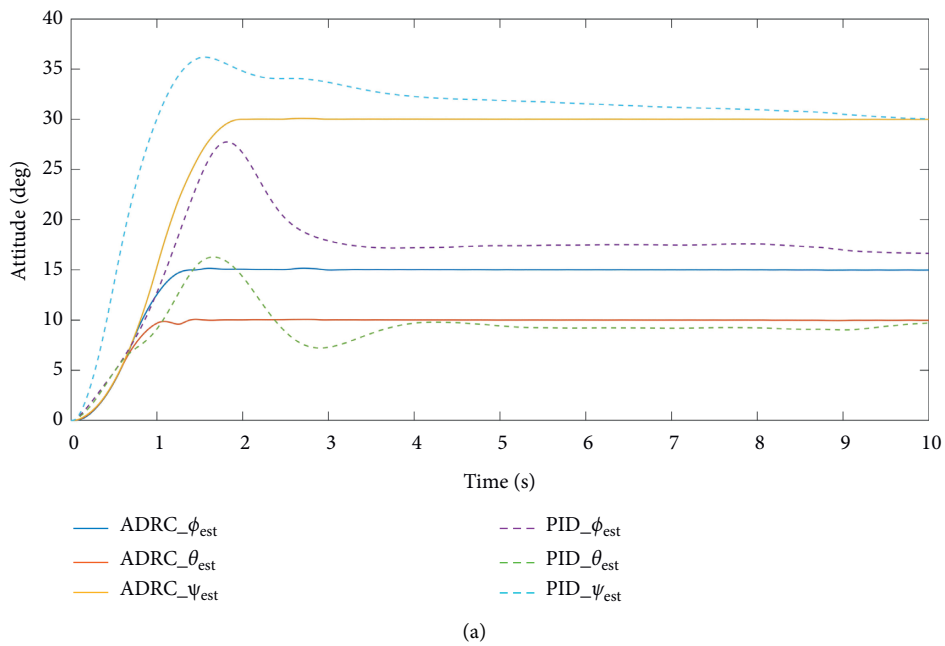


FIGURE 10: Continued.

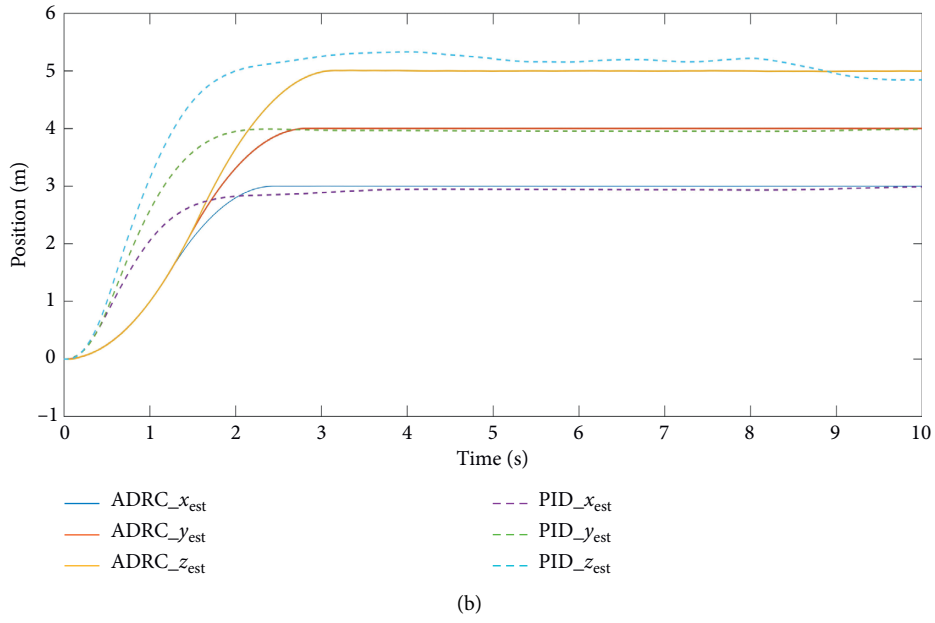


FIGURE 10: Attitude and position under turbulent wind and no continuous wind.

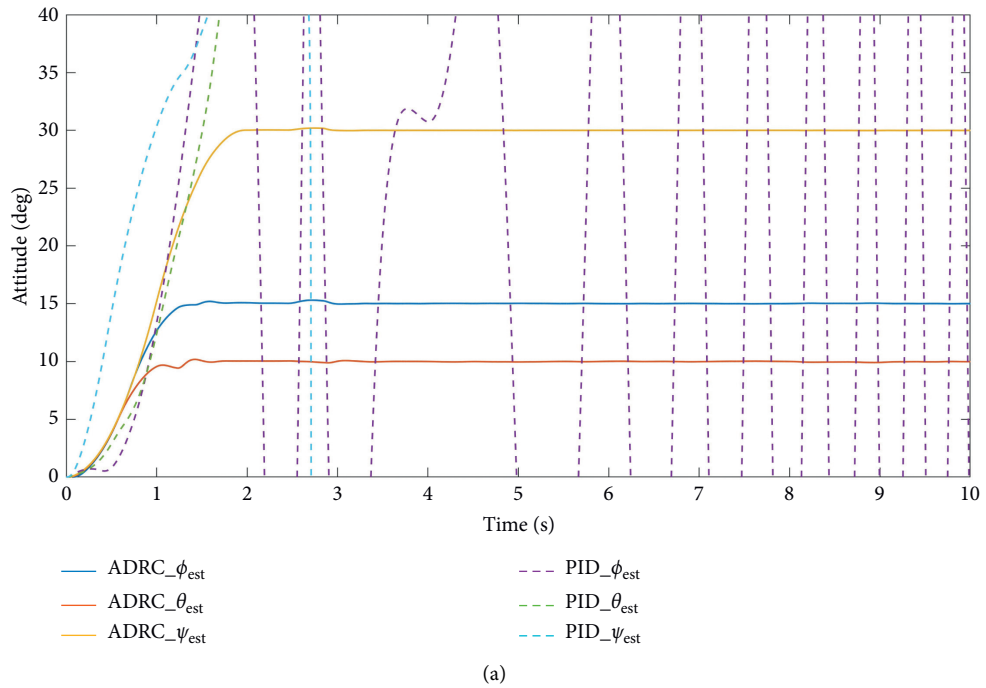


FIGURE 11: Continued.

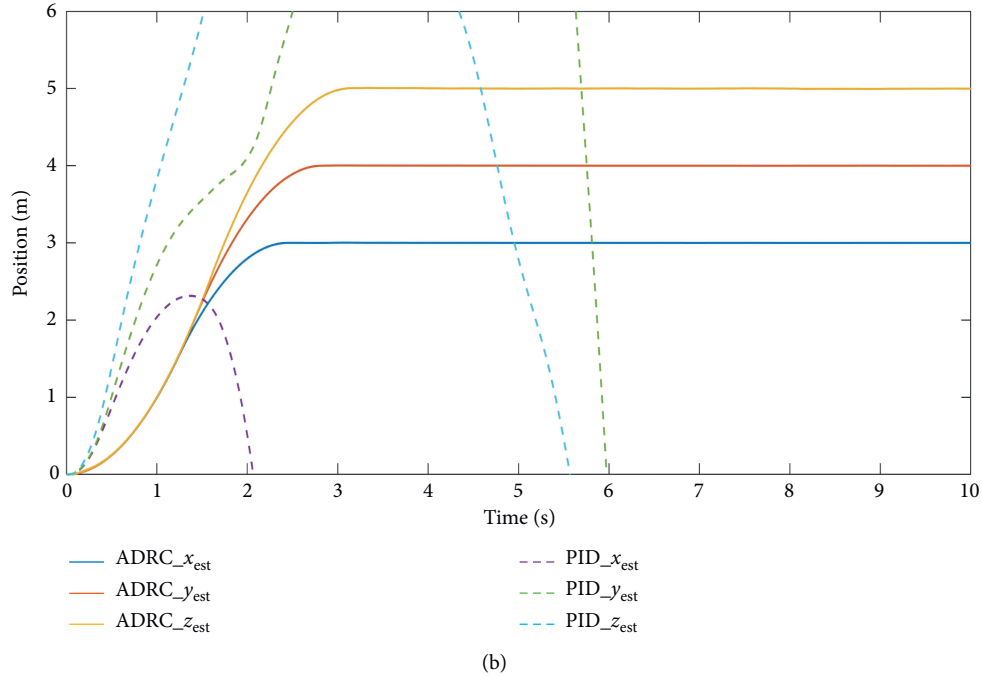


FIGURE 11: Attitude and position under turbulent wind and continuous wind.

PID controller parameters used for comparison are tuned based on particle swarm optimization.

- (i) No turbulent wind and no continuous wind
- (ii) Turbulent wind and no continuous wind
- (iii) Turbulent wind and continuous wind,  
 $V_w^E = [2 \ -2 \ 0]^T \text{ (m} \cdot \text{s}^{-1}\text{)}$

Set the initial position  $P_{\text{init}}^E = [0 \ 0 \ 0]^T \text{ (m)}$ , the initial attitude  $\eta_{\text{init}} = [0 \ 0 \ 0]^T \text{ (deg)}$ , the desired position  $P_{\text{des}}^E = [3 \ 4 \ 5]^T \text{ (m)}$ , and the desired attitude  $\eta_{\text{des}} = [10 \ 15 \ 30]^T \text{ (deg)}$ . The simulation results under the three conditions are as follows.

From Figure 9, the dynamic performance and steady-state performance of the PID controller are similar to those of the ADRC controller when there is no wind disturbance. In the presence of turbulent wind and no continuous wind (see Figure 10), the PID controller is unable to stably control the aircraft, and the altitude fluctuates continuously above and below the target value. When there is turbulent wind and continuous wind (see Figure 11), the PID controller cannot control the aircraft. Under the same wind disturbance, the ADRC controller can stably control the copter at the target height, and the control effect is almost the same as that without wind disturbance. From Figure 11 that when there is crosswind action, the lift of the rotor will increase more than when there is no wind. The PID controller cannot effectively suppress external disturbances such as wind disturbance. In contrast, the ESO of the ADRC controller can estimate the total disturbance, including wind disturbance in real time and control compensation.

All the above simulation experiments verify the effectiveness of the ADRC controller-based tilting six-rotor

aircraft flight control system. The results show that the aircraft has omnidirectional motion capability and can independently control the position and attitude. At the same time, under the regulation of ADRC, the aircraft can achieve stable control of the position and attitude. It has the characteristics of small overshoot, short adjustment time, strong robustness, and good external disturbance effect such as estimation and compensation of turbulent wind.

## 6. Conclusion

The focus of this paper is on a tilting hexacopter with an omnidirectional motion. The effects of large-scale uncertainty, parameter perturbation, strong coupling, and external environmental disturbances on the system are analyzed. Based on the aerodynamic influence and linearization control distribution matrix, a flight controller based on ADRC is designed and implemented. Independently controlled by the position and attitude with wind disturbance experiments compared with PID, the auto-disturbance suppression controller designed in this paper improves the antidisturbance capability of the hexacopter aircraft. In future work, we will design and manufacture tilting hexacopter based on open-source hardware, such as Pixhawk. At the same time, we will add more experiments, including control under failure conditions (rotor breakage, motor stalling, etc.).

## Data Availability

The figures, tables, and other data used to support this study are included within the article.



## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the Guizhou Provincial Science and Technology Projects (Grant no. [2020]2Y044), the Science and Technology Projects of China Southern Power Grid Co., Ltd. (Grant no. 066600KK52170074), the Key Laboratory Projects of Aeronautical Science Foundation of China (Grant no. 201928052006), and the National Natural Science Foundation of China (Grant no. 61473144).

## References

- [1] R. Mahony, V. Kumar, and P. Corke, "Multirotor aerial vehicles: modeling, estimation, and control of quadrotor," *IEEE Robotics & Automation Magazine*, vol. 19, no. 3, pp. 20–32, 2012.
- [2] Q. Quan, *Introduction to Multicopter Design and Control*, Springer, Berlin, Germany, 2017.
- [3] R. Rashad, F. Califano, and S. Stramigioli, "Port-hamiltonian passivity-based control on SE(3) of a fully actuated UAV for aerial physical interaction near-hovering," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4378–4385, 2019.
- [4] M. Ryll, G. Muscio, and F. Pierri, "6D physical interaction with a fully actuated aerial robot," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 5190–5195, Singapore, May 2017.
- [5] F. Ruggiero, V. Lippiello, and A. Ollero, "Aerial manipulation: a literature review," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1957–1964, 2018.
- [6] H. Bonyan Khamseh, F. Janabi-Sharifi, and A. Abdessameud, "Aerial manipulation—a literature survey," *Robotics and Autonomous Systems*, vol. 107, pp. 221–235, 2018.
- [7] B. Crowther, A. Lanzon, M. Maya-Gonzalez, and D. Langkamp, "Kinematic analysis and control design for a nonplanar multirotor vehicle," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 4, pp. 1157–1171, 2011.
- [8] E. Kaufman, K. Caldwell, D. Lee, and T. Lee, "Design and development of a free-floating hexrotor UAV for 6-DOF maneuvers," in *Proceedings of the IEEE Aerospace Conference*, pp. 1–10, Big Sky, MT, USA, March 2014.
- [9] Y. Tadokoro, T. Ibuki, and M. Sampei, "Maneuverability analysis of a fully-actuated hexrotor UAV considering tilt angles and arrangement of rotors," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8981–8986, 2017.
- [10] S. Rajappa, M. Ryll, H. H. Bulthoff, and A. Franchi, "Modeling, control and design optimization for a fully-actuated hexarotor aerial vehicle with tilted propellers," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 4006–4013, Seattle, WA, USA, May 2015.
- [11] D. Brescianini and R. D'Andrea, "Design, modeling and control of an omni-directional aerial vehicle," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3261–3266, Stockholm, Sweden, May 2016.
- [12] M. Kamel, S. Verling, O. Elkhatib et al., "The voliro omniorientational hexacopter: an agile and maneuverable tiltable-rotor aerial vehicle," *IEEE Robotics & Automation Magazine*, vol. 25, no. 4, pp. 34–44, 2018.
- [13] M. Ryll, H. H. Bulthoff, and P. R. Giordano, "Modeling and control of a quadrotor UAV with tilting propellers," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4606–4613, Saint Paul, MN, USA, May 2012.
- [14] M. Ryll, H. H. Bulthoff, and P. R. Giordano, "First flight tests for a quadrotor UAV with tilting propellers," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 295–302, Karlsruhe, Germany, May 2013.
- [15] G. Scholz and G. F. Trommer, "Model based control of a quadrotor with tiltable rotors," *Gyroscope and Navigation*, vol. 7, no. 1, pp. 72–81, 2016.
- [16] Y.-C. Choi and H.-S. Ahn, "Nonlinear control of quadrotor for point tracking: actual implementation and experimental tests," *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 3, pp. 1179–1192, 2015.
- [17] W. Wang, L. Yi-sha, and Y. Sheng-xuan, "An active disturbance-rejection flight control method for quad-rotor unmanned aerial vehicles," *Control Theory Applications*, vol. 32, no. 10, pp. 1351–1360, 2015.
- [18] G. V. Raffo, M. G. Ortega, and F. R. Rubio, "Nonlinear  $H_\infty$  controller for the quad-rotor helicopter with input coupling," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 13834–13839, 2011.
- [19] S. Bouabdallah and R. Siegwart, "Backstepping and sliding-mode techniques applied to an indoor micro quadrotor," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2247–2252, Barcelona, Spain, April 2005.
- [20] Sudhir and A. Swamp, "Second order sliding mode control for quadrotor," in *Proceedings of the 2016 IEEE 1st International Conference on Control, Measurement and Instrumentation, CMI*, pp. 92–96, Kolkata, India, January 2016.
- [21] J.-B. Liu, J. Cao, A. Alofi, A. Al-Mazrooei, and A. Elawi, "Applications of Laplacian spectra for  $n$ -prism networks," *Neurocomputing*, vol. 198, pp. 69–73, 2016.
- [22] J.-B. Liu, S. Wang, C. Wang, and S. Hayat, "Further results on computation of topological indices of certain networks," *IET Control Theory & Applications*, vol. 11, no. 13, pp. 2065–2071, 2017.
- [23] J.-B. Liu, C. Wang, S. Wang, and B. Wei, "Zagreb indices and multiplicative zagreb indices of eulerian graphs," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 42, no. 1, pp. 67–78, 2019.
- [24] J. Han, "From PID to active disturbance rejection control," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 3, pp. 900–906, 2009.
- [25] W. Xue and Y. Huang, "On performance analysis of ADRC for a class of MIMO lower-triangular nonlinear uncertain systems," *ISA Transactions*, vol. 53, no. 4, pp. 955–962, 2014.
- [26] Q. Zheng, L. Q. Gao, and Z. Gao, "On stability analysis of active disturbance rejection control for nonlinear time-varying plants with unknown dynamics," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 3501–3506, New Orleans, LA, USA, December 2007.
- [27] X. Yang and Y. Huang, "Capabilities of extended state observer for estimating uncertainties," in *Proceedings of the American Control Conference*, pp. 3700–3705, St. Louis, MO, USA, June 2009.
- [28] Z. X. Zhou, Y. Y. Qu, J. D. Yang, and D. J. Liu, "Suspension position control of precise planar magnetic bearing," *Kongzhi Yu Juece/Control and Decision*, vol. 25, no. 3, pp. 437–444, 2010.

- [29] Q. Zheng, L. Q. Gao, and Z. Gao, "On validation of extended state observer through analysis and experimentation," *Journal of Dynamic Systems, Measurement, and Control*, vol. 134, no. 2, pp. 1–6, 2012.
- [30] T. A. Johansen and T. I. Fossen, "Control allocation—a survey," *Automatica*, vol. 49, no. 5, pp. 1087–1103, 2013.
- [31] O. Härkegård and S. T. Glad, "Resolving actuator redundancy—optimal control vs. control allocation," *Automatica*, vol. 41, no. 1, pp. 137–144, 2005.
- [32] L. Wu, H. Bao, J. L. Du, and C. S. Wang, "A learning algorithm for parameters of automatic disturbances rejection controller," *Zidonghua Xuebao/Acta Automatica Sinica*, vol. 40, no. 3, pp. 556–560, 2014.
- [33] G. Kopasakis, "Modeling of atmospheric turbulence as disturbances for control design and evaluation of high speed propulsion systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 134, no. 2, pp. 1–12, 2012.
- [34] A. Jafar, S. Fasih-Ur-Rehman, S. Fazal-Ur-Rehman, N. Ahmed, and M. U. Shehzad, "A robust  $H_\infty$  control for unmanned aerial vehicle against atmospheric turbulence," in *Proceedings of the 2016 2nd International Conference on Robotics and Artificial Intelligence, ICRAI*, pp. 87–92, Rawalpindi, Pakistan, November 2016.

## Research Article

# The Cross-Entropy Method for the Winner Determination Problem in Combinatorial Auctions

Hanmi Lin <sup>1</sup>, Yongqiang Chen,<sup>2</sup> Changping Liu,<sup>3</sup> Jie Xie,<sup>4</sup> and Baixiu Ni <sup>4</sup>

<sup>1</sup>Qianjiang College, Hangzhou Normal University, Hangzhou 310016, China

<sup>2</sup>Alibaba Business College, Hangzhou Normal University, Hangzhou 310016, China

<sup>3</sup>Faculty of Management Engineering, Huaiyin Institute of Technology, Huaian 223200, China

<sup>4</sup>School of Finance and Mathematics, West Anhui University, Lu'an 237012, China

Correspondence should be addressed to Hanmi Lin; [ccnupolaris@163.com](mailto:ccnupolaris@163.com) and Baixiu Ni; [nibaixiu2000@163.com](mailto:nibaixiu2000@163.com)

Received 19 August 2020; Revised 16 September 2020; Accepted 17 September 2020; Published 12 October 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Hanmi Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The combinatorial auction is one of the important methods used for multi-item auctions, and the solution to the winner determination problem (WDP) is the key factor in the widespread application of combinatorial auctions. This paper explores the use of the cross-entropy method to solve the WDP, which is an NP problem. The performance of the proposed approach is evaluated on the basis of two well-known benchmark test cases. The experimental results show that, compared with the genetic algorithm and the particle swarm optimization algorithm, the cross-entropy (CE) method has the advantages of a higher success rate and a shorter time for solving the WDP. Therefore, the CE method provides a high-quality solution for the effective solution of the problem of determining winning bids in combined auctions.

## 1. Introduction

Combinatorial auctions, also called bundled auctions or packaged auctions, are a type of multi-item auctions that allow the bidders to bid on a combination of multiple items [1]. In multi-item auctions, bidders and the auctioneer can achieve a win-win situation through the combinatorial auction approach. On the one hand, bidders can express their preference, including complementarity and substitutability, as well as reducing the risks of bidding. On the other hand, the auctioneer can increase the economic revenue. Since the development of combinatorial auctions, they have been widely used in the real world, and the most well-known case is the spectrum licenses auction of America's Federal Communications Commissions [2]. In addition, the combinatorial auction approach has also been applied to a variety of resource allocation issues, such as transportation capacity resources [3], transportation line resources [4], airport time resources [5], logistics resources [6], industrial procurement resources [7], supply chain formation [8], collaborative rescue service resources [9], cloud computing

service resources [10], and network technology service resources [11].

The study of combinatorial auctions is a comprehensive application of interdisciplinary approaches, such as economics, game theory, operations research, computer science, and artificial intelligence. Experts and scholars in these fields have mainly focused on three core issues: combinatorial auctions-bidding language [12, 13], mechanism design [14–21], and the winner determination problem. In this paper, we will mainly discuss the issue of winner determination. According to the previous literature, approaches to the solution of the winner determination problem (WDP) generally use exact or inexact methods. Researchers [22–24] who use exact methods to solve the WDP have the advantage of being able to find the optimal solution and to prove its optimality. However, the shortcoming of the exact method is that it requires a large amount of calculations, and a huge amount of storage space is consumed in the calculation process. Therefore, in cases involving a large amount of data, the exact method may take a long time to find the optimal solution, which would introduce certain difficulties in real-world

applications. In contrast, using inexact methods to solve the WDP can find a local optimal solution within a certain period although they are unable to prove whether their solution has global optimality. In short, inexact methods are based on heuristics or metaheuristics, which help to find the optimal solution in the case of large amounts of data. In the past ten years, the inexact methods used in solving the WDP include the hybrid simulated annealing algorithm [25], the local search algorithm [26], the random local search algorithm [27], the memetic algorithm [28], the differential evolution algorithm [29], the biased random-key genetic algorithm [30], the discrete particle swarm algorithm [31], the ant colony algorithm and the graph-based ant colony algorithm [32, 33], and the deep-learning intelligent algorithm [34].

In this paper, the cross-entropy (CE) method is used to solve the WDP in combinatorial auctions. The advantage of CE is that it can not only estimate the probability of rare events but also can solve combinatorial optimization problems. This algorithm has a wide range of applications in computer science [35], mathematics [36], control science and engineering [37], mechanical engineering [38], and management science and engineering [39], among others.

The rest of this paper is organized into the following sections: Section 2 describes the WDP formulation. Section 3 introduces the CE algorithm. Section 4 gives the procedure of solving the WDP based on the CE algorithm. Section 5 reports experimental results, and the work is concluded in Section 6.

## 2. The Winner Determination Problem

The WDP refers to finding the winning bid that maximizes the revenue of the auctioneer under the constraint that each item is allocated to at most one bidder [29]. To what extent the solution of the WDP can guarantee quality and speed not only affects the final result of the auction but also affects the practicality of combinatorial auctions. The WDP is a complex combinatorial optimization problem, which has been proved to be an NP-complete problem, a weighted set packing problem [40].

The optimization of the WDP in combinatorial auctions can be stated as follows.

Suppose there is a set  $A$  comprised by  $a$  items, such that  $A$  can be stated as  $A = \{1, 2, \dots, a\}$ .  $B$  is a set of  $b$  bids, such that  $B$  can be stated as  $B = \{B_1, B_2, \dots, B_b\}$ . Each bid  $B_i$  is a tuple  $S_i, P_i$  where  $S_i$  is a subset of items and  $P_i$  is the price of bid  $S_i$  ( $i \geq 0$ ).  $M$  is an  $a \times b$  binary matrix where  $M_{ji} = 1$  if the object  $j \in S_i$ , otherwise  $M_{ji} = 0$ . Finally, we can obtain the decision variable as follows:  $x_i = 1$  if  $B_i$  is accepted, and  $x_i = 0$  otherwise.

Then, the optimal WDP in combinatorial auctions can be stated as the following integer programming model:

$$\max F(x) = \sum_{i=1}^b P_i \cdot x_i, \quad (1)$$

$$\text{s.t. } \sum_{i=1}^b M_{ji} \cdot x_i \leq 1; \quad j \in \{1, \dots, a\}; x_i \in \{0, 1\}. \quad (2)$$

## 3. The Cross-Entropy Method

In 1997, Israeli professor Reuven Y. Rubinstein proposed the cross-entropy (CE) method [41]. This method was first used to estimate rare-event probability and was later extended to solve optimization problems. The main idea is to transform the “deterministic” optimization problem into a related “random” optimization problem (accompanied by the stochastic optimization problem) and then to use the random simulation technology of rare-event simulation (importance sampling, IP) to solve this “random” optimization problem.

The framework of the cross-entropy method for solving an optimization problem can be described as follows.

Let us structure a general 0-1 integer maximization problem:

$$\max S(x): X \in R^n \longrightarrow R, \quad (3)$$

estimation problem. Let us denote probability distribution density functions  $\{f(\cdot; \nu), \nu \in V\}$  and indicator functions  $\{I\{S(X) \geq \gamma\}\}$  on  $X$ ; when  $u \in V$ , the corresponding auxiliary stochastic optimization problem can be stated as follows:

$$\begin{aligned} l(\gamma) &= P_u(S(X) \geq \gamma) = \sum_X I_{\{S(X) \geq \gamma\}} \cdot f(x; u) \\ &= E_u \left[ I_{\{S(X) \geq \gamma\}} \right], \end{aligned} \quad (4)$$

where  $E_u$  is the expectation operator,  $\gamma$  is the level parameter, and  $I = 1$  when  $S(X) \geq \gamma$ , otherwise  $I = 0$ .

$S(X) \geq \gamma$  is a rare event, and only when the sample size is large enough, can formula (4) be solved by the Monte Carlo method. Therefore, the CE algorithm usually uses importance sampling with density  $g(x)$  to reduce the number of samples; then formula (4) is transformed into

$$l(\gamma) = \frac{1}{N} \sum_{i=1}^N I_{\{S(x^i) \geq \gamma\}} \frac{f(x^i; u)}{g(x^i; V)}, \quad (5)$$

where  $x^i$  is a random sampling of  $g(x; \nu)$ .

In order to obtain the optimal density ( $g^*(x)$ ) of importance sampling, the CE introduces the Kullback–Leibler divergence to measure the distance between the two densities. Such that the optimal density ( $g^*(x)$ ) is obtained through minimizing the Kullback–Leibler divergence, and formula (6) expresses the maximization problem as follows:

$$\max_{\nu} \frac{1}{N} \sum_{i=1}^N I_{\{S(x^i) \geq \gamma\}} \ln f(x^i; \nu). \quad (6)$$

The main steps of the cross-entropy method can be summarized as follows:

Step 1: the parameters are set and initialized.

Step 2: a random sample is chosen according to the sampling distribution.

Step 3: the probability distribution parameters are modified on the basis of the so-called elite samples (the best scoring samples). The CE distance is involved in this step.

Step 4: The optimal probability distribution parameters are output, as long as the iteration termination conditions have been met, if not, return to Step 2.

#### 4. The CE Method for Solving the WDP

*4.1. Constraint Handling and Evaluation Function Construction.* In order to use CE to solve the WDP, we use the penalty method to transform the constrained optimization problem described in equations (1) and (2) into an unconstrained optimization problem and construct the function which evaluates the pros and cons of the individual samples of CE as follows:

$$\max F(x) = \frac{\sum_{i=1}^b P_i \cdot x_i}{1 + S}, \quad (7)$$

in which  $S$  is the penalty factor, whose definition is as follows:

$$S = \sum_{j=1}^a \left[ \sum_{i=1}^b M_{ji} \cdot x_i - 1 \right]^+, \quad (8)$$

in which  $[x]^+ = x$ , when  $x > 0$ ; otherwise,  $[x]^+ = 0$ . It can be seen that if the constraints are met, then

$$\sum_{j=1}^a \sum_{i=1}^b M_{ji} \cdot x_i \leq 1. \quad (9)$$

This means  $S = 0$ , and no penalty will be imposed on the evaluation function. If the constraint conditions are not met, then

$$\sum_{j=1}^a \sum_{i=1}^b M_{ji} \cdot x_i > 1, \quad (10)$$

and penalties are imposed on all the items which exceed 1.

*4.2. The CE Method for Solving the WDP.* The following main steps make up the process by which the CE method is used to solve the WDP problem:

Step 1: setting sample size  $N$ , the elite samples ratio  $\rho$ , adaptive parameters  $\alpha$ , the maximum number of iterations  $T$ , and parameter space dimension  $n$ .

Step 2: initializing probability distribution parameter  $p_0$  and number of iterations  $t=0$ .

Step 3: according to probability distribution, parameters  $p_t$  randomly choose sample vectors  $X_i = (x_1, x_2, \dots, x_n) \sim \text{Ber}(p_t)$ ,  $i = 1, 2, \dots, N$ , such that its probability distribution density function can be stated as follows:

$$f(X; p_t) = \prod_{i=1}^n p_t^{x_i} (1 - p_t)^{1-x_i}. \quad (11)$$

Step 4: evaluating all samples and sorting them to get the elite sample set and computing the means  $v_j$  of the elite samples  $X_e$  according to the following formula:

$$v_j = \frac{\sum_{s=1}^{\rho N} x_j^s}{\rho N}, \quad j = 1, \dots, n. \quad (12)$$

Step 5: using adaptive smoothing technology to modify probability distribution parameters  $p_t$ . Formula (13) is obtained as follows:

$$p_j^t = \alpha v_j + (1 - \alpha) p_j^{t-1}, \quad j = 1, \dots, n. \quad (13)$$

Step 6: setting  $t = t + 1$ , checking whether to stop iterating ( $t < T$ ), and outputting the optimal probability distribution parameter  $p^* = p_t$  when the criterion is met, otherwise return to Step 3.

#### 5. Experiments of the Data Sets

*5.1. Descriptions of the Test Cases.* In order to verify the performance of CE on the WDP, this paper selects two classic sets of data for testing [42, 43]. Test Case 1 is derived from a combined auction containing 10 bidders, 8 bids, and 30 bids. Its serial numbers, bids, and biddings are shown in Table 1.

Test Case 2 is derived from a combined auction containing 10 bidders, 10 bids, and 30 bids. Its serial numbers, bids, and biddings are shown in Table 2.

*5.2. Experimental Setup.* Experimental hardware environment used was Intel(R) Core(TM) i3 CPU M2.27 GHz, 2 GB RAM. Experimental software environment used was MATLAB 2018a. In order to evaluate the performance of CE on the WDP, we selected two classic algorithms: genetic algorithm (GA) [44] and particle swarm optimization algorithm (PSO) [45] for comparison.

The population size of the three algorithms is 1000, and the total number of valid bids was 30 in both test cases. The maximum number of iterations of Test Cases 1 and 2 is 200 and 500, respectively, and other related parameter settings are shown in Table 3, which determines the dimensionality  $n$  of the search space.

*5.3. Results and Analysis.* We used CE, GA, and PSO to solve Test Cases 1 and 2, respectively. Tables 4 and 5 give the statistical results of 30 independent experiments of the two test cases, respectively.

The numerical results in Table 4 and 5 show that CE can solve the WDP of combinatorial auctions. Compared with the genetic algorithm and the particle swarm optimization algorithm, CE can not only obtain the optimal solution consistent with them, but also obtain a larger average return than them, which verifies that CE is effective in solving the WDP. It can also be seen that CE is the fastest algorithm, and CE outperforms GA and PSO in terms of success rate. Above all, CE always outperforms GA and PSO in both solution quality and efficiency.

TABLE 1: Data of Test Case 1.

Serial number	Bids	Bidding	Serial number	Bids	Bidding
1	2, 4, 8	2401	16	8	903
2	2, 4, 6	2042	17	3	270
3	4	840	18	1, 4, 5	1658
4	4, 6, 8	2305	19	4, 5	1516
5	2, 4, 6	1996	20	1, 2	677
6	6	533	21	1, 7	569
7	4, 6, 8	2228	22	1	109
8	4, 6	1411	23	5, 6	1145
9	6, 8	1472	24	4, 8	1793
10	4	826	25	2, 6	1300
11	2, 4, 5	2135	26	6, 7	1032
12	5	629	27	2, 4	1479
13	4, 7, 8	2095	28	1, 4	937
14	7	421	27	3, 4	1075
15	3, 5, 8	1881	30	2	608

TABLE 2: Data of Test Case 2.

Serial number	Bids	Bidding	Serial number	Bids	Bidding
1	(B, D, H)	2407.19	16	(F, J)	1410.36
2	( F, H, I)	2309.00	17	(C, I)	1135.49
3	(F, H, J)	2224.43	18	(F, G)	1022.73
4	( B, D, E)	2145.06	19	(A, D)	943.72
5	(G, I, J)	2067.01	20	(H)	906.36
6	(B, D, F)	2040.48	21	(I)	840.568
7	(B, F, I)	1998.80	22	(D)	827.86
8	(C, E, H)	1881.22	23	(J)	771.94
9	(H, I)	1798.28	24	(E)	626.50
10	(A, D, E)	1652.16	25	(B)	607.23
11	(C, F, J)	1646.92	26	(A, G)	566.87
12	(B, F, G)	1620.68	27	(F)	531.89
13	(E, I)	1518.77	28	(G)	415.26
14	(F, H)	1478.81	29	(C)	273.97
15	(B, J)	1469.68	30	(A)	101.12

TABLE 3: Parameter setting of GA, PSO, and CE for solving the winner determination problem (WDP).

Algorithm	Parameter description	Symbol	Value
GA	Crossover percentage	$p_c$	0.8
	Mutation percentage	$p_m$	0.3
	Mutation rate	$mu$	0.02
PSO	Max inertia weight	$w_{max}$	0.9
	Min inertia weight	$w_{min}$	0.4
	Acceleration coefficient	$c_1, c_2$	2
CE	Rarity parameter	$\rho$	0.01
	Smoothing parameter	$\alpha$	0.7
	Number of elites	$N_e$	20

TABLE 4: Comparison of results from 30 experiments for Test Case 1.

The optimal solution	Algorithm	Average of time (s)	The best result	The worst result	Average of results	Success rate (%)
4590	GA	6.19	<b>4590</b>	<b>4558</b>	4578.87	53.33
	PSO	1.30	<b>4590</b>	4472	4575.07	66.67
	CE	<b>1.26</b>	<b>4590</b>	4501	<b>4584.30</b>	<b>86.67</b>

Figures 1 and 2 show the average change of the objective function during the 30 experimental iterations of the three algorithms to solve Test Cases 1 and 2, that is, the convergence process of the three algorithms. It is found from the

two figures that CE has the fastest convergence speed, followed by PSO, and GA is the slowest. Because the time running cost of CE is a linear function of the dimensionality of the search space, it is advantageous to use the CE



TABLE 5: Comparison of results from 30 experiments for Test Case 2.

The optimal solution	Algorithm	Average of time (s)	The best result	The worst result	Average of results	Success rate (%)
6216.82	GA	15.66	6216.82	<b>6141.24</b>	6181.89	23.33
	PSO	3.50	6216.82	6120.07	6186.71	20.00
	CE	<b>3.31</b>	<b>6216.82</b>	6134.10	<b>6202.54</b>	<b>53.33</b>

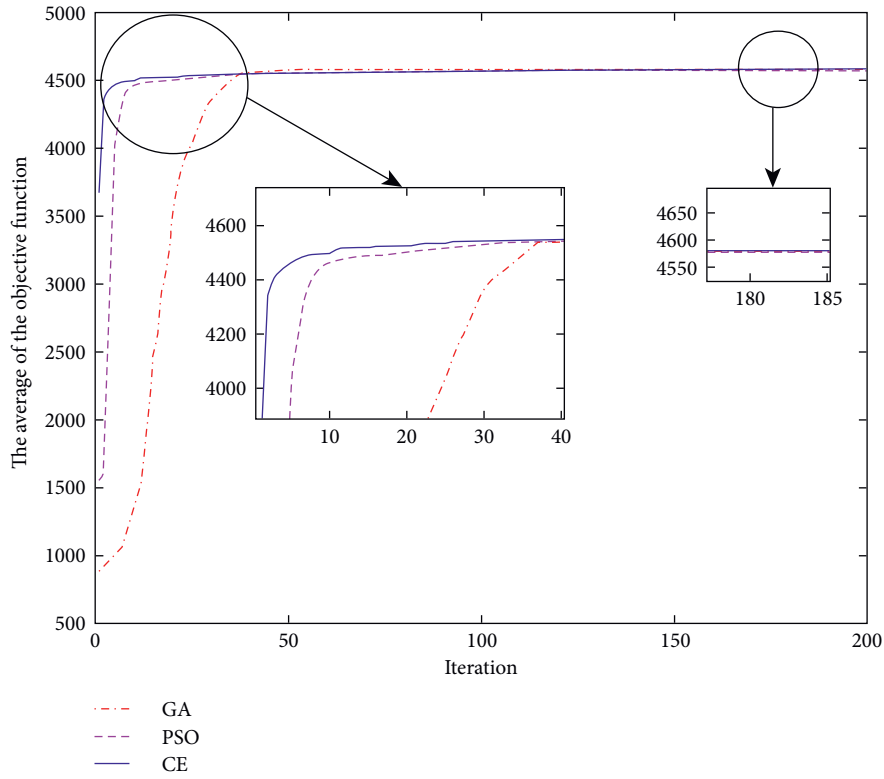


FIGURE 1: Comparison of the convergence characteristics of the three algorithms on Test Case 1.

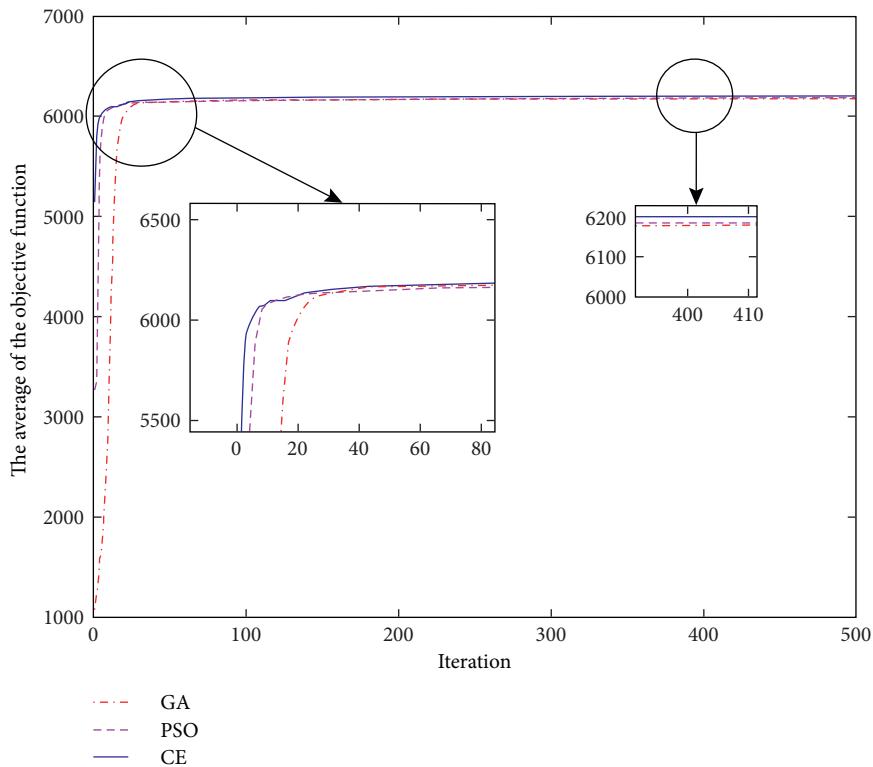


FIGURE 2: Comparison of the convergence characteristics of the three algorithms on Test Case 2.

algorithm to solve large-scale (large total number of quotes) WDP problems.

## 6. Conclusion

This paper attempts to use CE to solve the WDP of large-scale combinatorial auctions and uses the elite strategy and the importance sampling method to reduce the number of samples and speed up convergence. The proposed method was evaluated on two test data samples and compared with GA and PSO. The numerical results show that the CE method always outperforms the GA and PSO methods in both solution quality and efficiency. It is therefore advantageous to use CE to solve large-scale (large total number of quotes) WDP problems. All in all, the CE method provides a high-quality and effective solution of the WDP.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Hanmi Lin and Baixiu Ni designed the study. Baixiu Ni and Jie Xie implemented all experiments. Hanmi Lin wrote the first draft. Changping Liu performed the design of figures. Yongqiang Chen revised the manuscript. All authors contributed to the manuscript revision and read and approved the submitted version.

## Acknowledgments

This work was funded by the National Social Science Foundation of China (Grant no. 17BSH040) and the Natural Science Foundation of West Anhui University (Grant no. WXZR201818).

## References

- [1] P. Cramton, Y. Shoham, and R. Steinberg, *Combinatorial Auctions*, MIT Press, Cambridge, MA, USA, 2006.
- [2] P. Milgrom, "Putting auction theory to work: the simultaneous ascending auction," *Journal of Political Economy*, vol. 108, no. 2, pp. 245–272, 2000.
- [3] C. Triki, S. Oprea, P. Beraldi, and T. G. Crainic, "The stochastic bid generation problem in combinatorial transportation auctions," *European Journal of Operational Research*, vol. 236, no. 3, pp. 991–999, 2014.
- [4] A. Y. S. Lam, "Combinatorial auction-based pricing for multi-tenant autonomous vehicle public transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 859–869, 2016.
- [5] J. W. Song and S. B. Yang, "A faster optimal allocation algorithm in combinatorial auctions," in *Proceedings of the 2004 International Conference on Micai: Advances in Artificial Intelligence (MICAI 2004)*, Springer, Berlin, Heidelberg, pp. 362–369, 2004.
- [6] O. Gujo and M. Schwind, "Bid formation in a combinatorial auction for logistics services," in *Proceedings of the 2007 International Conference on Enterprise Information Systems (ICEIS 2007)*, Springer, Berlin Heidelberg, pp. 303–3155, 2007.
- [7] M. Bichler, A. Davenport, G. Hohner, and J. Kalagnanam, *Industrial Procurement Auctions*, MIT Press, Cambridge, MA, USA, 2005.
- [8] W. E. Walsh, M. P. Wellman, and F. Ygge, "Combinatorial auctions for supply chain formation," in *Proceedings of the 2nd ACM Conference on Electronic Commerce*, pp. 260–269, Minneapolis, MN, USA, 2000.
- [9] V. Prasad, S. Rao, and A. S. Prasad, "A combinatorial auction mechanism for multiple resource procurement in cloud computing," in *Proceedings of the 12th International Conference on Intelligent Systems Design and Applications (ISDA 2012)*, IEEE, Kochi, India, 2012.
- [10] L. Fawcett, M. Broadbent, and N. Race, "Combinatorial auction-based resource allocation in the fog," in *Proceedings of the Fifth European Workshop on Software-Defined Networks (EWSND 2016)*, IEEE, Hague, Netherlands, 2017.
- [11] K. Zhu, J. Tang, H. Guo et al., "Using a combinatorial auction-based approach for simulation of rescue operations in disaster relief," *International Journal of Modeling Simulation & Scientific Computing*, vol. 9, no. 4, pp. 1–21, 2018.
- [12] N. Nisan, "Bidding languages for combinatorial auctions," "Bidding languages for combinatorial auctions," in *Combinatorial Auctions*, P. Cramton et al., Ed., MIT Press, Cambridge, MA, USA, 2006.
- [13] B. Martin, "Compact bidding languages and supplier selection for markets with economies of scale and scope," *European Journal of Operational Research*, vol. 214, no. 1, pp. 67–77, 2011.
- [14] S. Assadi, "Combinatorial auctions do need modest interaction," in *Proceedings of the 2017 ACM Conference on Economics and Computation ACM Transactions on Economics and Computation (TEAC2017)*, Association for Computing Machinery, Cambridge, MA, USA, 2017.
- [15] B. Vangerven, D. R. Goossens, and F. C. R. Spieksma, "Using feedback to mitigate coordination and threshold problems in iterative combinatorial auctions," *Business & Information Systems Engineering*, 2020.
- [16] V. V. Narayan, G. Rayaprolu, and A. Vetta, *Risk-Free Bidding In Complement-Free Combinatorial Auctions*, pp. 123–136, Springer, Cham, Switzerland, 2019.
- [17] N. Muto and S. Yasuhiro, "Goods revenue monotonicity in combinatorial auctions," Graduate School of Economics, Hitotsubashi University, Kunitachi, Japan, Discussion Papers, 2013.
- [18] S. Dobzinski, N. Nisan, and M. Schapira, "Truthful randomized mechanisms for combinatorial auctions," *Journal of Computer and System Sciences*, vol. 78, no. 1, pp. 15–25, 2012.
- [19] P. Le, "Pareto optimal budgeted combinatorial auctions," *Theoretical Economics*, vol. 13, pp. 1–31, 2018.
- [20] L. M. Ausubel, P. Cramton, and P. Milgrom, *The Clock-Proxy Auction: A Practical Combinatorial Auction Design*, MIT Press, Cambridge, MA, USA, 2005.
- [21] A. Land, S. Powell, and R. Steinberg, *PAUSE: A Computationally Tractable Combinatorial Auction*, MIT Press, Cambridge, MA, USA, 2005.
- [22] R. Nabila, A. Amine, E. H. Issmail et al., "A robust optimization approach for the winner determination problem with uncertainty on shipment volumes and carriers' capacity,"

- Transportation Research Part B: Methodological*, vol. 123, pp. 127–148, 2019.
- [23] B. Vangerven, D. R. Goossens, and F. C. R. Spijksma, “Winner determination in geometrical combinatorial auctions,” *European Journal of Operational Research*, vol. 258, no. 1, pp. 254–263, 2017.
- [24] G. Lin, W. Zhu, and M. M. Ali, “An effective discrete dynamic convexized method for solving the winner determination problem,” *Journal of Combinatorial Optimization*, vol. 32, no. 2, pp. 563–593, 2016.
- [25] Y. Guo, A. Lim, B. Rodrigues, and Y. Zhu, “Heuristics for a bidding problem,” *Computers & Operations Research*, vol. 33, no. 8, pp. 2179–2188, 2006.
- [26] H. Zhang, S. Cai, C. Luo et al., “An efficient local search algorithm for the winner determination problem,” *Journal of Heurs*, vol. 23, no. 2, pp. 1–30, 2017.
- [27] D. Boughaci, B. Benhamou, and H. Drias, “Stochastic local search for the optimal winner determination problem in combinatorial auctions,” in *Proceedings of the International Conference on Principles and Practice of Constraint Programming*, Springer, Berlin, Heidelberg, pp. 593–597, 2008.
- [28] D. Boughaci, B. Benhamou, and H. Drias, “A memetic algorithm for the optimal winner determination problem,” *Soft Computing*, vol. 13, no. 8-9, pp. 905–917, 2009.
- [29] D. Boughaci, “A differential evolution algorithm for the winner determination problem in combinatorial auctions,” *Electronic Notes in Discrete Mathematics*, vol. 36, pp. 535–542, 2010.
- [30] C. E. De Andrade, R. F. Toso, M. G. C. Resende, and F. K. Miyazawa, “Biased random-key genetic algorithms for the winner determination problem in combinatorial auctions,” *Evolutionary Computation*, vol. 23, no. 2, pp. 279–307, 2015.
- [31] F. S. Hsieh, “A discrete particle swarm algorithm for combinatorial auctions,” in *Proceedings of the 2017 International Conference on Swarm Intelligence (ICSI 2017)*, Springer, Cham, Switzerland, pp. 201–208, 2017.
- [32] A. Ray and M. Ventresca, “An ant colony approach for the winner determination problem S,” in *Proceedings of the 2018 European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCOP 2018)*, Springer, Cham, Switzerland, pp. 174–188, 2018.
- [33] A. Ray, M. Ventresca, and K. N. Kannan, “A graph based ant algorithm for the winner determination problem in combinatorial auctions,” *SSRN Electronic Journal*, 2018.
- [34] G. Brero, B. Lubin, and S. Seuken, “Deep learning-powered iterative combinatorial auctions,” <https://arxiv.org/pdf/1911.08042.pdf>.
- [35] F. Mehdi, M. Asghar, H. Michael et al., “A cross-entropy method for optimising robotic automated storage and retrieval systems,” *International Journal of Production Research*, vol. 56, no. 19, pp. 6450–6472, 2018.
- [36] Y. Shi, M. Wang, W. Shi et al., “Accurate and efficient estimation of small P-values with the cross-entropy method: applications in genomic data analysis,” *Bioinformatics*, vol. 35, no. 14, 2018.
- [37] A. Kazemdehdashti, M. Mohammadi, and A. R. Seifi, “The generalized cross-entropy method in probabilistic optimal power flow,” *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5738–5748, 2018.
- [38] K. Abolfazl, M. Mohammad, and S. A. Reza, “Application of generalized cross-entropy method in probabilistic power flow,” *IET Generation Transmission & Distribution*, vol. 12, no. 11, pp. 2745–2754, 2018.
- [39] K. Sebaa, A. Tlemcani, and H. Nouri, “FACTS location and tuning using the cross-entropy method,” in *Proceedings of the Power Engineering Conference (UPEC), 2014 49th International Universities*, IEEE, Cluj-Napoca, Romania, 2014.
- [40] M. H. Rothkopf, A. Pekeč, and R. M. Harstad, “Computationally manageable combinatorial auctions,” *Management Science*, vol. 44, no. 8, pp. 1131–1147, 1998.
- [41] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte Carlo Simulation and Machine Learning*, Springer-Verlag, New York, NY, USA, 2004.
- [42] P. Chen and D. Wang, “Genetic algorithm for solving winner determination in combinatorial auctions,” *Journal of Northeastern University (Natural Science)*, vol. 24, pp. 7–10, 2003.
- [43] P. Chen and D. Wang, “Chaotic search algorithm for winner determination in combinatorial auctions,” *Journal of Management Science in China*, vol. 5, pp. 24–28, 2003.
- [44] R. L. Haupt and S. E. Haupt, *The Binary Genetic Algorithm Practical Genetic Algorithms*, John Wiley & Sons, Hoboken, NJ, USA, Second Edition, 2004.
- [45] S. Mirjalili and A. Lewis, “S-shaped versus V-shaped transfer functions for binary Particle Swarm Optimization,” *Swarm and Evolutionary Computation*, vol. 9, pp. 1–14, 2013.

## Research Article

# Efficient Processing of Image Processing Applications on CPU/GPU

**Najia Naz,<sup>1</sup> Abdul Haseeb Malik,<sup>1</sup> Abu Bakar Khurshid,<sup>1</sup> Furqan Aziz,<sup>2</sup> Bader Alouffi,<sup>3</sup> M. Irfan Uddin ,<sup>4</sup> and Ahmed AlGhamdi<sup>5</sup>**

<sup>1</sup>Department of Computer Science, University of Peshawar, Peshawar, Pakistan

<sup>2</sup>Department of Computer Science, Institute of Management Sciences, Peshawar, Pakistan

<sup>3</sup>Department of Computer Science, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

<sup>4</sup>Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan

<sup>5</sup>Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

Correspondence should be addressed to M. Irfan Uddin; [irfanuddin@kust.edu.pk](mailto:irfanuddin@kust.edu.pk)

Received 16 July 2020; Accepted 27 September 2020; Published 10 October 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Najia Naz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Heterogeneous systems have gained popularity due to the rapid growth in data and the need for processing this big data to extract useful information. In recent years, many healthcare applications have been developed which use machine learning algorithms to perform tasks such as image classification, object detection, image segmentation, and instance segmentation. The increasing amount of big visual data requires images to be processed efficiently. It is common that we use heterogeneous systems for such type of applications, as processing a huge number of images on a single PC may take months of computation. In heterogeneous systems, data are distributed on different nodes in the system. However, heterogeneous systems do not distribute images based on the computing capabilities of different types of processors in the node; therefore, a slow processor may take much longer to process an image compared to a faster processor. This imbalanced workload distribution observed in heterogeneous systems for image processing applications is the main cause of inefficient execution. In this paper, an efficient workload distribution mechanism for image processing applications is introduced. The proposed approach consists of two phases. In the first phase, image data are divided into an ideal split size and distributed amongst nodes, and in the second phase, image data are further distributed between CPU and GPU according to their computation speeds. Java bindings for OpenCL are used to configure both the CPU and GPU to execute the program. The results have demonstrated that the proposed workload distribution policy efficiently distributes the images in a heterogeneous system for image processing applications and achieves 50% improvements compared to the current state-of-the-art programming frameworks.

## 1. Introduction

GPUs (graphical processing units) are becoming popular to exploit data-level parallelism [1] in embarrassingly parallel applications [2] because of the SIMD (single instruction multiple data) [3, 4] architecture. However, task-level parallelism [5–7] is better exploited in general-purpose processors, because of MIMD (multiple instruction multiple data) [8] architecture. More recently, general-purpose processors are combined with GPUs and provide general-

purpose computation accelerated with GPU (commonly referred as GPGPU) [9]. A heterogeneous cluster has many CPUs and GPUs and can exploit both task-level and data-level parallelism in applications.

The amount of data generated in the form of images and videos is enormous because of the fact that many surveillance cameras, smart phones, and many other devices that capture images/videos are installed/used everywhere. These devices are constantly recording scenes and can be processed for different types of computer vision, image processing,



machine learning, and data science tasks. Images inherently have data-level parallelism [1, 10] (i.e., individual images can be processed in embarrassingly parallel fashion), because all pixels can be processed independently and therefore GPUs are commonly used for image processing applications. A heterogeneous cluster (i.e., a cluster containing CPUs and GPUs) can exploit both task-level (across multiple images) and data-level parallelism in images. GPUs enable heterogeneous clusters to accelerate the processing intensive operations of images in big data and machine learning applications.

Data processing applications are commonly processed in a cloud environment using the MapReduce [11] parallel processing model for efficient execution. The scheduler of the MapReduce influences the performance in different applications when utilized in a heterogeneous cluster. The dynamic nature of the cluster and the computing workload affect the execution time of the application. Data locality is an essential part to reduce the total application execution time and hence improve the overall throughput.

In current technology, it is very challenging to provide a mechanism that can efficiently utilize the computing resources in a heterogeneous cluster [12]. Traditionally, applications that use both CPUs and GPUs are created using low level programming languages. Limited support is available to programmers to efficiently exploit parallelism in these clusters. A heterogeneous cluster normally has nodes of different computational capabilities [13]. For instance, a node may have 4 CPUs and 1 GPU, and another node may have 16 CPUs and 8 GPUs. Therefore, static distribution of workload amongst these nodes without taking into account their processing capabilities is not justified [14] as one node will be overloaded, and the other node will remain idle for most of the time [15]. Load balancing can be done dynamically, but it requires extra overheads at runtime [16].

In this paper, we present a new technique for the efficient distribution of images in a heterogeneous cluster. The goal is to maximize the utilization of the processing resources [17–19] (i.e., both CPUs and GPUs) and throughput. We provide a programming framework that ensures efficient workload distribution amongst the nodes by dividing the data into equal size splits and then distribute the split data between CPU and GPU cores based on their computational capabilities [20]. The aim is to achieve maximum resource utilization, gain high performance by dividing the data into equal size splits, allocate the data locally to the computing units, and minimize data migration to GPUs [21].

The rest of the paper is organized as follows. Related studies are given in Section 2. In Section 3, we provide a background to existing work that use CPU and GPU integration to efficiently solve different problems. We also provide a brief overview of the Hadoop programming framework, as this framework is used to test the ideas presented in this paper. We highlight limitations in existing state-of-the-art frameworks and explain the problems due to which efficient utilization of computing resources is not possible. The details of the proposed framework are given in Section 4, and demonstration of experiments is given in Section 5. We conclude the paper in Section 6.

## 2. Related Work

In [22], image processing for face detection and tracking is performed using CPU and GPU integration in Hadoop framework and has improved the performance by 25%. In another research [23], GPU-based Hadoop framework is used for evaluation of Canny Edge Detection algorithm using the default scheduler of Hadoop for workload distribution and has demonstrated two times performance improvement than HIPI-based image processing [24]. In [25, 26], face detection in video frames is performed by using CUDA-based Hadoop framework. It is observed that actual data processing takes 55% of the processor time, and the remaining 45% is wasted as idle or busy in performing other management activities.

Another framework named SEIP (system for efficient image processing) [27] ensured high performance for image processing application by applying in-node pipeline framework. However, during the processing, it is observed that the number of load/store to the GPU is equal to the number of images, which results into overhead and performance loss in case of processing a large number of small images, and also there is no policy of workload distribution between CPU and GPU. In [28], an integrated framework based on Hadoop and GPU has been developed for processing massive amount of satellite images. In order to achieve application performance, image data are split into parts and then each part is allocated to processing units, but there is no support for efficient workload distribution between CPU and GPU.

An energy efficient runtime mapping and thread partitioning approach has been developed for distribution of concurrent OpenCL application between CPU and GPU cores and has demonstrated a 32% increase in the system performance [29]. The feature extraction algorithm SIFT [30] has been developed in OpenCL that distributes workload on CPU and GPU. It is demonstrated that features were extracted with more than 30 FPS (frames per second) on full HD images, and an average speed up of 2.69 was achieved. Another OpenCL-based framework single node vertically scaled system is developed in [31] where multiple GPUs are combined together and treated as a single computing device. It automatically distributes OpenCL kernel written for a single GPU into multiple CUDA kernels at runtime that is executed on multiple (eight) GPUs. The experiment was performed by combining 8 GPUs in a single node environment, and the performance speedup of 7.1x was achieved as compared to the performance of single GPU. An average overhead of 0.48% is reported.

In [32], an algorithm is proposed that improves the efficiency of Hadoop clusters. The experiments demonstrated that if a process is defined that can handle different use-case scenarios, the overall cost of computing can be reduced and get benefits from distributed system for fast executions. A reinforcement learning-based MapReduce scheduler is proposed in [33] for heterogeneous environment. The system observes the state of task execution and suggests speculative execution of slow tasks to other free nodes in the cluster for faster execution. The proposed approach does not need any prior knowledge of the

environment and adapts itself to the heterogeneous environment. The experiments demonstrate that over a few runs, the system can better map the tasks to the available resources in a heterogeneous cluster and hence improve the overall performance of the system. The workload partition and task granularity for a given application based on machine learning techniques are given in [34]. The machine learning model can train a predictive model off-line, and then the trained model can predict the data partition and task granularity for any program at runtime. The experiments demonstrate a 1.6× average speedup using a single MIC. Other studies that involve the improvement in parallel computation are given in [35–38].

In all techniques presented in this section, the objective is to improve the performance of execution in a heterogeneous environment. The current state-of-the-art techniques demonstrate that imbalanced distribution of tasks without considering the underlying computational capabilities results into inefficient execution of the applications. The proposed technique in this paper considers the underlying computational capability of the processor and then assigns tasks. This results into performance improvement as demonstrated in Section 5. To the best of our knowledge, no previous studies have addressed the problem in the same perspective as undertaken by this research.

### 3. Background

In this section, existing frameworks are explained along with the limitations.

**3.1. Programming Frameworks.** Different programming frameworks such as Hadoop [39], FastFlow [40], OpenMP [41], pthreads [42], OpenCL [43], DirectCompute [44], OpenGL [45], MapReduce [46], and Spark [47] have been developed for the efficient data processing in a heterogeneous environment. Hadoop is becoming very popular because it can efficiently process structured [48], semi-structured [49], and unstructured [50] data. Different image processing applications such as face and motion detection [51], face tracking [52], extracting text from video frames in an online lecture video [53], video processing for surveillance [54], and content-based image retrieval (CBIR) [55] have demonstrated that Hadoop can be efficiently used for image-based applications.

CUDA (compute unified device architecture) [56] is the most commonly used programming language for GPUs developed by Nvidia. CUDA integrated with Hadoop enhances the application throughput by using the distributed computing capability of Hadoop and parallel processing capability of GPU [57]. Mars framework [58], which has been used for processing of web documents (searches and logs), was the first framework which combined GPU with Hadoop. Some other popular frameworks that integrate GPUs with Hadoop are MAPCG [59], StreamMR [60], and GPMR [61]. However, these frameworks are developed for some specific projects and do not improve the performance

of image processing applications in Hadoop. Hadoop Image processing interface (HIPI) has been developed that can efficiently process a massive amount of small sized images but has no support of GPU [24].

A heterogeneous Hadoop cluster with nodes equipped with GPUs is shown in Figure 1. Hadoop is one of the famous and easy to use platforms which is a loosely coupled architecture and provides a distributed environment. Hadoop consists of the Hadoop distributed file system (HDFS) [62] and MapReduce [46] programming model. HDFS is an open-source implementation of the Google File System (GFS). It stores the data on different data nodes in a cluster. HDFS depends on the mechanism of the master-slave architecture. The access permission and data service to the slave nodes are provided by *master* node also known as *name* node, while the slaves, known as *data* nodes, are used as storage for the HDFS. Large files are handled efficiently by dividing them into chunks and then distributed amongst multiple data nodes. On each node, to process their local copies, a *map* processing job is located. The function of *name* node is only to keep record of the metadata and log information, while Hadoop API is used for the transfer of data to and from HDFS. MapReduce is an enhanced approach which provides an abstraction for data synchronization, load balancing, and dynamically allocation of tasks to different computing units in a reliable manner.

**3.2. Limitations in Existing Systems.** Some issues that lead to imbalanced workload distribution in a heterogeneous environment are discussed below.

**3.2.1. Data Locality.** Data locality means that the mapper and data are located on the same node. If data and mapper are on the same node, then it is easy for a mapper to efficiently map the data for computation, but if the data are on a different node than the mapper, then the mapper have to load data from different node over the network to be distributed. Suppose there are 50 mappers which try to copy data from other data nodes simultaneously. This situation leads to high network congestion which is not desirable because the overall performance of the application is affected. The situation where mapper and data are on the same node is shown in Figure 2(a), and the situation where mapper and data are on different nodes is shown in Figure 2(b). For efficient processing of applications, a programming framework should be able to ensure data locality. When the data stored in HDFS (Hadoop distributed file system) is distributed amongst the nodes, data locality needs to be handled very carefully. The data are divided into splits, and each split is provided to the data node in the cluster for processing. The MapReduce job is executed to map splits to individual mapper that will process the assigned split. That means that moving the computation closer to the data is better than moving the data closer to the computation. Hence, good data locality means good application performance.



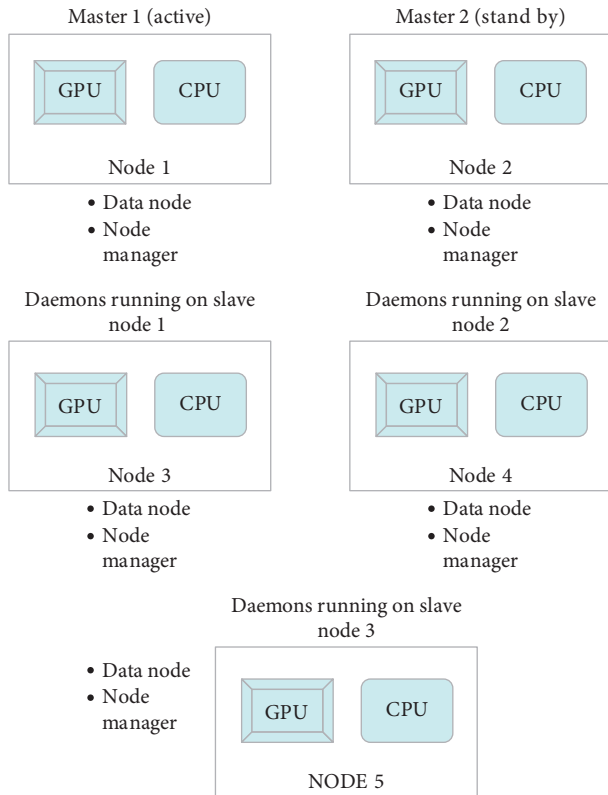


FIGURE 1: Heterogeneous Hadoop cluster with nodes equipped with GPUs.

**3.2.2. Split Size.** For efficient execution of programs in heterogeneous cluster, the programming framework should be able to distribute the data evenly into splits and then distribute amongst the available nodes. One of the main characteristics of MapReduce is to divide the whole data into chunks/input splits according to the block size of HDFS. As by default, Hadoop block size is 64 MB, and the issue of data locality arises when the input split size is larger than the block size. For better performance, input split size should be equal to or less than the block size of HDFS.

*Block size* and *split size* are not the same terms, as the block size is the physical chunk of data stored in disk, whereas input split size is the logical chunk of data with pointers for start and end locations in a block. When the split size is more or very small than the default block size, then uneven distribution of data happens, which leads to issues in data locality and memory wastage.

In Figure 3, the scenario of uneven split size is highlighted. Suppose the block size is 64 MB and each split size is 50 MB. The first split will easily fit in block 1, but the second split starts after the first split ending point and will not fully fit in the block 1, so the remaining part of the second split will be partially stored in block 1 and partially stored in block 2. When the mapper is assigned to block 1, it reads the first split, it will not read the second split data as it is not fully fitted in block 1 and cannot generate any final result of the second split data. According to [39, 63, 64], as quoted from the book “The logical records that *FileInputFormats* define do not usually fit neatly into HDFS blocks. For example, a

*TextInputFormat*’s logical records are lines, which will cross HDFS boundaries more often than not. This has no bearing on the functioning of your program—lines are not missed or broken, for example—but it’s worth knowing about, as it does mean that data-local maps (that is, maps that are running on the same host as their input data) will perform some remote reads. The slight overhead this causes is not normally significant.” If a record/file span across the HDFS boundaries of two nodes, then one of the nodes will perform some remote reads to fetch the missing piece. And it will read the data and generate final results but with the overhead of communication between the two nodes. Hence, the communication overhead arises because the data are not evenly distributed according to the block size, and most of the time, mapper waits for other mappers to generate the result and then to synchronize with each other for final result. This problem can be solved by arranging the whole data into ideal input split size. By dividing the data into equal and suitable split size that is less than or equal to the default block size, the mapper of each block will read its data easily and will not wait for other mapper to send data of the split that is partially stored in different blocks.

**3.2.3. Data Migration and Inefficient Resource Utilization.** Data migration is the process of transferring data from one node or processor to another node or processor as shown in Figure 4. Data migration between systems is usually performed programmatically to achieve better performance, but in heterogeneous systems, where a node contains CPU and GPU, the GPU being the faster computing processor will complete its task quickly and will fetch the data from CPU, a slow computing processor. Due to this data migration, the scheduler will always be busy in managing the tasks scheduling, the GPU will be idle, and hence performance in applications is affected.

Above are some of the problems that need to be tackled while using heterogeneous systems, so that application performance can be increased. In this paper, we will integrate CUDA with Hadoop to increase the performance in processing images in heterogeneous clusters. We will integrate the Hadoop platform that is used for distributed processing on clusters with libraries that allow code to be executed on GPUs.

## 4. Efficient Workload Distribution Based on Processor Speed

The issues of data locality, input split size, data migration, and inefficient resource utilization discussed in Section 3.2 lead to imbalanced workload distribution in heterogeneous environment, which results into inefficient execution of applications. The proposed framework will distribute the data in a balanced form amongst the nodes according to their computing capabilities, as shown in Figure 5. The distribution of data in the framework consists of two phases. In Phase I, the data are distributed amongst the nodes and in Phase II the workload is equally distributed between CPU

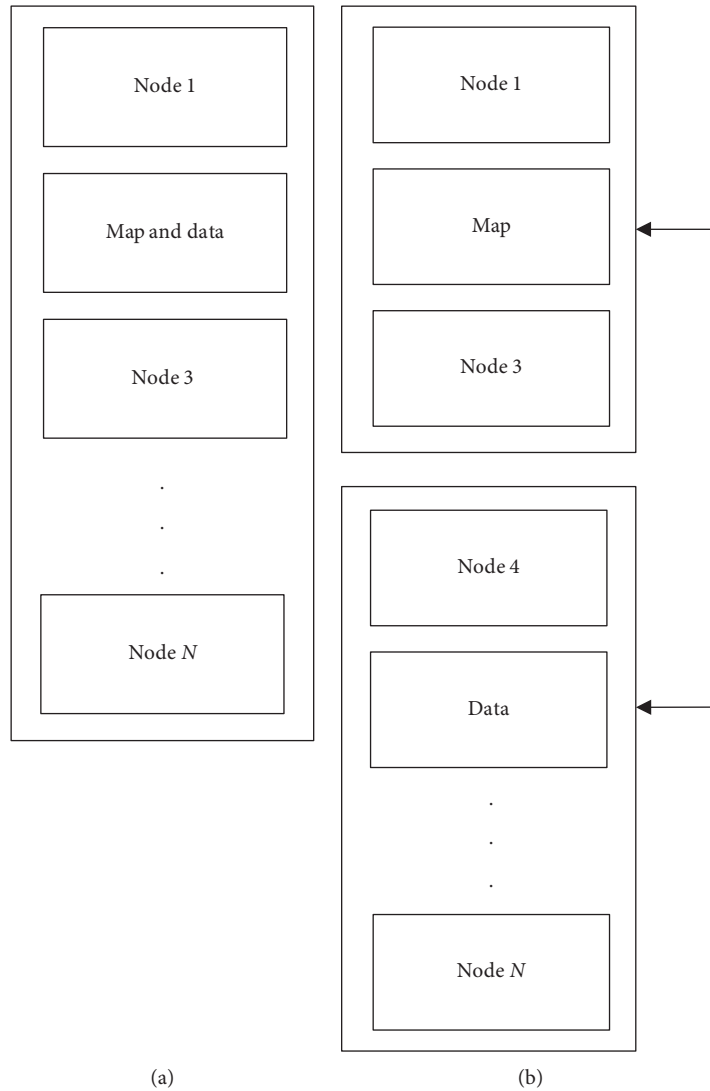


FIGURE 2: (a) Data and Mapper on the same node; (b) Data and Mapper on different nodes.

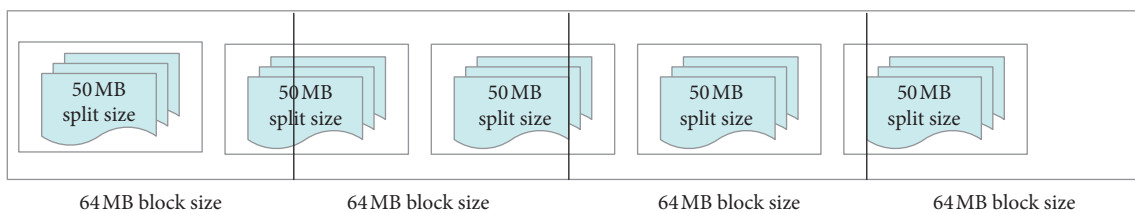


FIGURE 3: Mismatch between split size and block size.

and GPU based on their computing capability. Both phases are explained below in detail.

**4.1. Phase I: Data Distribution amongst Nodes.** In Hadoop, workload is organized in splits which are then distributed amongst nodes in the cluster to be processed. However, this workload is not evenly distributed and hence processors with lower processing capabilities are overwhelmed. In the proposed framework, the input is in the form of images and

is distributed evenly amongst cluster nodes in order to utilize computation and memory resources efficiently. We have developed a novel distribution policy for the even distribution of same size images in splits, where a split contains one or more images. We are focusing on same sized images only, in order to avoid communication overhead when images of different sizes are loaded. With images of different sizes, the ratio calculations explained in Section 4.2 are useless. This idea is mainly inspired from *arrays*, where continuous blocks of the same size are gathered together,

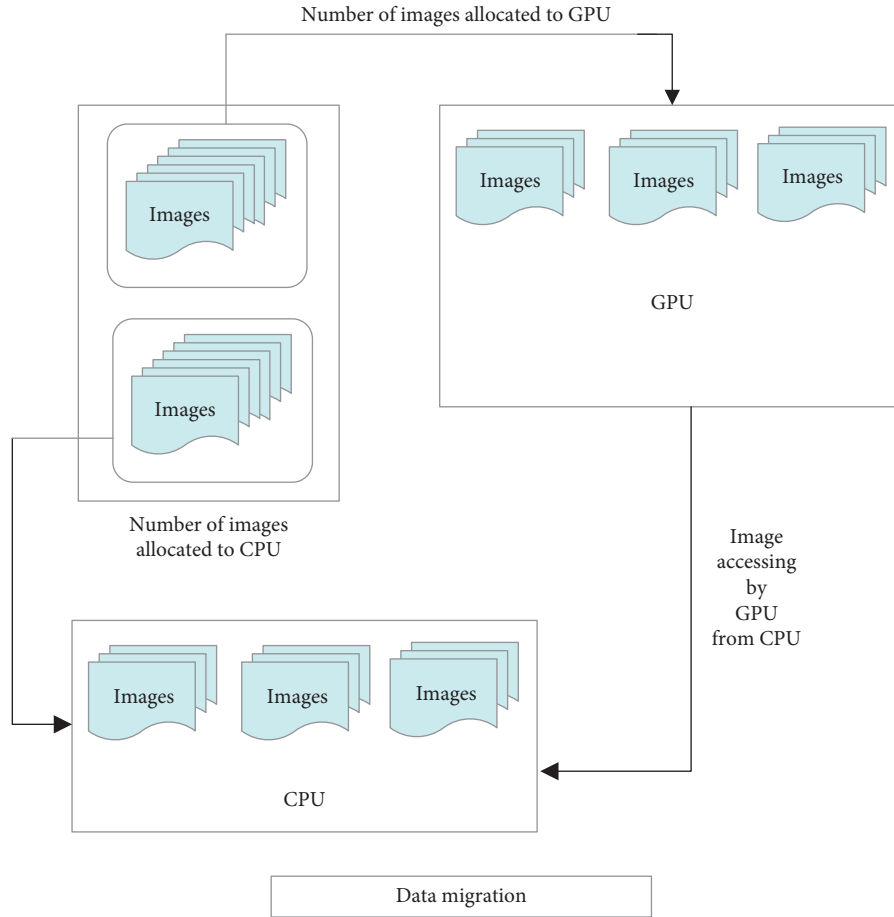


FIGURE 4: Data migration between CPU and GPU.

providing simplicity and performance. The distribution policy does not allow a single image to be distributed to multiple splits, because of data locality issue. The ideal split size is set according to the size of the images, so that an image does not exceed the boundary of that split. Multiple images evenly grouped together in a split and ready to be distributed amongst nodes is shown in Figure 6, and the ideal split size is calculated as per the default block size (i.e., 64 MB in Hadoop) as shown in Figure 7.

To avoid the problem of uneven splits, i.e., when an image is distributed in multiple splits, the split size is set very carefully based on the image size. We first measure the size of one image and then select an input split size where multiple images of that size will be placed. Let  $I$  be the ideal input split size which need to be computed,  $d$  be the default split size in Hadoop,  $s$  be the size of one of the input images,  $n_o$  be the number of images that can be fully accommodated by the default input split,  $T_i$  be the total number of images in a dataset, and  $S_n$  be the total number of splits in which the data are divided equally.  $n_o$  is calculated by dividing  $d$  on  $s$ , ignoring the fractional part by taking the floor.  $I$  is computed by multiplying the image size  $s$  with  $n_o$ , and  $S_n$  is computed by dividing  $T_i$  on  $n_o$ , as shown in equation (1). This equation calculates an ideal input split size, and no image can occur across two input splits. For instance, we have an input image of size 4.2 MB ( $s$ ), and 64 MB is the default input size ( $d$ ), and

then the ideal input split size  $I = 15 \times 4.2 = 63$  MB ( $n_o = \lfloor 64/4.2 \rfloor = 15$ ). To calculate the number of splits, the data should be arranged in  $S_n = 90/15 = 6$ , where 90 is the total number of input images. So, we will have a total of 6 input splits each of size 63 MB to store 90 input images.

$$n_o = \left\lfloor \frac{d}{s} \right\rfloor,$$

$$I = n_o \times s, \quad (1)$$

$$S_n = \frac{T_i}{n_o}.$$

**4.2. Phase II: Distribution of Workload on CPU and GPU within a Node.** In a heterogeneous cluster, every node is equipped with a GPU, which is much faster than the CPU. Therefore, for efficient execution of applications, it is important that tasks are distributed based on the computing capabilities of processors. The proposed workload distribution scheme for heterogeneous Hadoop cluster is shown in Figure 8. A split is a container of a group of images of the same sizes. For every split, the map function is invoked, which takes the input  $\langle key, value \rangle$  pair, where the key contains log file of images in a split and the value contains

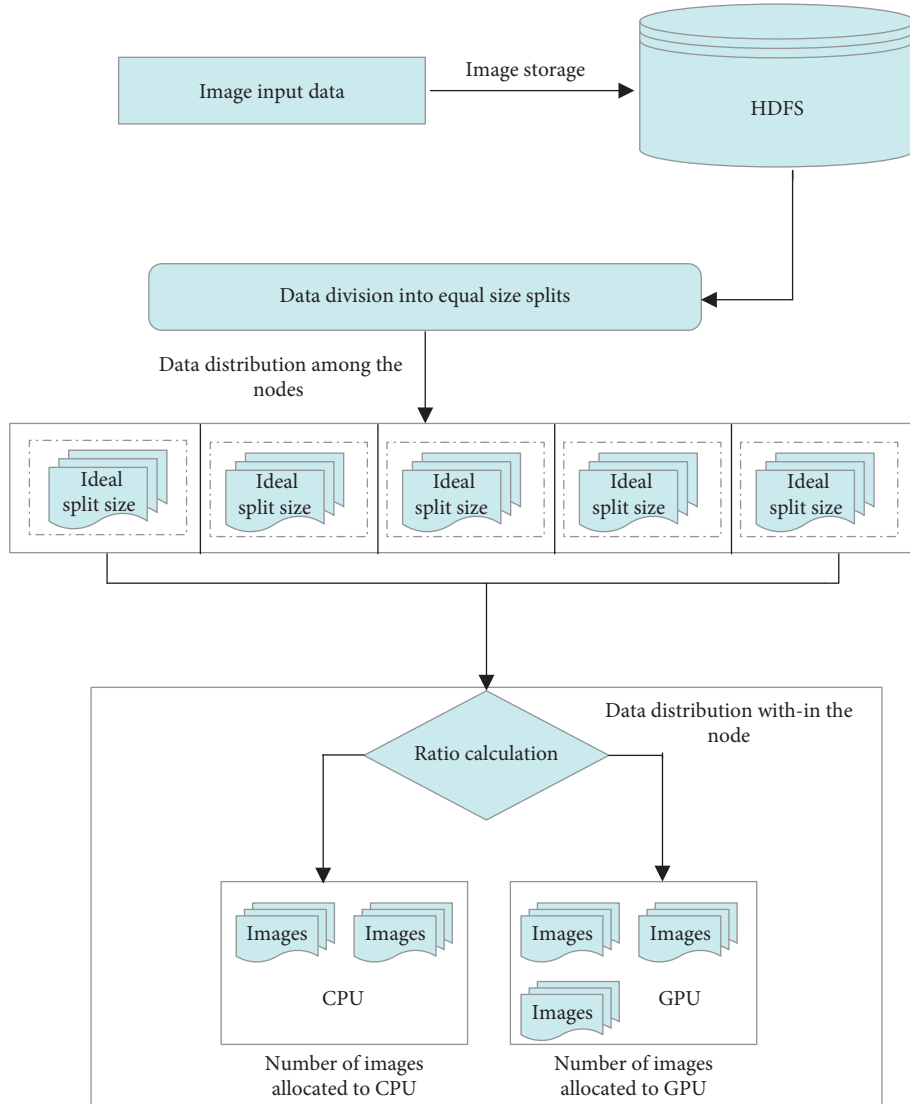


FIGURE 5: Proposed framework for workload distribution.

contents of images (bytes). The map function processes the split and reads each image in the split. In the proposed algorithm, the map function takes the split and checks the ratio for all the available images in that split so that images can be assigned to CPU and GPU according to their computing capability. Before assigning images to the CPU and GPU, a sample image is executed on both the CPU and GPU to find out the execution time of both the processors on a specified algorithm. From this execution time, the processing power of each processor is identified, and a ratio is calculated, demonstrating the number of images assigned to CPU and GPU.

In order to compute the computing capability of processors in a node, we initially assign a raw image to both CPU and GPU to find the execution time of both processors. Let  $c$  be the execution time in CPU and  $g$  be the execution time in GPU to execute an image processing algorithm on a raw input image. We take the ratio of  $\max(g, c)$  and  $\min(g, c)$ . The device with larger execution time (i.e., slower

processor) is assigned a value 1, and the device with smaller execution time (i.e. faster processor) is assigned an integer value of the execution time of slow processor divided by execution time of fast processor. When the processor fetches images, we assign  $nP_x$  number of images to the slower processor and  $nP_y$  images to the faster processor, as shown in the following equation:

$$\begin{aligned}
 x &= 1, \\
 y &= \left( \frac{\max(g, c)}{\min(g, c)} \right), \\
 nP_x &= \left( \frac{x}{x+y} \right) \times n_0, \\
 nP_y &= \left( \frac{y}{x+y} \right) \times n_0.
 \end{aligned} \tag{2}$$

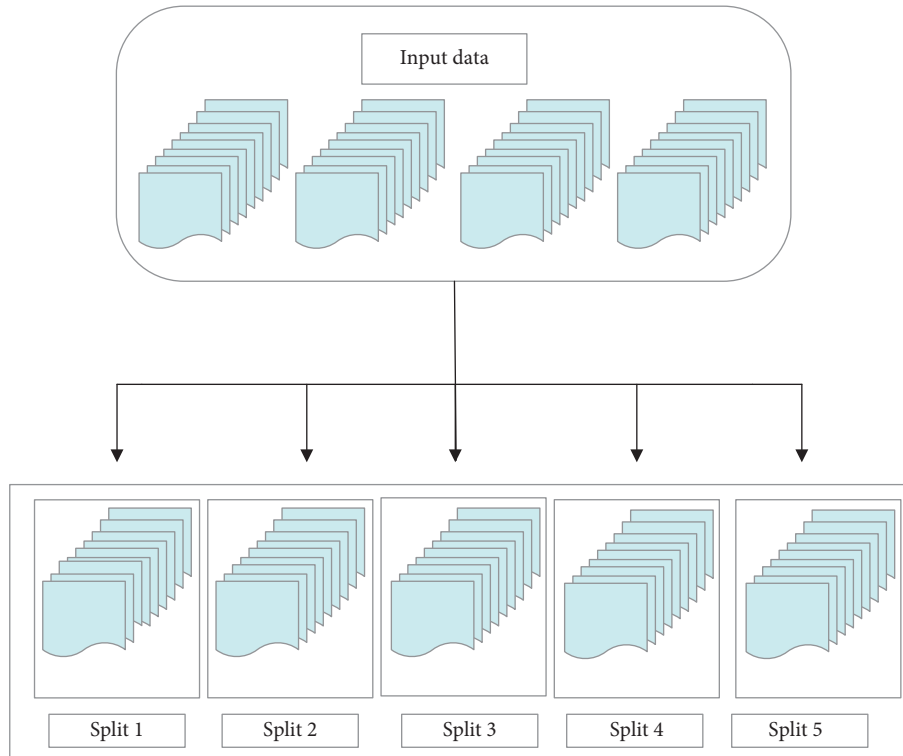


FIGURE 6: Grouping of input data to be partitioned in splits.

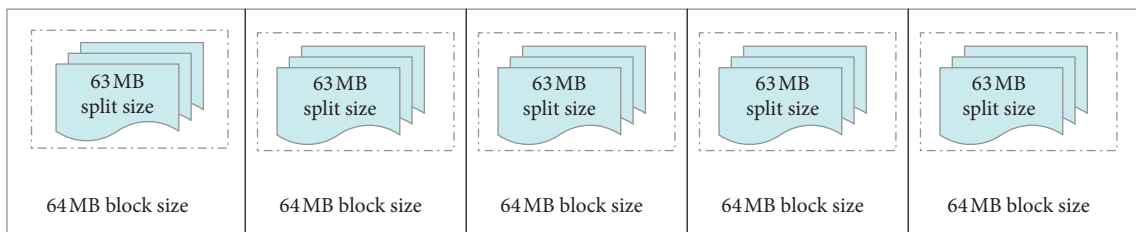


FIGURE 7: Image data partitioning into ideal split size.

Suppose  $n_0$  is 15 (as explained above), GPU execution time ( $g$ ) for a raw image is 25 ms, and CPU execution ( $c$ ) is 160 ms. In this example,  $x=1$  and  $y=160/25=6$ , which means that CPU and GPU are assigned images in the ratio of 1:6. Therefore,  $nP_x=(1/(1+6)) \times 15=2$  and  $nP_y=(6/(6+1)) \times 15=13$ , i.e., in a single fetch, 2 images are assigned to CPU and 13 images are assigned to GPU. This novel distribution of workload ensures that images are distributed based on the computing capability of the processor, which can speedup the execution of images in heterogeneous nodes. The flow chart of the proposed framework is shown in Figure 9.

## 5. Evaluation

In Section 4, the proposed framework is introduced, to efficiently handle the imbalanced workload distribution in a heterogeneous Hadoop cluster for image processing applications, and the implementation of such policy is evaluated by using the commodity computer systems accelerated with GPU. As the heterogeneous systems increase the

performance of applications by processing a massive amount of data in parallel, in the future, these heterogeneous systems will be commonly used and provide efficient execution of programs, by adopting policies for efficient workload distribution. In this section, the evaluation of the proposed programming framework is discussed. The detail of the dataset used for the evaluation is given in Table 1. The test environment includes a master node having a corei5 processor with speed 2.5 GHz, 8 GB RAM, and 4 processing cores. Two worker nodes are used each having a CPU, corei3, 1.8 GHz, 4 GB RAM, and 4 cores and a GPU NVIDIA 802 M, 64 cores, and 2 GB RAM.

*5.1. Processing Images of Different Sizes.* The Figure 10(a) shows the average execution time calculated in milliseconds for edge detection algorithm on four images of different sizes on a CPU and GPU, respectively. This experiment demonstrates that, on CPU, execution time of the application increases when the size of the image is increased. However, on GPU, the increase in execution time when processing an

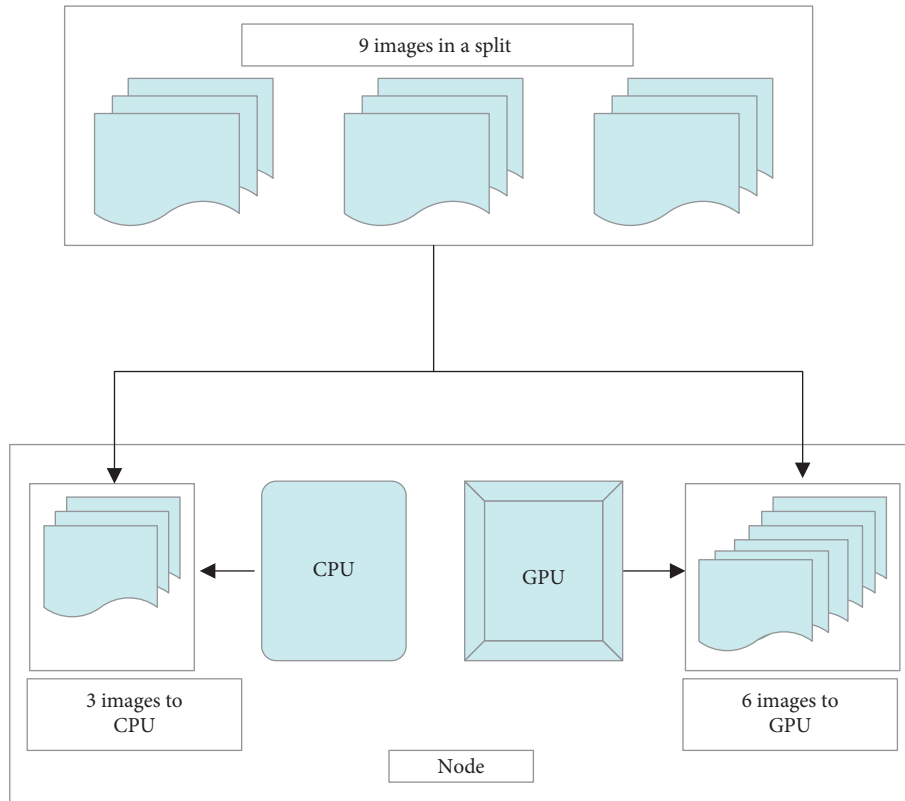


FIGURE 8: Distribution of workload within a node between CPU and GPU according to their computing capabilities.

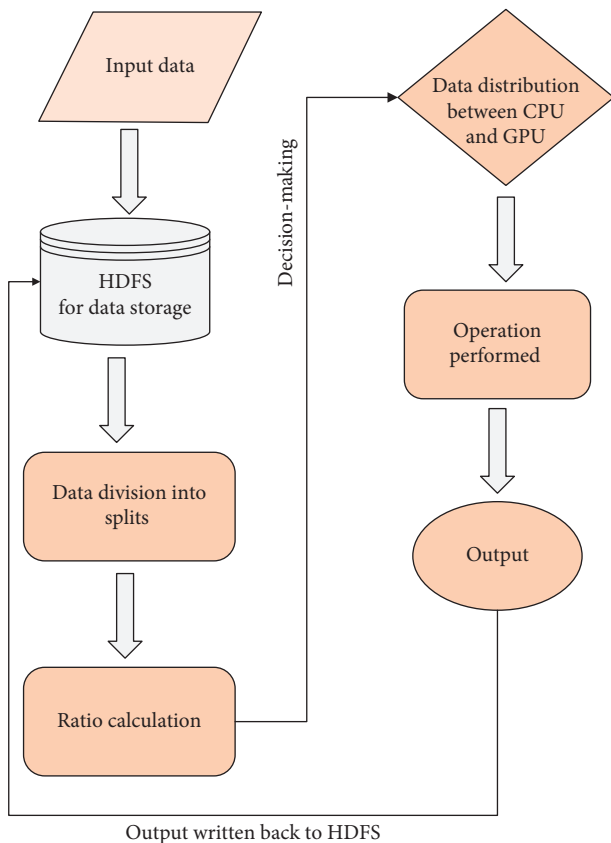


FIGURE 9: Basic flowchart of the proposed framework.

application on a very large image is not significant. This analysis demonstrates that GPU is a better choice for processing larger images.

5.2. *Processing Different Number of Images.* In Figure 10(b), the performance of the application in processing different number of images on a GPU is shown. Y-axis shows execution time in millisecond and x-axis shows the number of images of different sizes. Four images with different resolution sizes are grouped together as 1 image, 2 images, 3 images, and 4 images. For instance, when the application processes a single image of size  $1024 \times 768$ , the execution time on GPU is 42 milliseconds. But, in processing four images of the same size, the GPU takes 51 milliseconds. Similarly, a single image of size  $2560 \times 1440$  takes 69 ms, but processing four images of the same size takes 87 ms. This increase in execution time is mainly because of the communication overhead, as each image is migrated from CPU memory to GPU memory and then the result is written back to the CPU memory.

5.3. *Performance Comparison.* We compare the performance of our approach, i.e., modified split sizes and optimized workload distribution based on the computation capabilities of processors in a node of heterogeneous system with existing state-of-the-art techniques such as HIPI, HIPI executed on GPU, and Hadoop executed on GPU, as shown in Figure 10(c). The proposed framework is processing



TABLE 1: Information of images in the dataset.

S. no.	Image resolution	Total size (MB)	Total images	Individual image size (MB)
1	1024 × 768	216	90	2.4
2	1600 × 900	387	90	4.3
3	1920 × 1080	558	90	6.2
4	2560 × 1440	999	90	11.1

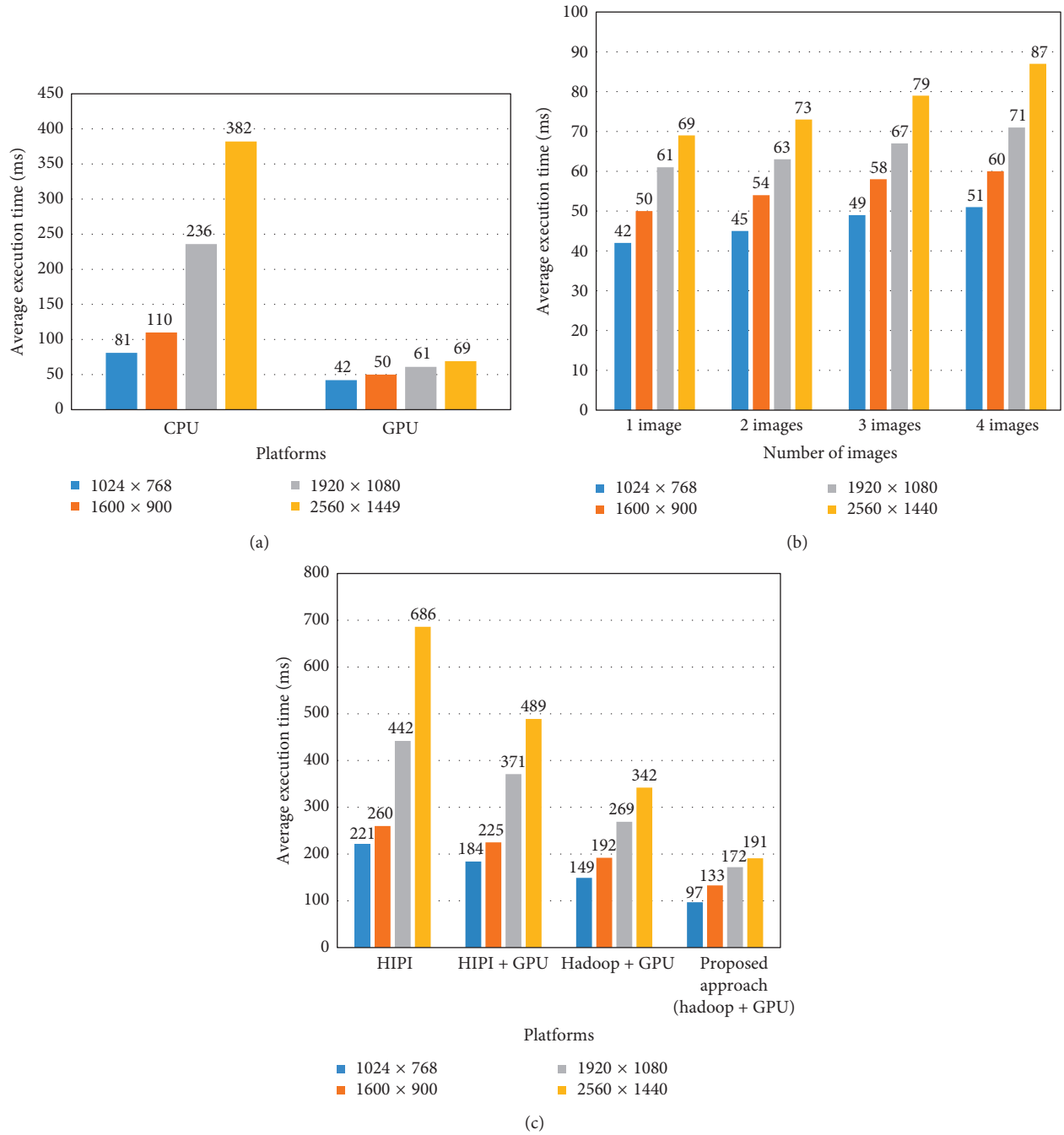
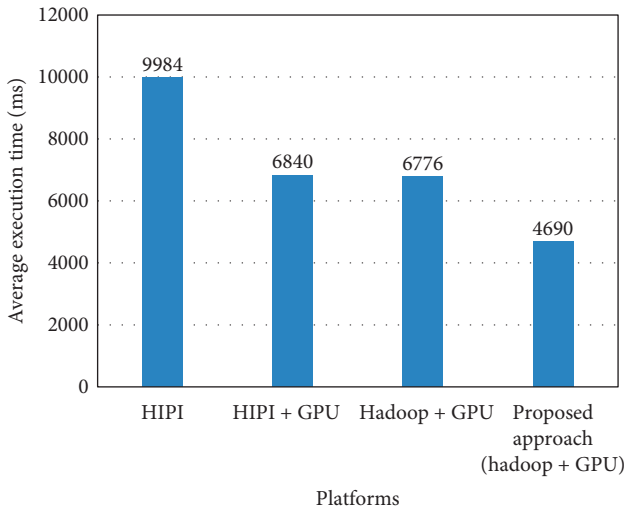
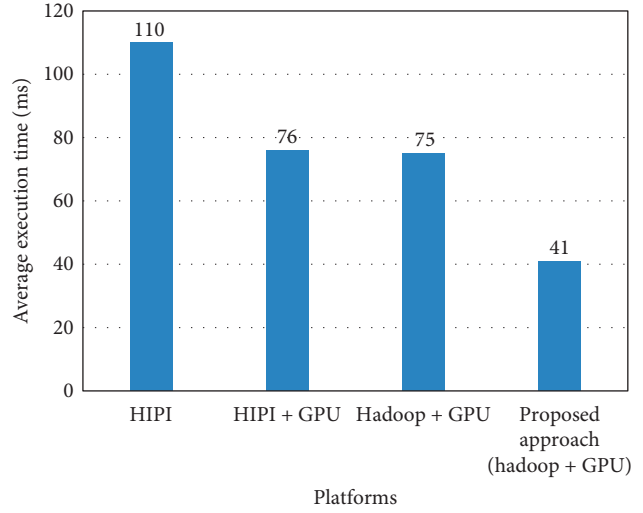


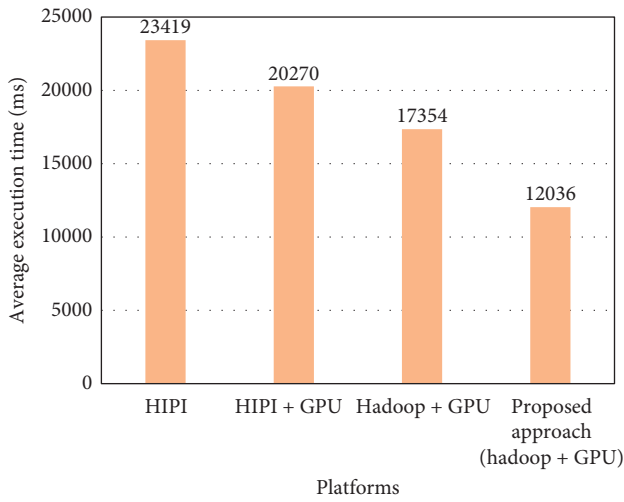
FIGURE 10: The execution time of CPU and GPU, different number of images, and average performance comparison with existing platforms. (a) Average execution time in CPU and GPU with different number of images for edge detection algorithm. (b) Average execution time of application in processing different number of images on a GPU. (c) Average execution time (milliseconds) for proposed framework (Hadoop + GPU) vs existing platforms.



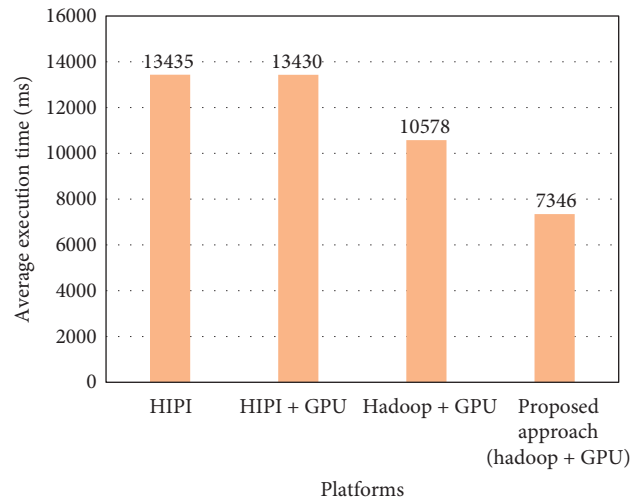
(a)



(b)



(c)



(d)

FIGURE 11: (a) Execution time of a split on a node. (b) Execution time calculated for each split. (c) Execution time of image dataset processed by edge detection application. (d) Execution time of a single image from split to processor.

images faster than the other state-of-the-art frameworks. For instance, processing 90 images of resolutions  $2560 \times 1440$  each of size 11.1 MB takes 686 ms in HIPI, but only 191 ms in the proposed framework, i.e., a speedup of 3.5, is achieved. This speedup is possible mainly because of the efficient workload distribution and the efficient arrangement of images in input splits which results into the efficient utilization of resources.

**5.4. Effect of Assigning Different Loads to CPU and GPU.** In the proposed approach, we have developed a novel technique to compute the ratio of images to be assigned to CPU and GPU in a heterogeneous node. The performance of this calculation is shown in Figure 11(a) along with comparison with other state-of-the-art techniques. It is observed that application processing the dataset of different images of different sizes executes efficiently on the proposed framework. For instance, the speedup in the proposed approach is

2.12x compared to HIPI. Hence, it is proved that, by efficient load balancing techniques in heterogeneous systems, we can increase the performance by two times.

**5.5. Execution Time of Each Split.** Input data are divided in different input splits and distributed amongst nodes, and from each split, images are accessed and processed. The average execution time taken by images of sizes  $(1024 \times 768, 1600 \times 900, 1920 \times 1080, \text{ and } 2560 \times 1440)$  in an input split is shown in Figure 11(b) for all the four platforms while processing edge detection application. It is demonstrated that the proposed framework processes the images more than two and half times faster than HIPI.

**5.6. Execution Time in Processing the Whole Dataset.** In Figure 11(c), an average of total execution time for the whole dataset shown in Table 1 is calculated for image processing

application performing edge detection operation. The total execution time includes partitioning of image data into splits, distribution amongst the nodes, and on each node, further distribution between CPU and GPU, and after processing, the result is written back to the HDFS. From the figure, it is clearly shown that, by adopting the proposed approach (Hadoop + GPU) where data is divided into ideal split size and then distributed to the processors according to their computing capabilities, the total execution time calculated in milliseconds is two times less than the other existing platforms.

*5.7. Minimization of the Communication Overhead by Ideal Split Size Selection.* The communication overhead of images during processing is shown in Figure 11(d). The access time of an image in this figure includes the time from split to the assigned processor. The execution time taken by a single image is recorded. From the experiment, it is observed that, by arranging the data local to the computing processor, the images can be easily accessed and processed. As HIPI has an overhead of data compression and decompression; therefore, the results of both HIPI and HIPI + GPU are relatively same and high. By using Hadoop + GPU, the image access time is less compared to the HIPI and HIPI + GPU, but by overcoming the data locality issue in the proposed framework (Hadoop + GPU), the image access time is almost two times reduced.

## 6. Conclusion and Future Work

Distributed systems provide parallel frameworks where massive amount of data can be efficiently processed. Hadoop is one of those frameworks that provides large amount of data storage and effective computational capability to handle massive amount of data in a parallel and distributed manner by using the cluster of commodity computer systems. These clusters also contain GPUs, as they are specialized to handle SIMD efficiently. For image processing applications or applications involving the process of big data in different domains such as healthcare, heterogeneous systems are commonly used and have shown improvement over single processor and distributed systems. However, imbalanced workload distribution between the processor causes data locality and inefficient workload distribution between slow and fast processors can affect the performance of applications. To deal with this imbalanced workload distribution in a heterogeneous cluster, this paper has proposed a novel technique of dividing data into ideal input splits so that an image is included in one split and does not exceed the boundary of that split. This paper also introduces a technique of distributing data as per the computing power of the processors in the node. This distribution maximizes the utilization of available resources and tackles the issue of data migration between the fast and slow computing processors. The results have demonstrated that the proposed framework achieves almost two times improvement compared to the current state-of-the-art programming frameworks. The proposed framework provides an efficient mechanism to

compute an ideal split size that is suitable for fixed size images but does not provide support for partitioning variable size images into splits. In the future, we will investigate techniques that can compute ideal split size for variable sized images. In the future, we will investigate techniques that can compute ideal split size for variable sized images which will enable us to process images from sources. The proposed model can be extended for big data applications which have inherent data-level parallelism in the form of arrays, matrices, images, or tables.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

- [1] L. Baumstark and L. Wills, "Exposing data-level parallelism in sequential image processing algorithms," in *Proceedings of the Ninth Working Conference on Reverse Engineering*, pp. 245–254, New York, NY, USA, 2002.
- [2] B. Vinter, "Embarrassingly parallel applications on a java cluster," in *Proceedings of the 8th International Conference on High-Performance Computing and Networking, HPCN Europe 2000*, Springer-Verlag, London, UK, pp. 614–617, 2000.
- [3] B. Furht, *SIMD (Single Instruction Multiple Data Processing)*, Springer US, Boston, MA, USA, 2008.
- [4] I. Uddin, "Multiple levels of abstractions in the simulation of microthreaded many-core architectures," *Open Journal of Modelling and Simulation*, vol. 3, 2015.
- [5] M. Girkar and C. D. Polychronopoulos, "Extracting task-level parallelism," *ACM Transactions on Programming Languages and Systems*, vol. 17, pp. 600–634, 1995.
- [6] I. Uddin, "High-level simulation of concurrency operations in microthreaded many-core architectures," *GSTF Journal on Computing (JoC)*, vol. 4, p. 21, 2015.
- [7] I. Uddin, "One-IPC high-level simulation of microthreaded many-core architectures," *International Journal of High Performance Computing Applications*, vol. 4, 2015.
- [8] R. Riesen and A. B. Maccabe, *MIMD (Multiple Instruction, Multiple Data) Machines*, Springer US, Boston, MA, USA, 2011.
- [9] L. Hu, X. Che, and S.-Q. Zheng, "A closer look at GPGPU," *ACM Computing Surveys*, vol. 48, 2016.
- [10] M. Fatima, "Reliable and energy efficient mac mechanism for patient monitoring in hospitals," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, 2018.
- [11] J. V. Bibal Benifa, "Performance improvement of Mapreduce for heterogeneous clusters based on efficient locality and replica aware scheduling (ELRAS) strategy," *Wireless Personal Communications*, vol. 95, no. 3, pp. 2709–2733, 2017.
- [12] N. Elgandy and A. Elragal, "Big data analytics: a literature review paper," in *Advances in Data Mining, Applications and Theoretical Aspects*, P. Perner, Ed., pp. 214–227, Springer International Publishing, Berlin, Germany, 2014.

- [13] A. Khan, M. A. Gul, M. I. Uddin et al., "Summarizing online movie reviews: a machine learning approach to big data analytics," *Scientific Programming*, vol. 2020, pp. 1–14, 2020.
- [14] F. Aziz, H. Gul, I. Muhammad, and I. Uddin, "Link prediction using node information on local paths," *Physica A: Statistical Mechanics and its Applications*, vol. 2020, Article ID 124980, 2020.
- [15] J. Zhu, H. Jiang, J. Li, E. Hardesty, K.-C. Li, and Z. Li, "Embedding GPU computations in hadoop," *International Journal of Networked and Distributed Computing*, vol. 2, pp. 211–220, 2017.
- [16] W. Chen, S. Xu, H. Jiang et al., "GPU computations on hadoop clusters for massive data processing," in *Proceedings of the 3rd International Conference on Intelligent Technologies and Engineering Systems (ICITES2014)*, J. Juang, Ed., Springer International Publishing, Berlin, Germany, pp. 515–521, 2016.
- [17] S. Amin, M. I. Uddin, S. Hassan et al., "Recurrent neural networks with TF-IDF embedding technique for detection and classification in tweets of dengue disease," *IEEE Access*, vol. 8, pp. 131522–131533, 2020.
- [18] M. I. Uddin, S. A. A. Shah, and M. A. Al-Khasawneh, "A novel deep convolutional neural network model to monitor people following guidelines to avoid COVID-19," *Journal of Sensors*, vol. 2020, pp. 1–15, 2020.
- [19] Z. Ullah, A. Zeb, I. Ullah et al., "Certificateless proxy re-encryption scheme (CPRES) based on hyperelliptic curve for access control in content-centric network (CCN)," *Mobile Information Systems*, vol. 2020, pp. 1–13, 2020.
- [20] A. Khan, I. Ibrahim, M. I. Uddin et al., "Machine learning approach for answer detection in discussion forums: an application of big data analytics," *Scientific Programming*, vol. 2020, 2020.
- [21] F. Aziz, T. Ahmad, A. H. Malik, M. I. Uddin, S. Ahmad, and M. Sharaf, "Reversible data hiding techniques with high message embedding capacity in images," *PLoS One*, vol. 15, no. 1–24, 2020.
- [22] B. Sharma, R. Thota, N. Vydyanathan, and A. Kale, "Towards a robust, real-time face processing system using CUDA-enabled GPUS," in *Proceedings of the 2009 International Conference on High Performance Computing (HiPC)*, pp. 368–377, New York, NY, USA, 2009.
- [23] H. Patel and K. Panchal, "Incorporating CUDS in hadoop image processing interface for distributed image processing," in *Proceedings of the International Journal of Advance Research and Innovative Ideas*, pp. 93–98, New York, NY, USA, 2016.
- [24] C. Sweeney, L. Liu, S. Arietta, and J. Lawrence, "Hipi: a hadoop image processing interface for image-based mapreduce tasks," 2011.
- [25] R. Malakar and N. Vydyanathan, "A cuda-enabled hadoop cluster for fast distributed image processing," in *Proceedings of the 2013 National Conference on Parallel Computing Technologies (PARCOMPTECH)*, pp. 1–5, New York, NY, USA, 2013.
- [26] M. I. Uddin, N. Zada, F. Aziz et al., "Prediction of future terrorist activities using deep neural networks," *Complexity*, vol. 2020, pp. 1–16, 2020.
- [27] T. Liu, Y. Liu, Q. Li et al., "Seip: system for efficient image processing on distributed platform," *Journal of Computer Science and Technology*, vol. 30, pp. 1215–1232, 2015.
- [28] M. S. Gowda and V. G. Hulyal, "Parallel image processing from cloud using cuda and hadoop architecture: a novel approach," 2015.
- [29] A. K. Singh, A. Prakash, K. R. Basireddy, G. V. Merrett, and B. M. Al-Hashimi, "Energy-efficient run-time mapping and thread partitioning of concurrent OPENCL applications on CPU-GPU MPSOCS," *ACM Transactions on Embedded Computing Systems*, vol. 16, 2017.
- [30] W. Burger, M. J. Burge, and M. J. Burge, *Scale-Invariant Feature Transform (SIFT)*, Springer, London, UK, 2016.
- [31] J. Kim, H. Kim, J. H. Lee, and J. Lee, "Achieving a single compute device image in OPENCL for multiple GPUS," *SIGPLAN Not*, vol. 46, pp. 277–288, 2011.
- [32] P. Dadheech, D. Goyal, S. Srivastava, and A. Kumar, "Performance improvement of heterogeneous hadoop clusters using query optimization," *SSRN Electronic Journal*, vol. 46, 2018.
- [33] N. S. Naik, A. Negi, and V. Sastry, "Performance improvement of MapReduce framework in heterogeneous context using reinforcement learning," *Procedia Computer Science*, vol. 50, pp. 169–175, 2015.
- [34] C. Lai, Y. Chen, X. Shi, M. Huang, and G. Chen, "Performance improvement on heterogeneous platforms: a machine learning based approach," in *Proceedings of the 2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, IEEE, Berlin, Germany, 2018.
- [35] P. Czarnul, J. Proficz, and K. Drypczewski, "Survey of methodologies, approaches, and challenges in parallel programming using high-performance computing systems," *Scientific Programming*, vol. 2020, pp. 1–19, 2020.
- [36] A. Rubio-Largo, J. C. Preciado, and L. Iribarne, "Data-driven computational intelligence for scientific programming," *Scientific Programming*, vol. 2019, pp. 1–4, 2019.
- [37] A. Kanan, F. Gebali, A. Ibrahim, and K. F. Li, "Low-complexity scalable architectures for parallel computation of similarity measures," *Scientific Programming*, vol. 2019, 2019.
- [38] W. Ma, W. Yuan, and X. Hu, "Implementation and optimization of a CFD solver using overlapped meshes on multiple MIC coprocessors," *Scientific Programming*, vol. 2019, pp. 1–12, 2019.
- [39] T. White, *Hadoop: The Definitive Guide*, O'Reilly Media, Inc., New York, NY, USA, 1st edition, 2009.
- [40] M. Aldinucci, M. Danelutto, P. Kilpatrick, and M. Torquati, "Fastflow: high-level and efficient streaming on multi-core," *Programming Multi-core and Many-core Computing Systems, ser. Parallel and Distributed Computing*, vol. 13, 2012.
- [41] L. Dagum and R. Menon, "OPENMP: an industry-standard api for shared-memory programming," *Computing in Science & Engineering*, vol. 5, pp. 46–55, 1998.
- [42] B. Nichols, D. Buttler, and J. P. Farrell, *Pthreads Programming*, O'Reilly & Associates, Inc., Sebastopol, CA, USA, 1996.
- [43] A. Munshi, B. Gaster, T. G. Mattson, J. Fung, and D. Ginsburg, *OpenCL Programming Guide*, Addison-Wesley Professional, Boston, MA, USA, 1st edition, 2011.
- [44] J. S. Harbour, D. Calkins, and R. Meuth, *Multi-Core Programming with CUDA and OpenCL*, Course Technology Press, Boston, MA, USA, 1st edition, 2011.
- [45] D. Shreiner and T. K. O. A. W. Group, *OpenGL Programming Guide: The Official Guide to Learning OpenGL, Versions 3.0 and 3.1*, Addison-Wesley Professional, New York, NY, USA, 7th edition, 2009.
- [46] J. Dean and S. Ghemawat, "Mapreduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, pp. 107–113, 2008.
- [47] M. Zaharia, R. S. Xin, P. Wendell et al., "Apache spark: a unified engine for big data processing," *Communications of the ACM*, vol. 59, pp. 56–65, 2016.
- [48] R. V. Guha, D. Brickley, and S. MacBeth, "Schema.org: evolution of structured data on the web," *Queue*, vol. 13, p. 10, 2015.

- [49] D. Suciú, *Information Organization and Databases*, Kluwer Academic Publishers, Norwell, MA, USA, 2000.
- [50] J. P. Isson, *Unstructured Data Analytics: How to Improve Customer Acquisition, Customer Retention, and Fraud Detection and Prevention*, Wiley Publishing, New York, NY, USA, 1st edition, 2018.
- [51] H. Tan and L. Chen, "An approach for fast and parallel video processing on Apache hadoop clusters," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1–6, New York, NY, USA, 2014.
- [52] C. Ryu, D. Lee, M. Jang, C. Kim, and E. Seo, "Extensible video processing framework in Apache hadoop," in *Proceedings of the 2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, pp. 305–310, New York, NY, USA, 2013.
- [53] M. Husain, A. K. Sabarad, H. Hebballi, S. M. Nagaralli, and S. Shetty, "Counting occurrences of textual words in lecture video frames using Apache hadoop framework," in *Proceedings of the 2015 IEEE International Advance Computing Conference (IACC)*, pp. 1144–1147, London, UK, 2015.
- [54] H. Tan and L. Chen, "An approach for fast and parallel video processing on Apache hadoop clusters," in *Proceedings of the 2014 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, London, UK, 2014.
- [55] U. S. N. Raju, S. George, V. S. Praneeth, R. Deo, and P. Jain, "Content based image retrieval on hadoop framework," in *Proceedings of the 2015 IEEE International Congress on Big Data*, pp. 661–664, New Jearsey, NJ, USA, 2015.
- [56] S. Cook, *CUDA Programming: A Developer's Guide to Parallel Computing with GPUs*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2013.
- [57] Z. Wang, P. Lv, and C. Zheng, "Cuda on hadoop: a mixed computing framework for massive data processing," in *Foundations and Practical Applications of Cognitive Systems and Information Processing*, F. Sun, D. Hu, and H. Liu, Eds., pp. 253–260, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
- [58] B. He, W. Fang, Q. Luo, N. K. Govindaraju, and T. Wang, "Mars: a mapreduce framework on graphics processors," in *Proceedings of the 17th International Conference on Parallel Architectures and Compilation Techniques, PACT '08*, ACM, New York, NY, USA, pp. 260–269, 2008.
- [59] C. Hong, D. Chen, W. Chen, W. Zheng, and H. Lin, "MAPCG: Writing Parallel Program Portable between CPU and GPU," in *Proceedings of the 2010 19th International Conference on Parallel Architectures and Compilation Techniques (PACT)*, pp. 217–226, New York, NY, USA, 2010.
- [60] M. Elteir, H. Lin, W. Feng, and T. Scogland, "Streammr: an optimized mapreduce framework for amd GPUS," in *Proceedings of the 2011 IEEE 17th International Conference on Parallel and Distributed Systems*, pp. 364–371, New York, NY, USA, 2011.
- [61] J. A. Stuart and J. D. Owens, "Multi-GPU mapreduce on GPU clusters," in *Proceedings of the 2011 IEEE International Parallel Distributed Processing Symposium*, pp. 1068–1079, Berlin, Germany, 2011.
- [62] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, "The hadoop distributed file system," in *Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST), MSST '10*, IEEE Computer Society, Washington, DC, USA, pp. 1–10, 2010.
- [63] I. Uddin, A. Baig, and A. A. Minhas, "A controlled environment model for dealing with smart phone addiction," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 9, 2018.
- [64] S. A. A. Shah, I. Uddin, F. Aziz, S. Ahmad, M. A. Al-Khasawneh, and M. Sharaf, "An enhanced deep neural network for predicting workplace absenteeism," *Complexity*, vol. 2020, pp. 1–12, 2020.



## Research Article

# Research on Management of Doctor-Patient Risk and Status of the Perceived Behaviors of Physician Trust in the Patient in China: New Perspective of Management of Doctor-Patient Risk

Jiang Jie Sun <sup>1,2</sup>, Zhi Bo Zheng,<sup>3</sup> Xue Li Jiang,<sup>1</sup> Wei Wei Hu,<sup>1</sup> Jun Liu,<sup>4</sup> Nan Zhen Ma,<sup>5</sup> Meng Ying Li,<sup>6</sup> Xiao-Zhe Yan,<sup>7</sup> Cheng sen He,<sup>6,7</sup> and Li Ping Zhang <sup>4,6</sup>

<sup>1</sup>Health Management College, Anhui Medical University, Hefei 230032, China

<sup>2</sup>Anhui Higher Medical Education Cooperation Committee, Hefei 230032, China

<sup>3</sup>Department of Mathematics, Baoshan University, Baoshan 678000, China

<sup>4</sup>School of Humanistic Medicine, Anhui Medical University, Hefei, Anhui 230032, China

<sup>5</sup>Hospital of Anhui Medical University, Hefei, Anhui, China

<sup>6</sup>Clinical Medical College, Anhui Medical University, Hefei 230601, China

<sup>7</sup>Department of Psychology, Anhui Medical University, Hefei, Anhui, China

Correspondence should be addressed to Li Ping Zhang; zhangliping@ahmu.edu.cn

Received 15 July 2020; Accepted 18 September 2020; Published 8 October 2020

Academic Editor: M Javaid

Copyright © 2020 Jiang Jie Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the situation of physician trust in the patient (PTP), we explored the differences in perceived behaviors of physician trust in the patient (PBPTP). We used the PTP scale as a research tool, taking physicians of the hospitals in Anhui region as the research object to carry out the investigation of PTP, Python software was applied to explore the status of PTP, and the differences of PBPTP distribution rate with different demographic characteristic variables were compared by testing based on theory of planned behavior. We get six results as follows: (1) the overall PTP level was low, and nearly 50% of doctors doubt the integrity of patients. "Patients will not be driven by improper interests" becomes the most reluctant problem or the most distrustful option for doctors. (2) In terms of patients' participation in disease management and regular follow-up visits, PTP rate in male was higher than that in female ( $P_s < 0.018$ ). (3) PBPTP was affected by age ( $P_s < 0.017$ ). (4) In terms of the behavior of patients who did not follow the treatment plans, the PTP rate of postgraduates and above physicians was higher than that of undergraduates and below ( $P = 0.017$ ). (5) In terms of providing diagnosis and treatment information, timely notification of illness, medication information, doctor-patient communication behaviors, and compliance with doctors' treatment plans, PBPTP was affected by doctors' professional titles and annual income levels ( $P_s \leq 0.001$ ). At the same time, PTP levels of different professional titles showed differences in patients' respect for doctors' time and bottom line ( $P_s \leq 0.001$ ). (6) In terms of doctor-patient communication behaviors, PBPTP was affected by physician departments ( $P \leq 0.001$ ). Hence, demographic characteristics variable may be one of the factors affecting PBPTP, and PBPTP is associated with doctor-patient risk. It makes sense for us to propose a new model of physician-patient risk management from the perspective of PTP about "official-individual-social" triple action.

## 1. Introduction

The article "Cultivating the good root of doctor-patient trust" in the People's Daily pointed out that the current domestic doctor-patient trust is fragile; medical violence and injury to physicians occur frequently; the mutual trust and harmonious relationship between doctors and

patients (RBDP) that originally "health is related to life" has changed. How to enhance the trust between doctors and patients and avoid the risk that doctors and patients have become a research hotspot in academia. The outline of the "healthy China 2030" program emphasizes on strengthening humanistic care for medical services and building a harmonious RBDP. These articles proposed that the doctor-



patient trust crisis is the important reason for the deterioration of RBDP, and the lack of doctor-patient trust is the root cause [1–3]. Through the questionnaire survey, it is found that the current doctor-patient trust status and doctor-patient trust level have an impact on RBDP. The results are as follows (Figure 1).

The results show that the overall level of doctor-patient trust is not high. Both doctors and patients highly believe that doctor-patient trust has a great impact on RBDP, while doctors pay more attention to doctor-patient trust. The early manifestation of the doctor-patient risk is the deterioration of the RBDP. Therefore, doctor-patient trust is a direct factor influencing doctor-patient risk. Krot conducted an online study about polish patients and concluded that doctor-patient trust positively affects RBDP [4].

The research on doctor-patient trust originated earlier, mainly including the research of doctor-patient trust connotation and doctor-patient trust measurement methods.

Doctor-patient trust is a belief, an expectation, a fulcrum, or a feeling of security (reflecting the emotional nature), etc. For example, Thom proposed that doctor-patient trust is the patient's expectation of the medical services provided by the doctor [5]; Haslam believed that trust is the fulcrum of RBDP [6]; Howe believed that doctor-patient trust means that the patient trusts the doctor [7]. This view dominated for a long time; however, the degree of doctors' trust in patients affects the perception of communication quality and is directly related to doctor-patient risk [8]. At present, there is almost no research on the status of doctor-patient trust and doctor-patient risk management. This paper explores the new idea of doctor-patient risk management from the perspective of doctor-patient trust, which has certain theoretical value.

Doctor-patient trust measurement research originated in 1973. Wallston introduced the "interpersonal trust scale" into the medical field and proposed the "nurse trust patient scale", which apparently solved the problem of nurse trust patient measurement [9]. Because of the lack of reliability and validity as well as the Cronbach coefficient of retest reliability is only 0.32, the scale is considered invalid. In 1991, Anderson first proposed the scale of trusting doctors, which includes three dimensions, trustworthiness, confidence, and information, and 11 items. Using quantitative methods to measure the trust of patients in doctors has become one of the main reference standards for such studies since that [10]. In 2015, Gopichandran proposed the trust doctor scale from the five dimensions of the doctor's perception, treatment assurance, confidence in the doctor, loyalty to him/her, and respect for him/her and verified that the overall model had a good statistical fitting effect [11]. However, most of them lack integrity reliability and validity test [12–14]. Until 2016, Dong translated the physician trust patient scale developed by Thom and revised the Chinese version [15]. The application of measurement of doctor-patient trust appears in China.

As a contradiction between the supply of medical resources in China and the growing demand for medical services in patients have intensified, the risks of doctors and patients have gradually increased, and the tension between

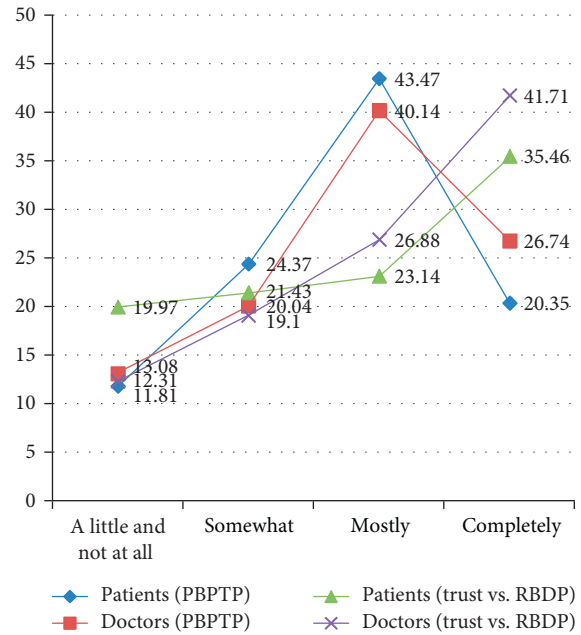


FIGURE 1: The results of PBPTP and trust level vs. RBDP (%). Note. RBDP, relationship between doctors and patients.

doctors and patients has become a social hotspot. Trust between doctors and patients is the basis of building a harmonious RBDP as well as an important guarantee for obtaining the best medical effect [16]. The crisis of doctor-patient trust has become an urgent social problem in China while the study of doctor-patient risk in Anhui has certain representativeness in the country. Investigating and studying the current situation of physicians' trust in patients in Anhui hospitals is a key step in solving this core problem. The Theory of Planned Behavior holds that human behavior is the result of deliberate and planned thinking and that perceptual behavior, attitude, and subjective norms affect behavioral intention and contribute to the process of actual behavior [17]. Based on this, this study explores the status quo of cognitive of physician trust in the patient (PBPTP) and the representation of PBPTP in demographic variables and then explores a new doctor-patient risk management model based on doctor-patient Trust. Thom put forward the importance of PBPTP and pointed out that doctors' trust in patients can positively affect patients' trust in doctors, and the lack of trust in related doctors will lead patients to think that doctors' behaviors is lax and negative and may affect patients' behavior [18]. He proposed that the current RBDP is tense [19], and Sun interpreted the economic incentives for the frequent doctor-patient risks [20]. On this basis, we propose Hypothesis 1.

*Hypothesis 1.* The overall trust of physicians in trusting patients is not high; Sun proposed that physicians' cognitive level of RBDP evaluation is affected by factors such as age, education, job title, and income level [21]. Some studies have proposed that the risk of surgical doctors and patients is relatively high [22], and the conflict rate between doctors

and patients may negatively affect the PTP rate. On this basis, we propose Hypothesis 2.

*Hypothesis 2.* PBPTP differs in gender, age, job title, education level, and annual income level; Zhang's research shows that only 2.4% of physicians are satisfied with their working environment [23]. Doctors, who are dissatisfied with their working environments, are more likely to avoid high-risk diagnosis and treatment at work, whose attitude may further stimulate patients' dissatisfaction with doctors and affect the level of doctor-patient mutual trust. However, because of different departments, the working environment is relatively different, so we propose Hypothesis 3 here.

*Hypothesis 3.* There are department differences in PBPTP.

## 2. Objects and Methods

*2.1. Research Samples.* This paper took doctors from 10 hospitals in Anhui province as the objects of investigation and conducted a questionnaire survey on doctors from various departments including internal medicine, surgery, obstetrics and gynecology, and pediatrics (choose the opportunity of weekly meeting of corresponding hospital to conduct questionnaire survey). The research sample size is not less than 10 times of the number of observed variables to ensure the validity of the research results [24]. Therefore, the sample size selected in this study is 30 times more than the observed variables. A total of 350 copies were distributed and 329 valid questionnaires were collected, with an effective recovery rate of 94%. The survey was conducted by postgraduate students of psychology, lecturer of psychology, and associate professor of psychology. Before the questionnaire was distributed, the respondents received a unified professional training and obtained the informed consent of the respondents.

### 2.2. Research Tools

*2.2.1. Physician Trust in the Patient Scale (PTP Scale).* This scale was developed by Thom et al. [5]. Until 2016, Dong translated the physician trust patient scale developed by Thom and revised the Chinese version [15]. In 2018, Sun used the doctor-patient trust scale to explore the relationship between doctor-patient trust and mental resilience of doctors and concluded that the higher degree of doctor-patient trust, the better doctor-patient relationship [25]. This scale is mainly used to measure the level of PBPTP, including patient roles (8 items in total) and respect for interpersonal relationships (4 items in total). Each item has 5 options from (no trust at all) to (full trust), measuring the levels of doctors' perception of patients' trust. The Cronbach coefficient of the scale is 0.92, which has a good reliability. Each item of this scale reflects PBPTP, and the reliability and validity of each item is high, meeting the needs of this study.

*2.2.2. Self-Compiled Demographic Data Questionnaire.* Demographic information about the respondents (gender, age, education level, professional title, and income level)

and hospital departments were included (internists, surgeon, gynecologist, pediatrics, and other clinical departments).

*2.2.3. Statistical Methods.* After all questionnaires were reviewed by two graduate students, EpiData3.1 was used to input data, python software was used to explore the status quo of PTP, SPSS19.0 software was used for analysis [26], frequency and percentage of counting data were used for statistical description, and chi-squared test was used to compare the differences of PBPTP among different demographic characteristics.  $P < 0.05$  was considered statistically significant.

## 3. Results

*3.1. Descriptive Statistics.* This study collected data from 329 doctors, including 162 males (49.24%) and 167 females (50.76%). This sample was predominantly doctors from graduate/professional school (69.60%), but there were also college graduate (26.44%) and junior college doctors (3.96%). 95 people (28.88%) have primary professional titles, 81 people (24.62%) have intermediate professional titles, 105 people (31.91%) have sub-senior professional titles, and 48 people (14.59%) have senior professional titles. There were 72 patients (21.88%) in internal medicine, 72 patients (21.88%) in surgery, 62 patients (18.84%) in gynecology, 64 patients (19.45%) in pediatrics, and 59 patients (17.95%) in other departments. There were 90 persons (27.36%) with an annual income of less than ¥70,000, 104 persons (31.61%) with an annual income of between ¥70,000 and ¥120,000, and 135 persons (41.03%) with an annual income of 120,000 (¥) and above; average age was 37.8, standard deviation was 11.2, with the youngest being 23 and the oldest 61.

*3.2. Distribution of Perceived Behaviors of Doctors Trusting Patients.* This study uses Python-2.7.15 software to analyze the PBPTP data in a heat map, and the results show that the overall trust of each PBPTP project is not high (See Figure 2 for details).

Physicians strongly believe that the patient's perceived behavior is "the patient is willing to inform you of important changes in his/her condition in a timely manner", accounting for only 54.1%. The perception behavior with the highest complete trust rate is "the patient will provide all the medical information you need", accounting for only 21.6%. On the contrary, the perceived behavior with the highest total distrust rate is "patients will not be driven to seek treatment by improper benefits", as high as 44.7%. Nearly half of the participants questioned the patient's integrity issues, which became the most unwilling physician or the least trusted option. Combined with PBPTP heat map, Hypothesis 1 is confirmed.

Next, we will use the chi-square test of  $R \times C$  cross-table data to explore the PBPTP distribution rates of different demographic characteristics. Its statistics is:

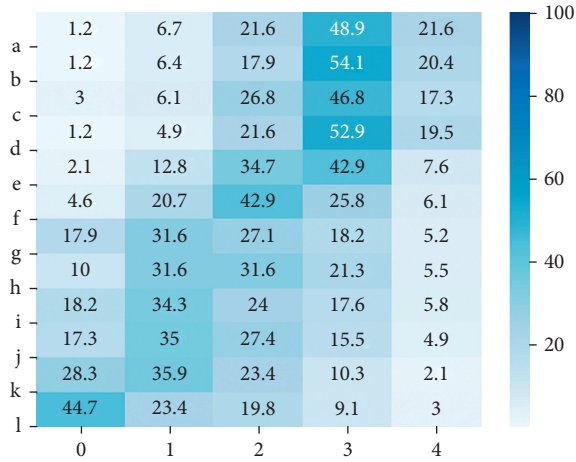


FIGURE 2: Heat map of the distribution of physician's trust rate in patients' perception behavior (%). Note. (a) the patient will provide all the medical information you need; (b) the patient is willing to inform you of important changes in his/her condition in a timely manner; (c) the patient will tell you all the drugs he/she is taking or the treatment he/she is receiving; (d) the patient can understand what you say to him/her; (e) the patient will follow the treatment plan you suggest; (f) the patient will actively participate in his/her condition management; (g) the patient will tell you when they do not follow the treatment plan; (h) the patient will have regular follow-up visits; (i) the patient will respect your time; (j) the patient will respect your personal bottom line; (k) patient will not make unreasonable demands; (l) patient will not be driven by improper interests to see a doctor; 0, no trust at all; 1, a little trust; 2, a certain trust; 3, a great trust; 4, full trust.

$$\chi^2 = \sum_{i=1}^R \sum_{j=1}^C \frac{(A_{ij} - T_{ij})^2}{T_{ij}}, \quad (1)$$

where  $A_{ij}$  is the actual frequency of each cell,  $T_{ij}$  is the theoretical frequency of each cell, and its parametric freedom is calculated as  $(R-1)(C-1)$ ,  $R$  is the number of rows, and  $C$  is the number of columns.

There are differences in perceived behaviors of PTP with different demographic characteristics.

We performed the chi-squared test on PBPTP distribution rates of different genders/ages. "Patient will actively participate in his/her disease management" and "patients will visit regularly" male physicians have a higher PTP rate than female physicians (all  $P < 0.018$ ), and there is no significant difference in other items. The age difference of PBPTP distribution rate was statistically significant (all  $P < 0.017$ ), which partially verified Hypothesis 2. Detailed data are shown in Table 1.

We performed a chi-squared test on the PBPTP distribution rate for different levels of education/title. In terms of perceived behaviors of "patients will tell you when they do not follow the treatment plan", the PTP rate of postgraduates and above is higher than undergraduates and below ( $P = 0.017$ ), and there is no significant difference in other programs. In addition to the items "patient will actively participate in his/her condition management", "patient will tell you when he does not follow the treatment plan",

"patient will return regularly," and "patient will not make unreasonable requests" with the exception of "patients are not driven by improper benefits," the remaining perceived behaviors are statistically significant for different job title levels ( $P_s \leq 0.001$ ), partially verifying Hypothesis 2, and the detailed data are shown in Table 2.

We conducted the chi-squared test on the trust rate of patients' perceived behaviors among physicians with different annual income/department. In the project "patients will provide you with all of the medical information you need", "patient inform you of the important changes about his/her illness timely", "patient will tell you the kinds of drugs he/she is taking or the treatment he/she is accepting", "the patient can understand what you said to him/her," and "patients will follow the treatment plans proposed by you", the perceived behaviors and PBPTP distribution income are different ( $P \leq 0.001$ ). The difference in other items is not obvious, which partially confirms the conjecture of Hypothesis 2; in terms of "the patient can understand what you say to him/her," the PBPTP distribution rate was different among departments ( $P \leq 0.001$ ). The conjecture of Hypothesis 3 is partially confirmed, and the detailed data are shown in Table 3.

## 4. Discussion

This study found that the overall level of PTP was not high, and most PBPTPs were skeptical. There are three possible reasons for this. Firstly, it will take some time for the medicine to take effect, and in such a semistructured medical market environment [27], it is very common for patients to change therapists because they cannot adhere to the medication, which aggravates the suspicion between doctors and patients. Secondly, it may be related to the current domestic doctor-patient tension, frequent doctor-patient disputes, frequent occurrence of medical reparations, and severe impact on PBPTP caused by medical trouble and medical violence, which is consistent with the literature [19, 21]. Furthermore, it may be related to domestic medical habits. For example, in the face of major medical activities such as the diagnosis and treatment plan and operation, doctor-patient communication and medical respect are mostly process-based and off-site behaviors, which have adversely affected the patients/families. As a result, the patients'/family's bad emotions will be directly fed back to the physician, which will affect PBPTP cognition, so the overall low PBPTP level can be understood. Based on such issues, we propose strengthening the top-level design and official intervention, reducing the imbalance of doctor-patient information, and improving doctor-patient cognition.

The results of this study showed that the distribution rate of PBPTP was correlated with factors such as gender, age, education level, professional title, and department of physicians, which was consistent with the results of the literature [14, 25]. The reason for the association with gender may be the gender differences in personality traits and the consequences of gender differences in China over the last century. One of the reasons associated with age may be that the formation of individual thinking and values is influenced by

TABLE 1: Comparison of gender/age differences between physicians and patients' perceived behavior trust (cases (%)).

Projects	PBPTP <i>n</i> (%)		$\chi^2$	<i>P</i>	PBPTP <i>n</i> (%)				$\chi^2$	<i>P</i>
	Male	Female			21~30	31~40	41~50	>51		
a	114 (70.4)	118 (70.6)	0.003	0.997	36 (40.9)	72 (66.7)	82 (96.5)	42 (87.5)	72.073	0.000
b	119 (73.3)	126 (75.5)	0.172	0.706	39 (44.3)	82 (75.9)	82 (96.5)	42 (87.5)	68.124	0.000
c	109 (67.3)	102 (61.1)	1.377	0.252	40 (45.5)	58 (53.7)	77 (90.6)	36 (75.0)	46.781	0.000
d	117 (72.2)	121 (72.5)	0.002	0.962	43 (48.9)	71 (65.7)	79 (92.9)	45 (93.8)	55.615	0.000
e	83 (51.2)	83 (49.7)	0.077	0.781	28 (31.8)	28 (25.9)	73 (85.9)	37 (77.1)	94.514	0.000
f	64 (39.5)	41 (24.6)	8.464	0.004	25 (28.4)	25 (23.1)	41 (48.2)	14 (29.2)	14.904	0.002
g	45 (27.8)	32 (19.2)	3.405	0.069	14 (15.9)	18 (16.7)	32 (37.6)	13 (27.1)	15.474	0.001
h	53 (32.7)	35 (31.0)	5.802	0.018	15 (17.0)	22 (20.4)	34 (40.0)	17 (35.4)	15.930	0.001
i	45 (27.8)	32 (19.2)	3.405	0.069	12 (13.6)	10 (9.3)	34 (40.0)	21 (43.8)	40.881	0.000
j	40 (24.7)	27 (16.2)	3.684	0.057	12 (13.6)	8 (7.4)	27 (31.8)	20 (41.7)	33.879	0.000
k	26 (16.1)	15 (9.0)	3.765	0.066	8 (9.1)	7 (6.5)	15 (17.6)	11 (22.9)	11.362	0.010
l	24 (14.8)	16 (9.6)	2.109	0.177	9 (10.2)	6 (5.6)	16 (18.8)	9 (18.8)	10.205	0.017

Note. The connotation of the project serial number is the same as that in Figure 3. Significance level  $P < 0.05$  (2-tailed). Referring to the literature [4], doctors with "great trust" and "complete trust" are denoted as patients trusted by doctors.

TABLE 2: Comparison of differences in trust in patients' perceived behavior among physicians with different education levels/professional titles (cases (%)).

Projects	CPTP <i>n</i> (%)		$\chi^2$	<i>P</i>	CPTP <i>n</i> (%)				$\chi^2$	<i>P</i>
	UGB	PGA			Primary	Intermediate	Subsenior	Senior		
a	72 (72.0)	160 (69.9)	0.152	0.793	42 (44.2)	52 (64.2)	97 (92.4)	41 (85.5)	62.539	0.000
b	78 (78.0)	167 (73.0)	0.943	0.410	46 (48.4)	58 (71.6)	100 (95.2)	41 (85.5)	61.098	0.000
c	63 (63.0)	148 (64.7)	0.080	0.803	47 (49.5)	42 (51.9)	82 (78.1)	40 (83.3)	30.778	0.000
d	76 (76.0)	162 (70.7)	0.962	0.351	45 (47.4)	58 (71.6)	90 (85.7)	45 (93.7)	50.011	0.000
e	55 (55.0)	111 (48.5)	1.187	0.284	30 (31.6)	34 (41.9)	64 (60.9)	38 (79.2)	36.328	0.000
f	27 (27.0)	78 (34.1)	1.597	0.247	29 (30.6)	23 (28.4)	37 (35.2)	16 (33.3)	1.124	0.771
g	15 (15.0)	62 (27.1)	5.661	0.017	17 (17.9)	17 (21.0)	29 (27.6)	14 (29.1)	3.802	0.284
h	25 (25.0)	63 (27.5)	0.224	0.686	17 (17.9)	22 (27.2)	34 (32.4)	15 (31.2)	6.004	0.111
i	19 (19.0)	58 (25.3)	1.555	0.258	10 (10.5)	14 (17.3)	35 (33.4)	18 (37.5)	21.576	0.000
j	16 (16.0)	51 (22.3)	1.688	0.234	11 (11.6)	10 (12.4)	28 (26.7)	18 (37.5)	18.995	0.000
k	13 (13.0)	28 (12.2)	0.038	0.857	7 (7.4)	7 (8.7)	18 (17.1)	9 (18.8)	7.191	0.066
l	11 (11.0)	29 (12.7)	0.180	0.718	9 (9.5)	8 (9.9)	15 (14.3)	8 (16.7)	2.394	0.495

Note. The connotation of the item serial number is the same as that in Figure 3; the significance level is  $P < 0.05$  (2-tailed). With reference to the literature [17], doctors with "great trust" and "full trust" are recorded as doctors trusting patients; UGB: undergraduates and below; PGA: postgraduates and above.

TABLE 3: Comparison of differences in patients' perceived behavior trust among physicians of different annual incomes/departments (cases (%)).

Projects	CPTP <i>n</i> (%)			$\chi^2$	<i>P</i>	CPTP <i>n</i> (%)					$\chi^2$	<i>P</i>
	< 70000	70000~120000	> 120000			Internists	Surgeon	Gynecologist	Pediatrics	Others		
a	45 (50.0)	63 (60.6)	124 (91.8)	52.721	0.000	55 (76.3)	47 (65.3)	51 (82.3)	43 (67.2)	36 (61.0)	9.158	0.057
b	51 (56.6)	68 (65.4)	126 (93.4)	44.783	0.000	59 (81.9)	53 (73.6)	47 (75.8)	50 (78.1)	36 (61.0)	8.268	0.082
c	49 (54.4)	53 (51.0)	109 (80.7)	27.704	0.000	54 (75.0)	39 (54.2)	40 (64.5)	43 (67.2)	35 (59.3)	7.663	0.105
d	45 (50.0)	75 (72.1)	118 (87.4)	37.768	0.000	45 (62.5)	40 (55.5)	51 (82.3)	55 (85.9)	47 (79.6)	24.164	0.000
e	28 (31.1)	49 (47.2)	89 (65.9)	26.863	0.000	32 (44.5)	38 (52.8)	33 (53.2)	37 (57.8)	26 (44.1)	2.891	0.576
f	31 (34.4)	31 (29.8)	43 (31.9)	0.478	0.787	18 (25.0)	29 (40.3)	15 (24.2)	25 (39.1)	18 (30.5)	7.161	0.128
g	16 (17.8)	26 (25.0)	35 (25.9)	2.216	0.330	16 (22.2)	22 (30.5)	10 (16.1)	14 (21.9)	15 (25.4)	4.158	0.385
h	20 (22.2)	30 (28.8)	38 (28.1)	1.310	0.520	17 (23.6)	22 (30.5)	18 (29.1)	18 (28.1)	13 (22.0)	1.791	0.774
i	13 (14.4)	29 (27.9)	35 (25.9)	5.674	0.059	16 (22.2)	15 (20.8)	15 (24.2)	14 (21.9)	17 (28.8)	1.390	0.846
j	12 (13.3)	20 (19.2)	35 (25.9)	5.401	0.067	13 (18.1)	14 (19.4)	12 (19.4)	8 (12.5)	20 (33.9)	9.418	0.051
k	8 (8.9)	13 (12.5)	20 (14.9)	1.738	0.419	5 (7.0)	13 (18.1)	8 (12.9)	9 (14.0)	6 (10.2)	4.520	0.340
l	12 (13.3)	10 (9.7)	18 (13.3)	0.921	0.631	7 (9.7)	9 (12.5)	5 (8.1)	11 (17.2)	8 (13.6)	3.005	0.557

Note. The connotation of the project serial number is the same as that in Figure 3. Significance level  $P < 0.05$  (2-tailed). Referring to [17], doctors with "great trust" and "complete trust" are denoted as patients trusted by doctors. Others: other clinical departments.

age, and the other may come from the formation of individual cognition and the differences in the influence of social environment and doctor-patient risk events. The possible

reason for the influence of education level and professional title is the management system and promotion mechanism of domestic health service. Differences among departments



may result from different incidence rates of doctor-patient risk events in departments, which is consistent with the literature [11]. Based on such problems, we propose to actively respond to personalized needs, analyze personalized differences, weigh the pros and cons, improve the advantages of policy management, eliminate the disadvantages of the old pattern, improve the medical effect, and reduce the doctor-patient risk from the root.

The results of this study found that the trust rate of doctors in patients' perceived behaviors fluctuated around 20% for five items, such as "not following the advice of treatment plan", "patients will respect your time", "patients will respect your personal bottom line", "patients will not make unreasonable demands," and "patients will not be driven by illegitimate interests". In particular, "patients will not make unreasonable demands" and "patients will not be driven by illegitimate interests" became the most reluctant problems for doctors to face or the least trusted option. There are several possible reasons for this. Firstly, unfair media coverage of doctor-patient disputes has seriously affected doctors' personal cognition. Secondly, hospitals and even the officials regard the affected party as vulnerable groups; nonprinciple compromises and financial compensation have made the doctor be swayed by considerations of gain and loss of consciousness. Thirdly, the absence of relevant regulations and laws leads to the lack of rules and regulations of patients/family members in doctor-patient disputes, which seriously affects the basic image of patients, thus giving PBPTP a bias. Based on such issues, we propose to enhance the effect of third-party supervision, increase social participation, and face up to the risks of doctor-patient supply and demand.

## 5. Conclusion

Through the above analysis and research, in order to improve RBDP, enhance doctor-patient trust rate, and promote RBDP harmony, doctor-patient risk management models of "Official-Individual-Social" Triple Action should be formed, as shown in Figure 3.

Official actions, first of all, further standardize the development of medical and health services, strengthen the reform of medical and health care, implement scientific pricing of pharmaceutical products, simplify the distribution process, save health resources, and form a good situation of "the people can afford to be ill and is optimistic about the disease". Secondly, improve the training mechanism for medical workers, optimize the mechanism for staff promotion, and explore a "green" salary system that links doctors' cure rates, patient satisfaction rates, and bonuses; finally, improve the legal system for doctor-patient disputes, regulate doctor-patient rights and responsibilities, and establish dispute adjustment mechanism.

Individual response clarified the consequences of medical atrocity and clarified the significance of doctor-patient mutual trust. First of all, to explore the relationship between "doctor-patient trust" and "social trust" by exploring the validity of doctor-patient trust

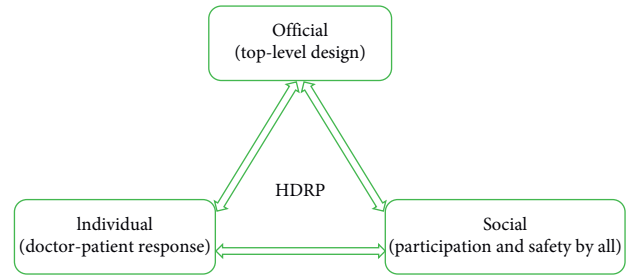


FIGURE 3: Model of "official-individual-social" triple action. *Note.* HDRP: harmonious RBDP.

and behavioral intervention and to open up a theoretical framework to improve individual trust rate from an academic perspective is an attempt to solve the crisis of social trust and doctor-patient trust; secondly, carry out the publicity of laws and regulations on doctor-patient disputes, so that patients/family members can clearly understand the serious consequences of medical violence, smooth the channels of doctor-patient dispute mediation, and release the extreme emotions of patients/family members; Finally, the concept of medical personnel should be changed, the professional spirit of medical personnel should be improved, and the passive should be changed into the active one, so as to form a medical team that is "dare to treat and able to treat" diseases.

The social supervision medical process gradually realizes "full openness". Firstly, promote the popularization of medical science to the general public, try to make the whole people agree with the correct medical treatment concepts such as "treatment is not equal to cure" and "doctors do not cure all diseases," so as to guide patients/families to have reasonable expectations on the treatment effect and reduce conflicts. Secondly, social media should be impartial, scientific, and reasonable in shouldering corresponding social responsibilities, be impartial, rational, and tolerant in reporting medical disputes, and have clear rights and responsibilities, while analyzing doctor-patient contradictions fairly, so as to make the due contribution to construct the good health care environment.

To summarize, demographic characteristics variable may be one of the factors affecting PBPTP; PBPTP is associated with doctor-patient risk. It makes sense for us to propose a new model of physician-patient risk management from the perspective of PTP about "official-individual-social" triple action.

This study expands the research on the relationship between PBPTP and demographic characteristics, provides a new research perspective for RBDP improvement and doctor-patient risk management in hospitals, and expands the application of planned behavior theory in doctor-patient risk management in hospitals, which makes certain theoretical contributions. Its deficiency is that it has not conducted quantitative management of doctor-patient risk management based on differences in PBPTP level, which is also an in-depth study that we will carry out in the future.

## Data Availability

All datasets generated for this study are included in the article and/or the supplementary files.

## Ethical Approval

Ethical approval for the study was received from the Ethics Committee of Anhui Medical University, and this study meets the ethical requirements.

## Disclosure

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Jiang Jie Sun was involved in the study design, data-entry, and original writing. Zhi Bo Zheng was involved in the study design. Xue Li Jiang was responsible for the data entry and language editing. Wei Wei Hu was responsible for data entry. Nan Zhen Ma and Xiao-Zhe Yan collected the data. Jun Liu was responsible for language editing. Meng Ying Li was involved in the data analysis and language editing. Cheng sen He financially supported the study. Li ping Zhang was involved in the language editing, data collection, and financial support.

## Acknowledgments

We would like to thank Matthias Alfons Theis for collecting part of the data. This work was supported in part by the Natural Science Foundation of Anhui Province of China under Grant no. 1908085MG233 to JS, Quality Engineering for research projects of the Anhui Department of Education about Wisdom classroom 2018zhkt180 to JS, the Natural Science Foundation for the Higher Education Institutions of Anhui Province of China under Grant no. KJ2019A0945 to LZ, and Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province under Grant nos. KJ2016A372 and KJ2016SD30 to LZ and CH. The work was supported in part by the Joint Special Foundation on basic research in Local Colleges and Universities for the Department of Science and Technology of Yunnan Province of China under Grant no. 2017FH001-106 to ZZ.

## References


- [1] C. Zhou and X. Xu, "Review on the studies of doctor-patient trust in both China and overseas countries," *Journal of Kunming University of Science and Technology(Social Science Edition)*, vol. 15, no. 1, pp. 8–14, 2015.
- [2] J. Liu, C. Zhou, and F. Yang, "A review on domestic and oversea literature about doctor-patient trust scale research," *Journal of Kunming University of Science and Technology (Social Science Edition)*, vol. 15, no. 3, pp. 11–18, 2015.
- [3] A. A. Sewell, "Disaggregating ethnoracial disparities in physician trust," *Social Science Research*, vol. 54, no. 11, pp. 1–20, 2015.
- [4] W. Augustyniak, "The role of trust in doctor-patient relationship: qualitative evaluation of online feedback from polish patients," *Economics & Sociology*, vol. 9, no. 3, pp. 76–88, 2015.
- [5] D. H. Thom, M. A. Hall, and L. G. Pawlson, "Measuring patients' trust in physicians when assessing quality of care," *Health Affairs*, vol. 23, no. 4, pp. 124–132, 2004.
- [6] D. Haslam, "Trust is the fulcrum of the doctor-patient relationship," *The Practitioner*, vol. 259, no. 1783, p. 35, 2015.
- [7] E. G. Howe, "How to retain the trust of patients and families when we will not provide the treatment they want," *The Journal of Clinical Ethics*, vol. 26, no. 2, pp. 89–99, 2015.
- [8] H. Pin, H. Jin, J. Eshin et al., "Socially transmitted placebo effects," *Nature Human Behavior*, vol. 3, pp. 1295–1305, 2019.
- [9] K. A. Wallston, B. S. Wallston, and S. Gore, "Development of a scale to measure nurses' trust of patients: a preliminary report," *Nursing Research*, vol. 22, no. 3, pp. 232–235, 1973.
- [10] L. A. Anderson and R. F. Dedrick, "Development of the trust in physician scale a measure to assess interpersonal trust patient-physician relationships," *Psychological Reports*, vol. 67, no. 3, pp. 1091–1100, 1991.
- [11] V. Gopichandran and S. K. Chetlapalli, "Trust in the physician-patient relationship in developing healthcare settings: a quantitative exploration," *Indian Journal of Medical Ethics*, vol. 12, no. 3, pp. 141–148, 2015.
- [12] D. J. Burgess, "Addressing racial healthcare disparities: how can we shift the focus from patients to providers?" *Journal of General Internal Medicine*, vol. 26, no. 8, pp. 828–830, 2011.
- [13] E. M. A. Smets, M. A. Hillen, K. F. L. Douma, L. J. A. Stalpers, C. C. E. Koning, and H. C. J. M. de Haes, "Does being informed and feeling informed affect patients' trust in their radiation oncologist?" *Patient Education and Counseling*, vol. 90, no. 3, pp. 330–337, 2013.
- [14] S. Ozawa and P. Sripad, "How do you measure trust in the health system? a systematic review of the literature," *Social Science & Medicine*, vol. 91, no. 1, pp. 10–14, 2013.
- [15] Z. Dong and C. Chen, "A preliminary study of validity and reliability of the Chinese version of the Physician Trust in the Patient Scale," *Chinese Mental Health Journal*, vol. 30, no. 7, pp. 481–485, 2016.
- [16] A. Banerjee and D. Sanyal, "Dynamics of doctor-patient relationship: a cross-sectional study on concordance, trust, and patient enablement," *Journal of Family and Community Medicine*, vol. 19, no. 1, pp. 12–19, 2012.
- [17] I. Ajzen and M. Fishbein, *Understanding Attitudes and Predicting Social Behavior*, Prentice-Hall, Upper Saddle River, NJ, USA, 1980.
- [18] D. H. Thom, S. T. Wong, D. Guzman et al., "Physician trust in the patient: development and validation of a new measure," *The Annals of Family Medicine*, vol. 9, no. 2, pp. 148–154, 2011.
- [19] C. He, "Inspiration of the evolution of doctor-patient relationship to today's medical care reform," *Jianghuai Tribune*, no. 2, pp. 117–121, 2015.
- [20] J. Sun, L. Zhang, C. He et al., "Interpretation of the Behavioral economics and countermeasures of the Medical risk," *Journal of Nanjing Medical University(Social Sciences)*, vol. 73, no. 2, pp. 141–145, 2016.
- [21] J. J. Sun, L. P. Zhang, P. K. Mu et al., "Differences in the cognitive evaluation of the doctor-patient Relationship



- between the medical side and the contracting parties,” *Chinese Mental Health Journal*, vol. 30, no. 7, pp. 486–491, 2016.
- [22] J. Sun, L. Zhang, and C. He, “The related factors of physician trust in patients,” *Chinese Mental Health Journal*, vol. 32, no. 5, pp. 407–409, 2018.
- [23] X. Zhang and M. Sleeboom-Faulkner, “Tensions between medical professionals and patients in mainland China,” *Cambridge Quarterly of Healthcare Ethics*, vol. 20, no. 3, pp. 458–465, 2011.
- [24] J. Sun, R. Sun, Y. Jiang et al., “The relationship between psychological health and social support: evidence from physicians in China,” *PLoS One*, vol. 15, no. 1, Article ID e0228152, 2020.
- [25] J. Sun, L. Zhang, R. Sun et al., “Exploring the influence of resiliency on physician trust in patients: an empirical study of Chinese incidents,” *PLoS One*, vol. 13, no. 12, Article ID e0207394, 2018.
- [26] J. Sun, P. Wang, Y. Du, and J. Liu, “Analysing the influence factors of single task pricing based on public packet system: an Empirical Study in China,” *Journal of Physics: Conference Series*, vol. 1437, pp. 1–6, 2020.
- [27] Q. Yang and J. Pan, “Control under times of uncertainty: the relationship between hospital competition and physician-patient disputes,” *International Journal for Equity in Health*, vol. 16, no. 1, pp. 205–217, 2017.

## Research Article

# Multimodal Semisupervised Deep Graph Learning for Automatic Precipitation Nowcasting

Kaichao Miao,<sup>1</sup> Wei Wang,<sup>2</sup> Rui Hu,<sup>3</sup> Lei Zhang,<sup>1</sup> Yali Zhang,<sup>1</sup> Xiang Wang,<sup>1</sup> and Fudong Nian <sup>4</sup>

<sup>1</sup>Anhui Public Meteorological Service Center, Anhui Meteorological Bureau, Hefei, Anhui, China

<sup>2</sup>Center for Assessment and Demonstration Research, Academy of Military Science, Beijing, China

<sup>3</sup>School of Electrical Engineering and Automation, Anhui University, Hefei, Anhui, China

<sup>4</sup>School of Advanced Manufacturing Engineering, Hefei University, Hefei, Anhui, China

Correspondence should be addressed to Fudong Nian; nianfd@hfuu.edu.cn

Received 21 July 2020; Accepted 21 September 2020; Published 1 October 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Kaichao Miao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Precipitation nowcasting plays a key role in land security and emergency management of natural calamities. A majority of existing deep learning-based techniques realize precipitation nowcasting by learning a deep nonlinear function from a single information source, e.g., weather radar. In this study, we propose a novel multimodal semisupervised deep graph learning framework for precipitation nowcasting. Unlike existing studies, different modalities of observation data (including both meteorological and nonmeteorological data) are modeled jointly, thereby benefiting each other. All information is converted into image structures, next, precipitation nowcasting is deemed as a computer vision task to be optimized. To handle areas with unavailable precipitation, we convert all observation information into a graph structure and introduce a semisupervised graph convolutional network with a sequence connect architecture to learn the features of all local areas. With the learned features, precipitation is predicted through a multilayer fully connected regression network. Experiments on real datasets confirm the effectiveness of the proposed method.

## 1. Introduction

The goal of precipitation nowcasting is to predict the future rainfall intensity in a local region over a relatively short period. This method plays an important role in land security and emergency management of natural calamities. Though it is very crucial, precipitation nowcasting still depends on the real-time manual analysis of meteorological observation data by forecasters.

To address this issue, researchers are increasingly trying to replace forecasters with computers, e.g., [1]. In essence, precipitation nowcasting is a spatiotemporal sequence forecasting problem with a sequence of past meteorological observation data as input and a sequence of fixed numbers (usually larger than 1) of future precipitation as output. However, traditional optical flow-based methods have unsatisfactory performance.

Precipitation nowcasting has three long-standing problems. The first challenge is spatiotemporal correlations. Precipitation at any location is influenced by the weather condition of its neighboring region over a relatively short period. The second challenge is diversified observation data. With the progress of technology, we can obtain various meteorological or nonmeteorological observation data of a certain region, e.g., radar echo maps, satellite images, topographic map, temperature, and air humidity. However, handling these data in a unified framework is an open problem. The third challenge is semisupervision. It is generally known that the distribution of meteorological stations for observing precipitation is uneven, implying that precipitation at many locations is unknown.

Since the recent advances in deep neural networks [2, 3], their high capability in various classification and regression tasks has been successfully demonstrated, e.g., image

classification by [4], object detection by [5], video representation by [6], and speech recognition by [7]. There are also some studies on precipitation nowcasting using deep learning techniques. Xingjian et al. [8] and Shi et al. [9] first introduced a convolutional long short-term memory (ConvLSTM) network to capture spatiotemporal correlations, which has been shown to outperform traditional optical flow-based methods for precipitation nowcasting, indicating that deep learning models have a huge potential for solving this problem. Subsequently, Singh et al. [10] and Karevan and Suykens [11] introduced a more complicated LSTM structure, and Shi et al. [12] used convolutional neural networks (CNNs) to extract more efficient representations.

Although the methods above can achieve some progress, their inputs are merely weather radar echo maps, and they only focus on addressing the problem of spatiotemporal correlations, neglecting the latter two problems.

Inspired by the success of Deep Learning Vision for Nonvision Tasks by [13], we propose a novel multimodal semisupervised deep graph learning framework for precipitation nowcasting in this study. In contrast to previous studies, different modalities of observation data (including both meteorological and nonmeteorological data) are modeled jointly, and thus, benefiting each other. All information is converted to an image structure. Then, precipitation nowcasting is deemed as a computer vision task to be optimized. We convert all observation information into a graph structure and introduce a semisupervised graph convolutional network (GCN) with a sequence connect architecture to learn the features of all local areas for handling areas without available precipitation. With the learned features, precipitation is predicted through a multilayer fully connected regression network.

We summarize the contributions of this work as follows:

- (1) We introduce a novel short-term precipitation nowcasting solution by leveraging multimodal observation data (including meteorological and non-meteorological data)
- (2) A unified multimodal semisupervised deep graph learning is proposed, aiming at simultaneously addressing three long-standing problems in precipitation nowcasting, i.e., spatiotemporal correlations, diversified observation data, and semisupervision
- (3) To the best of our knowledge, this is the first attempt to address the precipitation nowcasting problem using a GCN
- (4) The experimental results on the collected large-scale real dataset validate the effectiveness of the proposed solution

The rest of the paper is organized as follows. In Section 2, related work is reviewed. Section 3 presents the problem formulation and introduces the proposed semisupervised deep regression framework in detail. In Section 4, we report and analyze the experimental results. Finally, we conclude the paper and discuss future work in Section 5.

## 2. Related Work

In this section, we briefly review the closely related methods in two folds: deep learning-based precipitation nowcasting methods and graph-based semisupervised learning techniques.

*2.1. Deep Learning-Based Precipitation Nowcasting.* Xingjian et al. [8] formulated precipitation nowcasting as a spatiotemporal sequence forecasting problem and proposed a ConvLSTM model, which extends the LSTM by [14], by adding convolutional structures into both input-to-state and state-to-state transitions to solve the problem. Using radar echo sequences for model training, the authors showed that ConvLSTM was better at capturing spatiotemporal correlations than fully connected LSTM and provided more accurate predictions than the real-time optical flow via variational methods for echoes of radar algorithm by [15]. However, the convolutional recurrence structure in ConvLSTM-based models is location invariant, whereas natural motion and transformation (e.g., rotation) are location variant in general. To address this, Shi et al. [9] proposed a trajectory-gated recurrent unit model that can actively learn a location-variant structure for recurrent connections. To obtain a more robust representation, Karevan and Suykens [11] proposed a two-layer spatiotemporal stacked LSTM model. Unlike LSTM-based methods which were used widely on sequence learning and time series prediction, Shi et al. [12] used recurrent dynamic CNNs to handle the spatiotemporal information of radar data.

Unfortunately, all techniques mentioned above only consider radar echo maps as the input data. In addition to radar echo maps, other types of meteorological observed data, such as satellite images, temperature, and air humidity, and nonmeteorological observation data, such as “topographic maps” are available, and these data were incorporated in this study. Notably, data from different sources are complementary. In this work, the proposed model is capable of not only handling spatiotemporal information from different observation sources simultaneously but also unifying labeled and unlabeled data.

*2.2. Graph-Based Semisupervised Learning.* Existing semisupervised learning methods using graph representations can be roughly be divided into two categories: methods that use some form of explicit graph Laplacian regularization and graph embedding-based approaches.

Typical graph Laplacian regularization techniques include label propagation by [16], manifold regularization by [17], and deep semisupervised embedding by [18]. Typical graph embedding-based methods include DeepWalk by [19], LINE by [20], and node2vec by [21]. Perozzi et al. [19] learned embedding via the prediction of the local neighborhood of nodes, sampled from random walks on a graph. Tang et al. and Grover and Leskovec [20, 21] extend DeepWalk with more sophisticated random walk or breadth-first search schemes.

Although the methods mentioned above have achieved significant progress, a multistep pipeline is included in these frameworks, which means that they cannot be trained in an end-to-end way. Recently, Kipf and Welling [22] proposed a more simplified GCN for semisupervised learning by employing a first-order approximation of spectral filters. The GCN and subsequent variants have achieved state-of-the-art results in various application areas, including multilabel classification by [23], zero-shot learning by [24], social networks by [25], and natural language processing by [26].

In this paper, graph-based semisupervised learning is utilized for precipitation nowcasting. There are two main differences between the proposed method and that by [22] (including its variants). First, as observation data are obtained from different sources, graphs in our framework are multimodal. In contrast, most graphs in conventional graph-based semisupervised learning methods are built by a single modality (image, text, etc.). Second, precipitation nowcasting is a regression problem in our study, whereas previous techniques focus on classification tasks.

### 3. Method

In this section, we first define problems of precipitation nowcasting processed in this paper. Then, we introduce the proposed multimodal semisupervised deep graph learning in details.

**3.1. Problem Definition.** Our goal is forecasting precipitation at a place for the next several hours based on the weather conditions of its neighbors over the past several hours. Precipitation is related to many factors. As illustrated in Figure 1, according to the existing observation conditions, herein, we assume that future precipitation can be predicted via modeling the following factors: radar echo maps, air humidity images, satellite images, temperature images, and available precipitation data over the past several hours with the corresponding topographic map. In addition, as precipitation prediction of every longitude and latitude is impossible, we assume that if we can divide a certain region to many cells (that is, each cell represents a local area) then precipitation at different locations in the same cell can be deemed as almost the same.

As the factors belong to different modalities, precipitation nowcasting can be deemed as a multimodal analysis task. Moreover, from Figure 1, we can observe that precipitation in some cells may be unknown when the target area is divided into many grid cells because precipitation stations are not placed everywhere. Thus, precipitation nowcasting is also a semisupervised learning problem in this paper. Therefore, we present a multimodal semisupervised learning method to solve the precipitation nowcasting problems herein. The problem formulations are detailed below.

Let  $\mathbf{X}_t = [\mathbf{x}_{re}, \mathbf{x}_{si}, \mathbf{x}_{tp}, \mathbf{x}_{ti}, \mathbf{x}_{ah}]$  be the observation data at time  $t$ , where  $\mathbf{x}_{re} = [\mathbf{x}_{re_t}, \mathbf{x}_{re_{t-1}}, \dots, \mathbf{x}_{re_{t-m}}]$  are the radar echo maps of the target area.  $\mathbf{x}_{si} = [\mathbf{x}_{si_t}, \mathbf{x}_{si_{t-1}}, \dots, \mathbf{x}_{si_{t-m}}]$  are the satellite images of the same area.  $\mathbf{x}_{tp}$  is the topographic map

of the same area.  $\mathbf{x}_{ti} = [\mathbf{x}_{ti_t}, \mathbf{x}_{ti_{t-1}}, \dots, \mathbf{x}_{ti_{t-m}}]$  and  $\mathbf{x}_{ah} = [\mathbf{x}_{ah_t}, \mathbf{x}_{ah_{t-1}}, \dots, \mathbf{x}_{ah_{t-m}}]$  are the temperature and air humidity image sequences from a remote sensing satellite, respectively.  $m$  is the sample index before time  $t$ . If the target area is divided into  $N \times N$  grid cells, the precipitation of all cells is denoted as  $\mathbf{P} = [\mathbf{P}_t, \mathbf{P}_{t-1}, \dots, \mathbf{P}_{t-m}]$ , where  $\mathbf{p}_u = [p_{u_1}, p_{u_2}, \dots, p_{u_{NN}}]$  ( $0 \leq u \leq m$ ). It should be noted that  $p_{u_k}$  will be unknown if no precipitation stations are placed in the corresponding cell. If a cell has several precipitation stations, the cell precipitation values will be the average observation values from the stations. In addition, the precipitation of  $n$  time points in the future is denoted as  $\mathbf{Y} = [\mathbf{y}_{t+1}, \mathbf{y}_{t+2}, \dots, \mathbf{y}_{t+n}]$ , where  $\mathbf{y}_v = [y_{v_1}, y_{v_2}, \dots, y_{v_{NN}}]$  ( $1 \leq v \leq n$ ) are the predicted precipitation in all local areas.

Therefore, the problem can be formulated as follows:

$$\mathbf{Y} = g(\mathbf{X}_t, \mathbf{P}). \quad (1)$$

To address it, we propose a multimodal semisupervised deep graph learning framework herein.

#### 3.2. Multimodal Semisupervised Deep Graph Learning

**3.2.1. Graph Representation.** Multimodal and heterogeneous data are represented by a graph structure. The target region of precipitation nowcasting is divided into many grid cells, and each cell represents a local area (as shown in Figure 2, the rectangular area represents the target region of the precipitation nowcasting and the blue points represent the precipitation stations). Let  $G(\mathbf{V}, \mathbf{E})$  be the graph structure of the observation data of the target region, where nodes  $\mathbf{V}$  represent a certain local area and edges  $\mathbf{E}$  represent the relationship of two local areas. Here, red points mean the precipitation is available while the green means the precipitation is unknown.

Therefore, the first problem is vector representation of  $G(\mathbf{V}, \mathbf{E})$ . To obtain the node and edge representations of  $G(\mathbf{V}, \mathbf{E})$ , we utilize a convolutional autoencoder to learn the feature representation of a radar echo map, an air humidity image, a satellite image, a temperature image, and a topographic map. The convolutional autoencoder (as shown in Figure 3) architecture in our method consists of five parts: an input image, an encoder, a feature representation layer, a decoder, and a reconstructed image. It should be noted that the five convolutional autoencoder architectures have the same structure but are trained separately. Moreover, to ensure that the input images are as similar as possible to the reconstructed images, we utilize mean structural similarity (MSSIM) as our objective function in the convolutional autoencoder training step. If the input and reconstructed images are marked as  $\mathbf{I}$  and  $\mathbf{I}'$ , respectively, then the MSSIM of  $\mathbf{I}$  and  $\mathbf{I}'$  is formulated as follows:

$$\text{MSSIM}(\mathbf{I}, \mathbf{I}') = \frac{1}{K} \sum_{j=1}^K \text{SSIM}(\mathbf{I}_j, \mathbf{I}'_j), \quad (2)$$

where  $K$  represents the image similarity computed in  $K$  patches and  $\mathbf{I}_j$  and  $\mathbf{I}'_j$  represent the  $j$ th image patch. SSIM is

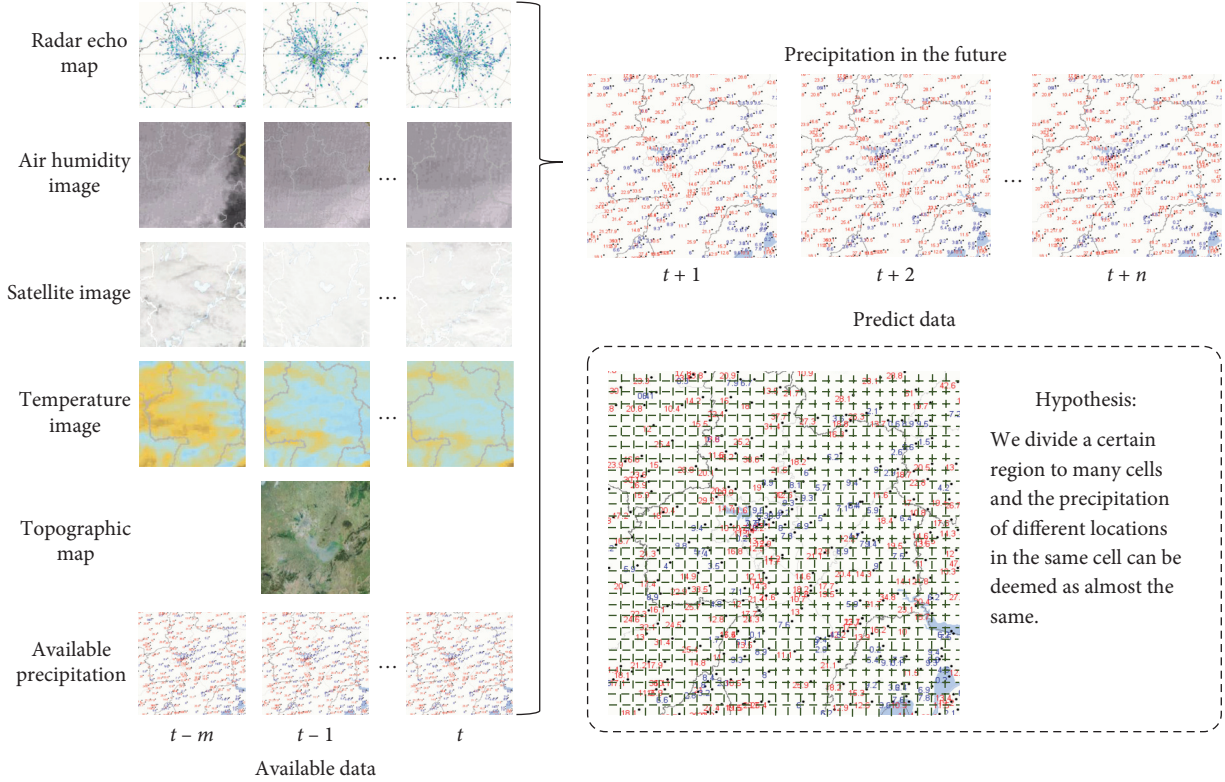


FIGURE 1: Motivation and the problem definition of precipitation nowcasting in this paper. In this paper, we assume that the future precipitation can be predicted via modeling the following factors: radar echo maps, air humidity images, satellite images, temperature images, and available precipitation data over the past several hours with the corresponding topographic map. In addition, we assume that if we can divide a certain region to many cells, then precipitation at different locations in the same cell can be deemed as almost the same.

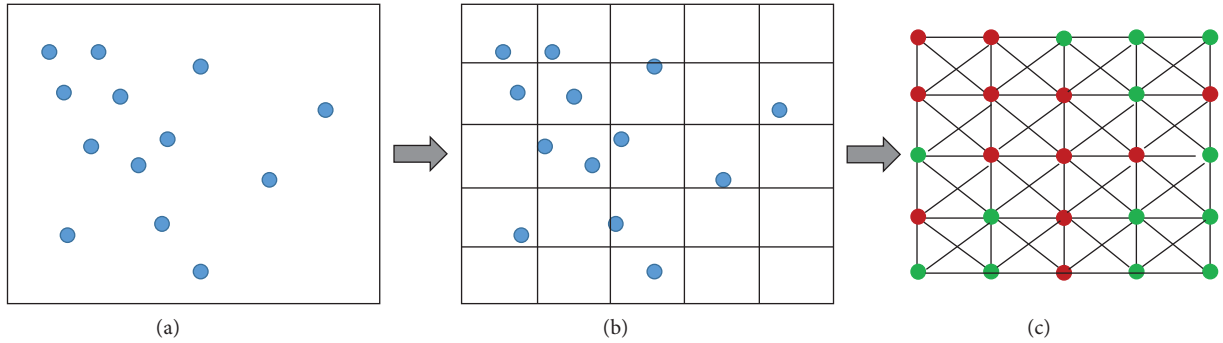


FIGURE 2: Illustration of the pipeline of graph construction. A certain area can be represented by a graph, where nodes represent a local area and the edges represent the relationship between two local areas. The red nodes mean the precipitation of the corresponding local area is available while the green means the precipitation is unknown.

the structural similarity measurement of two image patches formulated as

$$\text{SSIM}(I_j, I'_j) = \frac{2\mu_{I_j} \mu_{I'_j} + 2\sigma_{I_j, I'_j}}{(\mu_{I_j}^2 + \mu_{I'_j}^2)(\sigma_{I_j}^2 + \sigma_{I'_j}^2) + C}, \quad (3)$$

where  $\mu$  and  $\sigma$  are the mean and variance of the image patch and  $C$  is a small constant which avoids zero errors.

Once all convolutional autoencoders are trained, the node representations in graph  $G(\mathbf{V}, \mathbf{E})$  can be denoted as

$$\mathbf{f}_{v_i} = [\mathbf{f}_{\text{re}_{v_i}}, \mathbf{f}_{\text{si}_{v_i}}, \mathbf{f}_{\text{tp}_{v_i}}, \mathbf{f}_{\text{ti}_{v_i}}, \mathbf{f}_{\text{ah}_{v_i}}], \quad (4)$$

where  $\mathbf{f}_{v_i}$  is the representation of node  $v_i$ , and  $\mathbf{f}_{\text{re}_{v_i}}$ ,  $\mathbf{f}_{\text{si}_{v_i}}$ ,  $\mathbf{f}_{\text{tp}_{v_i}}$ ,  $\mathbf{f}_{\text{ti}_{v_i}}$ , and  $\mathbf{f}_{\text{ah}_{v_i}}$  are the representations obtained from the outputs of the decoders in Figure 3.

Then, the edges in graph  $G(\mathbf{V}, \mathbf{E})$  are denoted as

$$s_{i,j} = \frac{1}{\|\mathbf{f}_{v_i} - \mathbf{f}_{v_j}\|_2}, \quad (5)$$

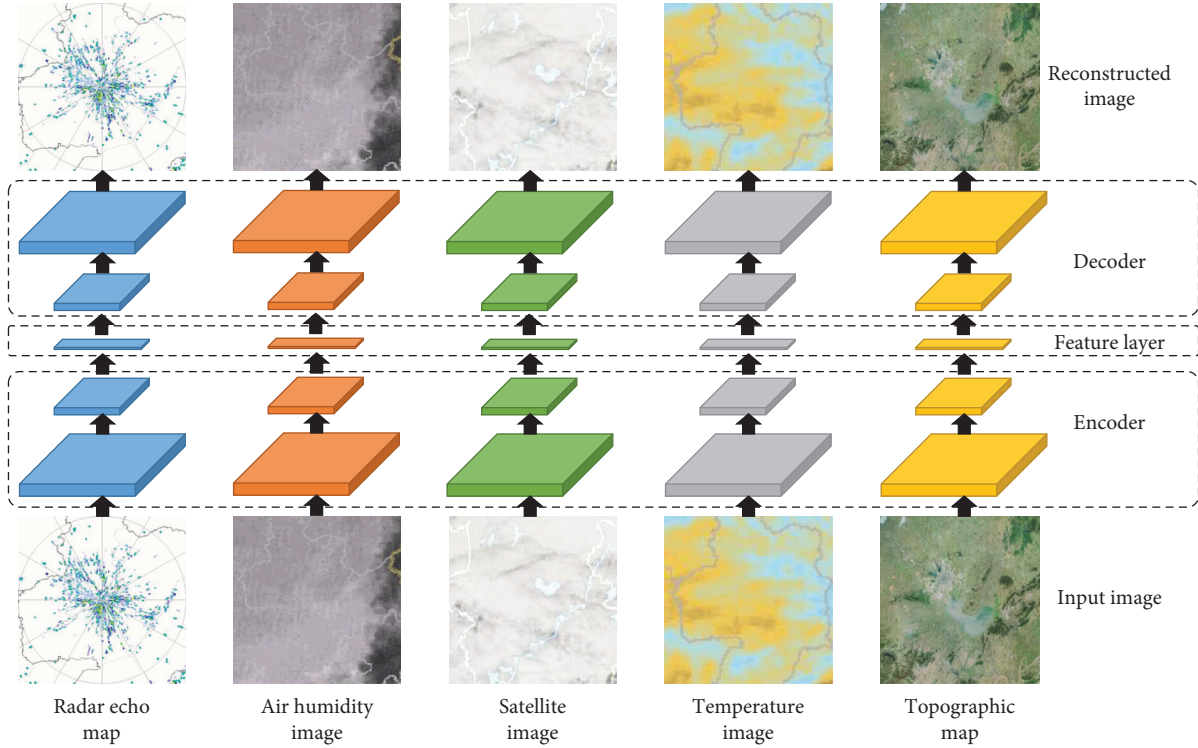


FIGURE 3: Architecture of the convolutional autoencoder used for feature representation learning in our framework.

where  $s_{i,j}$  is the similarity of two nodes  $v_i$  and  $v_j$  and  $\| * \|_2$  is the Euclidean distance of the two node representations.

Therefore, let  $G(\mathbf{H}, \mathbf{A})$  be the graph representation of graph  $G(\mathbf{V}, \mathbf{E})$ , which can be formulated as

$$\mathbf{H} = \begin{bmatrix} \mathbf{f}_{re_{v_1}} & \mathbf{f}_{si_{v_1}} & \mathbf{f}_{tp_{v_1}} & \mathbf{f}_{ti_{v_1}} & \mathbf{f}_{ah_{v_1}} \\ \mathbf{f}_{re_{v_2}} & \mathbf{f}_{si_{v_2}} & \mathbf{f}_{tp_{v_2}} & \mathbf{f}_{ti_{v_2}} & \mathbf{f}_{ah_{v_2}} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{f}_{re_{v_{N \times N}}} & \mathbf{f}_{si_{v_{N \times N}}} & \mathbf{f}_{tp_{v_{N \times N}}} & \mathbf{f}_{ti_{v_{N \times N}}} & \mathbf{f}_{ah_{v_{N \times N}}} \end{bmatrix}, \quad (6)$$

$$\mathbf{A} = \begin{bmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,N} \\ s_{2,1} & s_{2,2} & \dots & s_{2,N} \\ \dots & \dots & \dots & \dots \\ s_{N,1} & s_{N,2} & \dots & s_{N,N} \end{bmatrix}.$$

**3.2.2. Precipitation Nowcasting Model Learning.** Based on the graph representation mentioned above, the observation data of different sampled times can be deemed as a graph sequence, as shown in Figure 4. Therefore, the precipitation nowcasting task is converted to a sequence graph modeling problem, which aims to obtain the precipitation of each node according to a series of graphs.

To address it, as illustrated in Figure 5, a multimodal semisupervised deep graph learning framework is proposed as a solution. A GCN and LSTM are used to model spatial and temporal information, respectively.

GCN contains one input layer, several propagation (hidden) layers, and one final perceptron layer by [22].

Given an input  $\mathbf{H}$  and an adjacency matrix  $\mathbf{A}$ , the GCN conducts the following layerwise propagation in the hidden layer as

$$\mathbf{H}^{k+1} = \sigma(\mathbf{D}^{-(1/2)} \mathbf{A} \mathbf{D}^{-(1/2)} \mathbf{H}^k \theta^k), \quad (7)$$

where  $k = 0, 1, \dots, K - 1$ .  $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$  is a diagonal matrix with  $d_i = \sum_{j=1}^n \mathbf{A}_{ij}$ .  $\theta^k \in \mathbb{R}^{d_k \times d_{k+1}}$ ,  $d_0 = p$  is a layer-specific weight matrix needing to be trained.  $\sigma$  denotes a nonlinear activation function, and  $\mathbf{H}^{k+1} \in \mathbb{R}^{n \times d_{k+1}}$  denotes the activation output in the  $k + 1$ th layer.

As each row in  $\mathbf{H}$  represents a local area, the precipitation can be predicted as follows:

$$\mathbf{P}'_t = \text{Sigmoid}(h(\mathbf{H}'_t)), \quad (8)$$

where  $h$  is a multilayer fully connected neural network. It should be noted that equation (8) is a regression formulation, which is different from the commonly used semisupervised node classification, each row in  $\mathbf{H}$  represents a node, and GCN defines the final perceptron layer as  $\text{softmax}(\mathbf{D}^{-(1/2)} \mathbf{A} \mathbf{D}^{-(1/2)} \mathbf{H}^k \theta^k)$  to classify all nodes.

Considering that  $\mathbf{H}^0$  is crucial for computing equation (7), we design a fusion unit to integrate the five features  $(\mathbf{f}_{re_{v_i}}, \mathbf{f}_{si_{v_i}}, \mathbf{f}_{tp_{v_i}}, \mathbf{f}_{ti_{v_i}}, \mathbf{f}_{ah_{v_i}})$  into one single hidden feature representation. Here, we apply a fully connected layer using a hyperbolic tangent nonlinearity activation function. The detailed transformations are as follows:

$$\mathbf{H}^0 = \tanh\left(W_f \left[ \mathbf{f}_{re_{v_i}}, \mathbf{f}_{si_{v_i}}, \mathbf{f}_{tp_{v_i}}, \mathbf{f}_{ti_{v_i}}, \mathbf{f}_{ah_{v_i}} \right]\right), \quad (9)$$



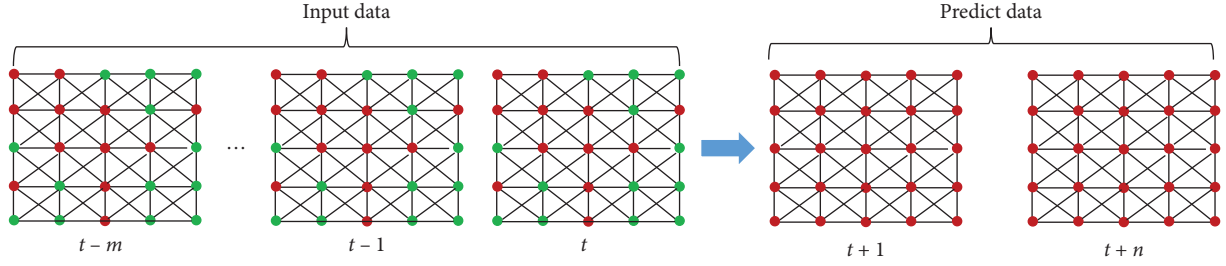


FIGURE 4: Illustration of the precipitation nowcasting which is deemed as a sequence graph modeling problem in this paper.

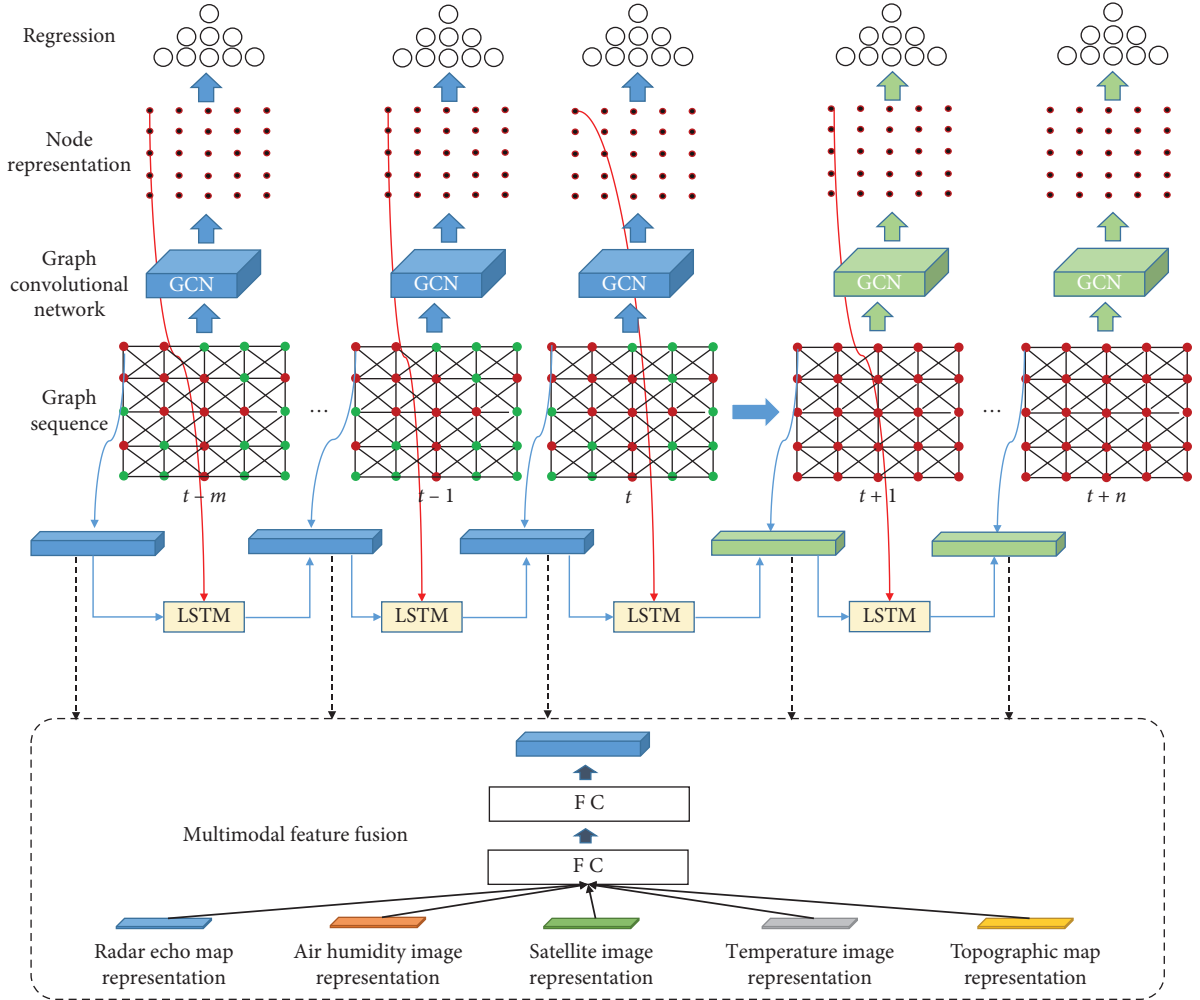


FIGURE 5: The framework of the proposed precipitation nowcasting model. The red nodes mean the precipitation of the corresponding local area is available while the green means the precipitation is unknown.

where  $W_f$  is the transformation parameter, that is, the weights of the fully connected layer, which are learned automatically in the training. This fusion unit is capable of learning the weights for different types of the features.

It is well known that precipitation is related to not only the current weather but also meteorological conditions over the past period. To handle this temporal influence, LSTM is utilized to establish relationships between the node at time  $t$  and the nodes before time  $t$ , which is formulated as

$$\mathbf{H}_t^0 = h(\mathbf{H}_{t-1}^0, \mathbf{H}_{t-1}^K, \mathbf{H}_t^0), \quad (10)$$

where  $h$  is a LSTM-based sequence model.

The optimal weight parameters are trained by minimizing the following loss function over all labeled nodes  $L$ , i.e.,

$$\mathcal{L} = - \sum_{i \in L} \sum_{j=0}^m \left\| \mathbf{P}_{(t-j)_i} - \mathbf{P}'_{(t-j)_i} \right\|, \quad (11)$$

where  $L$  indicates a set of the labeled nodes and  $j$  denotes the index of all observation graphs.

After the model training is completed, the future precipitation is easy to compute because  $\mathbf{y}_{t+n} = \mathbf{P}'_{t+n}$ .

## 4. Experiments

In this section, we first evaluate the proposed method on a large-scale real meteorological dataset. Then, we demonstrate extensive experimental results and analysis.

**4.1. Dataset.** The dataset used in the benchmark contains radar echo images, satellite images from 2016 to 2018, and topographic maps. The radar reflectivity images at 0.5° elevation with a resolution of  $416 \times 416$  pixels were collected from local new-generation weather radar stations (Hefei, Bengbu, and Huangshan) and cover a local area centered in each radar station. The topographic maps were collected from Baidu maps. Infrared and vapor channel images of FengYun-2 meteorological satellite images were chosen to increase the accuracy of precipitation nowcasting. The topographic maps and satellite images have the same spatial resolution and cover the same areas. Thus, a large-scale real meteorological dataset was obtained (partial raw data can be queried from <http://www.amo.org.cn/tcsj/zh/index.jsp>). All data were checked by five volunteers. Any sample data from consecutive 27 h can constitute a whole training sample (this paper aims to forecast precipitation at a place for the next 3 h based on the weather conditions of its neighbors over the past 24 h. Moreover, to match the sampling time, the precipitation herein is accumulated over the past 20 min). For each local radar station, we randomly selected 5000 samples as the training set, 500 samples as the validation set, and 2000 samples as the test set.

**4.2. Evaluation Metrics.** We utilize mean-square error (MSE) to evaluate model's effectiveness. The formulation is as follows:

$$\text{MSE} = \frac{1}{Q} \sum_{n=1}^N \|\mathbf{I}_{\text{pm}} - \mathbf{I}'_{\text{pm}}\|_2, \quad (12)$$

where  $Q$  is the number of test samples and smaller MSE means better performance.

Furthermore, to quantitatively evaluate our model's effectiveness influenced by the time variations, we compute MSE in different sampling time periods in the test phase:

$$\text{MSE}_{i \in \{0,8\}} = \frac{1}{Q} \left( \frac{1}{M \times N} \|\mathbf{I}_{\text{pm}_i} - \mathbf{I}'_{\text{pm}_i}\|_2 \right), \quad (13)$$

where  $i$  is the index of the  $i$ th precipitation map.

**4.3. Implementation Details.** We trained the proposed model from scratch on the training set and considered the model with the smallest mean-square error on the validation set as our final model. All images were normalized to  $[-1, +1]$ . Note that images from different sources have different resolutions. Therefore, we aligned and cropped all images

before the training. We first selected two key points as anchors and resized all images based on the anchor distance of each image. Then, we cropped all images to  $416 \times 416$  pixels based on the radar longitude and latitude. ResNet18 was utilized to extract feature representations of all images, and the feature dimension was set to 256 (that is, the average pooling layer of the ResNet18 was replaced by a 256-d fully connected layer).

The open source deep learning framework PyTorch was employed to implement the proposed semisupervised deep regression model. The parameter initialization for all layers was Xavier, and we used the Adam optimizer with an initial learning rate of 0.001.

## 4.4. Results and Discussion

**4.4.1. Ablation Study.** Table 1 presents the mean-square errors of different time nodes of the three radar stations. We can observe that the proposed framework predicts the precipitation in the next hour well. It also clearly shows that the error increases with time. Moreover, we can see that our model has the worst performance for Huangshan. We believe that the reason is that the causes of precipitation in mountainous areas are more complicated than elsewhere.

To the best of our knowledge, this is the first attempt to model various types of the observed data in a unified network. We describe the performed ablation study below to explore the performance influenced by different factors.

The inputs have five forms:

re: inputs that only include radar echo maps

re + si: inputs that include radar echo maps and satellite images

re + si + ah: inputs that include radar echo maps, satellite images, and air humidity images

re + si + ah + ti: inputs that include radar echo maps, satellite images, air humidity images, and temperature images

re + si + ah + ti + tm: inputs that include radar echo maps, satellite images, topographic map, temperature images, and air humidity images

From Table 2, we can see that all observed data used in this paper contribute to short-term precipitation nowcasting. Table 2 also shows that the topographic map usage improves the accuracy slightly, as compared to the other four factors. We think that the reason is that all other observed data are influenced by the topographic map. Therefore, the features extracted from the other factors represent the topographic map implicitly.

**4.4.2. Comparison with Existing Methods.** We note that some deep learning-based precipitation nowcasting methods have recently been proposed, e.g., [28–30]. However, they are all focused on supervised tasks. As we attempt to address the semisupervised automatic precipitation nowcasting problem, we compare our method with the two widely used semisupervised node representation methods,

TABLE 1: Mean-square errors (MSEs) of different time nodes in three radar stations.

Radar stations	MSE0	MSE1	MSE2	MSE3	MSE4	MSE5	MSE6	MSE7	MSE8	MSE
Hefei	0.06	0.11	0.19	0.28	0.59	1.17	1.93	2.61	3.6	1.17
Bengbu	0.04	0.08	0.15	0.29	0.64	1.12	1.86	2.57	3.16	1.1
Huangshan	0.07	0.12	0.22	0.40	0.92	1.58	2.11	2.86	3.74	1.34

TABLE 2: Mean-square errors (MSEs) of different methods in three radar stations.

Method	MSEs of different radar stations		
	Hefei	Bengbu	Huangshan
ICA by [27]	1.33	1.24	1.59
DeepWalk by [19]	1.67	1.46	2.01
Our (re)	3.62	3.42	3.98
Our (re + si)	2.48	2.51	2.64
Our (re + si + ah)	1.56	1.51	1.72
Our (re + si + ah + ti)	1.19	1.18	1.41
Our (re + si + ah + ti + tm)	<b>1.17</b>	<b>1.10</b>	<b>1.34</b>

e.g., DeepWalk by [19] and ICA by [27]. Specifically, the GCN utilized in the proposed framework is replaced by these two methods. The results are reported in Table 2. It can be observed that the proposed method achieves the best performance. We think that the major reason is that the proposed framework can handle the spatial (local region) relationship more effectively.

## 5. Conclusion and Future Work

In this paper, we propose a multimodal semisupervised deep graph learning framework for precipitation nowcasting. Particularly, multimodal factors, i.e., the radar echo maps, air humidity images, satellite images, temperature images, and available precipitation data over the past several hours with the corresponding topographic map are handled in a unified framework. In addition, we handle areas without precipitation stations, and thus, with unknown precipitation. A GCN with sequence information is presented herein this paper to model temporal and spatial information with semisupervised labels simultaneously. Using the methods above, we successfully implemented efficient precipitation nowcasting via a computer vision technique. To verify our method, we built large-scale real precipitation nowcasting datasets. Extensive experimentation demonstrated that our approach achieves superior performance to the baselines. In the future, we will explore how to embed attention mechanisms in our framework, which may improve the accuracy further.

## Data Availability

The data used to support the findings of this study are available at <http://data.cma.cn/site/index.html>.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Authors' Contributions

Kaichao Miao and Wei Wang contributed equally to this work and should be considered co-first authors.

## Acknowledgments

This work was supported by the Anhui Provincial Natural Science Foundation (no. 2008085QF295), Scientific Research Development Foundation of Hefei University (no. 19ZR15ZDA), and Talent Research Foundation of Hefei University (no. 18-19RC54).

## References

- [1] S. Otsuka, G. Tuerhong, R. Kikuchi et al., "Precipitation nowcasting with three-dimensional space-time extrapolation of dense and frequent phased-array weather radar observations," *Weather and Forecasting*, vol. 31, no. 1, pp. 329–340, 2016.
- [2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [3] J. Gao, T. Zhang, and C. Xu, "I know the relationships: zero-shot action recognition via two-stream graph convolutional networks and knowledge graphs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 8303–8311, Vancouver, Canada, July 2019.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [5] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7263–7271, Honolulu, HI, USA, July 2017.
- [6] Z. Qiu, T. Yao, and T. Mei, "Learning spatio-temporal representation with pseudo-3D residual networks," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5533–5541, Venice, Italy, October 2017.
- [7] C.-C. Chiu, T. N. Sainath, Y. Wu et al., "State-of-the-art speech recognition with sequence-to-sequence models," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, Calgary, Canada, pp. 4774–4778, April 2018.
- [8] S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 802–810, Montreal, Canada, December 2015.
- [9] X. Shi, Z. Gao, L. Lausen et al., "Deep learning for precipitation nowcasting: a benchmark and a new model," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 5617–5627, Long Beach, CA, USA, December 2017.
- [10] S. Singh, S. Sarkar, and P. Mitra, "Leveraging convolutions in recurrent neural networks for doppler weather radar echo

- prediction,” in *Proceedings of the International Symposium on Neural Networks*, Springer, Hokkaido, Japan, pp. 310–317, June 2017.
- [11] Z. Karevan and J. A. Suykens, “Spatio-temporal stacked LSTM for temperature prediction in weather forecasting,” 2018, <http://arxiv.org/abs/1811.063411>.
- [12] E. Shi, Q. Li, D. Gu, and Z. Zhao, “A method of weather radar echo extrapolation based on convolutional neural networks,” in *Proceedings of the International Conference on Multimedia Modeling*, Springer, Bangkok, Thailand, pp. 16–28, February 2018.
- [13] Pechyonkin, *Deep Learning Vision for Non-Vision Tasks*, 2018, <https://pechyonkin.me/deep-learning-vision-non-vision-tasks>.
- [14] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [15] W.-c. Woo and W.-k. Wong, “Operational application of optical flow techniques to radar-based rainfall nowcasting,” *Atmosphere*, vol. 8, no. 3, p. 48, 2017.
- [16] X. Zhu, Z. Ghahramani, and J. D. Lafferty, “Semi-supervised learning using gaussian fields and harmonic functions,” in *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 912–919, Washington, DC, USA, August 2003.
- [17] M. Belkin, P. Niyogi, and V. Sindhwani, “Manifold regularization: a geometric framework for learning from labeled and unlabeled examples,” *Journal of Machine Learning Research*, vol. 7, pp. 2399–2434, 2006.
- [18] J. Weston, F. Ratle, H. Mobahi, and R. Collobert, “Deep learning via semi-supervised embedding,” in *Neural Networks: Tricks of the Trade*, pp. 639–655, Springer, Heidelberg, Germany, 2012.
- [19] B. Perozzi, R. Al-Rfou, and S. Skiena, “Deepwalk: online learning of social representations,” in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, pp. 701–710, August 2014.
- [20] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, “Line: large-scale information network embedding,” in *Proceedings of the 24th International Conference on World Wide Web*, International World Wide Web Conferences Steering Committee, Florence, Italy, pp. 1067–1077, May 2015.
- [21] A. Grover and J. Leskovec, “Node2vec: scalable feature learning for networks,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, San Francisco, CA, USA, pp. 855–864, August 2016.
- [22] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” 2016, <http://arxiv.org/abs/1609.02907>.
- [23] Z.-M. Chen, X.-S. Wei, P. Wang, and Y. Guo, “Multi-label image recognition with graph convolutional networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5177–5186, Long Beach, CA, USA, June 2019.
- [24] X. Wang, Y. Ye, and A. Gupta, “Zero-shot recognition via semantic embeddings and knowledge graphs,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6857–6866, Salt Lake City, UT, USA, June 2018.
- [25] L. Wu, P. Sun, R. Hong, Y. Fu, X. Wang, and M. Wang, “Socialgcn: an efficient graph convolutional network based model for social recommendation,” 2018, <http://arxiv.org/abs/1811.02815>.
- [26] L. Yao, C. Mao, and Y. Luo, “Graph convolutional networks for text classification,” 2018, <http://arxiv.org/abs/1809.05679>.
- [27] Q. Lu and L. Getoor, “Link-based classification,” in *Proceedings of the 20th International Conference on Machine Learning*, ICML-03, Washington, DC, USA, pp. 496–503, August 2003.
- [28] Q.-K. Tran and S.-k. Song, “Computer vision in precipitation nowcasting: applying image quality assessment metrics for training deep neural networks,” *Atmosphere*, vol. 10, no. 5, p. 244, 2019.
- [29] J. Jing, Q. Li, and X. Peng, “MLC-LSTM: exploiting the spatiotemporal correlation between multi-level weather radar echoes for echo sequence extrapolation,” *Sensors*, vol. 19, no. 18, p. 3988, 2019.
- [30] Q.-K. Tran and S.-k. Song, “Multi-channel weather radar echo extrapolation with convolutional recurrent neural networks,” *Remote Sensing*, vol. 11, no. 19, p. 2303, 2019.

## Research Article

# Some New Bounds of Weighted Graph Entropies with GA and Gaurava Indices Edge Weights

Tiejun Wu,<sup>1</sup> Hafiz Mutee Ur Rehman ,<sup>2,3</sup> Yu-Ming Chu ,<sup>4,5</sup> Deeba Afzal,<sup>2</sup> and Jianfeng Yu<sup>1</sup>

<sup>1</sup>College of Mechanical and Electronics Engineering, Dongguan Polytechnic, Dongguan 523808, Guangdong, China

<sup>2</sup>Department of Mathematics and Statistics, The University of Lahore, Lahore, Pakistan

<sup>3</sup>Department of Mathematics, Division of Science and Technology, University of Education, Lahore, Pakistan

<sup>4</sup>Department of Mathematics, Huzhou University, Huzhou 313000, China

<sup>5</sup>Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science & Technology, Changsha 410114, China

Correspondence should be addressed to Yu-Ming Chu; [chuyuming@zjhu.edu.cn](mailto:chuyuming@zjhu.edu.cn)

Received 27 April 2020; Accepted 17 August 2020; Published 30 September 2020

Academic Editor: Alessandro Gasparetto

Copyright © 2020 Tiejun Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Motivated by the concept of Shannon's entropy, the degree-dependent weighted graph entropy was defined which is now become a tool for measurement of structural information of complex graph networks. The aim of this paper is to study weighted graph entropy. We used GA and Gaurava indices as edge weights to define weighted graph entropy and establish some bounds for different families of graphs. Moreover, we compute the defined weighted entropies for molecular graphs of some dendrimer structures.

## 1. Introduction

The branch of mathematics known as graph theory provides tools for solving problems of information theory, computer sciences, physics, and chemistry [1–3]. Among them, special places are reserved for so-called topological descriptors, which play an important role in mathematical chemistry, especially in QSPR/QSAR surveys. Many topological descriptors are introduced and studied in the literature, such as the Zagreb index [4–6], the Randić connectivity index [7], and the modified Zagreb index [8, 9].

In the year 2009, Vukičević and Furtula [10] introduced the geometrical arithmetic (GA) index as a molecular descriptor, and mathematical formula for this index is,

$$GA_1 = \sum_{uv \in E(G)} \frac{2\sqrt{d_u d_v}}{d_u + d_v}. \quad (1)$$

There are many interesting attributes on topological indices, and different physicochemical properties of

hydrocarbons can be obtained from this index [11–16]. The predictive power of the GA index is compared with others such as the famous Randić index [17]. Due to this reason, different versions of the GA index are now investigated and introduced in the literature [18–20]. The other famous indices are given in [21]. The mathematical formulae for first and second Gaurava indices are

$$GO_1 = \sum_{uv \in E(G)} [(d_u + d_v) + (d_u \cdot d_v)], \quad (2)$$

$$GO_2 = \sum_{uv \in E(G)} [(d_u + d_v) \cdot (d_u \cdot d_v)], \quad (3)$$

respectively.

Many problems of information theory, biology, computer sciences, chemistry, and discrete mathematics are directly solved by utilizing different kinds of graph measures and the graph entropy is one of the powerful tools [22–26] that help to understand the structural

complexity of graph networks [27]. Different molecular descriptors are used to introduce weighted graph entropies [28–31].

In this paper, we extended the work of [28, 30] and introduced weighted graph entropies by using GA and Gaurava indices as edge weights. We examine the extreme properties of these entropies for some special graph families. We also computed these entropies for different chemical structures.

Now, we define some basic notions about entropy of graphs. We always consider  $G$  being a connected graph with  $E$  as the set of edges,  $V$  as the set of vertices, and  $w$  to be the edge weight given to the edges of graph  $G$  that will be used to define weighted graph entropy. As mentioned above, GA and Gaurava indices are taken as edge weights in this paper.  $d_u$  denoted the degree of a vertex  $u$  that is defined as the total number of vertices of  $G$  that are at distance one from the vertex  $u$ . Consider

$$p_{ij} = \frac{w(v_i v_j)}{\sum_{j=1} d_i w(v_i v_j)}, \quad (4)$$

where  $w(v_i v_j)$  is the weight  $v_i v_j$  and  $w(v_i v_j) > 0$ .

Now, the weighted graph entropy can be defined by

$$H(v_i) = - \sum p_{ij} \log(p_{ij}). \quad (5)$$

*Definition 1.* For the graph  $G$ , the weighted entropy can be defined as follows [28, 30]:

$$I(G, w) = - \sum_{uv \in E(G)} p_{uv} \log(p_{uv}). \quad (6)$$

Here,  $p_{uv}$  is same as that given in (4).

## 2. Main Results

In this section, we are going to present our main results.

**Lemma 1.** Let  $G$  be a simple graph having  $n$  vertices and  $m$  edges and let  $\Delta$  and  $\delta$  be the maximum degree and minimum degree of a vertex, respectively; then,

$$\begin{aligned} & \frac{2\nu\sqrt{\Delta}}{\Delta+1} + \sqrt{\mu^2 - \frac{\mu}{4\delta_1^2} \left[ \sum_1 d_i^3 - 2M_2(G) - \nu(\delta_1 - 1)^2 \right] - \frac{\mu^2}{4} \left( 1 - \frac{2\sqrt{\Delta}\delta_1}{\Delta+1} \right)^2} \\ & \leq GA_1 \leq \frac{2\nu\sqrt{\delta_1}}{\delta_1+1} + \sqrt{\mu^2 - \frac{\mu}{4\Delta^2} \left[ \sum_1 d_i^3 - 2M_2(G) - \nu(\Delta - 1)^2 \right]}, \end{aligned} \quad (7)$$

with equality if and only if  $G$  is a regular graph or  $G$  is a bipartite graph.

**Lemma 2.** Consider  $G$  is a simple connected graph having  $m$  edges. Then, we have

$$GA_1 \leq \frac{2\nu\sqrt{\Delta}}{\Delta+1} + \frac{1}{\Delta} \sqrt{M_2(G) - \nu\Delta + (m - \nu)\delta_1^2}, \quad (8)$$

with equality if and only if  $G$  is isomorphic to  $K_{1, m-1}$  or  $G$  is isomorphic to regular graph or  $G$  is isomorphic to  $(\Delta, 1)$  semiregular graph.

**Lemma 3.** Consider  $G$  is a simple connected graph having  $m$  edges. Then, we have

$$GA_1 \geq \frac{2\nu\sqrt{\Delta}}{\Delta+1} \frac{8\sqrt{(\Delta+\delta_1)(\Delta\delta_1)}}{(\sqrt{\delta} + \sqrt{\Delta})^2} \sqrt{(m - \nu)^2 \left( \frac{m - \nu}{4\delta_1^2} \right) \left[ \sum_1 d_i^3 - 2M_2(G) - \nu(\delta_1 - 1)^2 \right]}. \quad (9)$$

Moreover, the equality holds if and only if  $G$  is isomorphic to a regular graph or  $G$  is isomorphic to  $(\Delta, 1)$  semiregular graph.

**Lemma 4.** Consider  $G$  is a simple connected graph having  $m$  edges. Then, we have

$$GA_1 \leq \frac{2m\sqrt{2}(n-1)}{n+1} - 2\nu \left( \frac{\sqrt{2(n-1)}}{n+1} - \frac{\sqrt{n-1}}{n} \right), \quad (10)$$

with equality if and only if  $G$  is isomorphic to  $K_{1, n-1}$  or  $G$  is isomorphic to a complete graph  $K_3$ .

**Lemma 5.** Consider  $G$  is a simple connected graph having  $m$  edges. Then, we have



$$GA_1 \leq \left\lceil \frac{m-1}{n-1} \right\rceil + \sqrt{\left\lceil \frac{m-1}{n-1} \right\rceil^2 + \frac{M_2(G)}{n-1}}. \quad (11)$$

**Theorem 1.** Consider  $G$  is a connected graph having  $n$  vertices with  $n \geq 3$ . Then, we have

$$\begin{aligned} \log(GA_1) + \log\left(\frac{2}{n}\right) &\leq I(G, GA_1) \leq \log(GA_1) \\ &+ \log\left(\frac{n}{4}\right); \end{aligned} \quad (12)$$

$$\begin{aligned} \log(GO_1) + \log 4(n-1) &\leq I(G, GO_1) \leq \log(GO_1) \\ &- \log(3(n-1)^2). \end{aligned}$$

*Proof.* We prove the result for  $GO_1$ , the other results can be proved in the same manner.

Since  $G$  is a connected graph with  $n$  number of vertices, for any vertex, the maximum possible degree can be  $n-1$  and the minimum possible degree can be one. Hence, for any edge  $uv$ , the minimum degrees for  $u$  and  $v$  can be 1 and 2 and maximum possible degrees for  $u$  and  $v$  can be  $n-1$  and  $n-1$ , and hence, we have

---


$$\begin{aligned} I(G, GO_1) &= \log(GO_1) - \frac{1}{GO_1} \sum_{uv \in E(G)} [(d_u + d_v) + (d_u \cdot d_v)] \cdot \log \left[ \sum_{uv \in E(G)} (d_u + d_v) + (d_u \cdot d_v) \right] \\ &\leq \log(GO_1) - [\log(3) + \log(n-1)^2] \\ &= \log(GO_1) - \log(3(n-1)^2) \\ &\geq \log(GO_1) - \frac{1}{GO_1} \sum_{uv \in E(G)} [(d_u + d_v) + (d_u \cdot d_v)] \cdot \log \left[ \sum_{uv \in E(G)} (d_u + d_v) + (d_u \cdot d_v) \right] \\ &\geq \log(GO_1) - [\log 2(n-2) + \log(2)] \\ &= \log(GO_1) + \log(4n-1). \end{aligned} \quad (13)$$

Therefore,

$$\begin{aligned} \log(GO_1) + \log 4(n-1) &\leq I(G, GO_1) \leq \log(GO_1) \\ &- \log(3(n-1)^2). \end{aligned} \quad (14) \quad \square$$

**Theorem 2.** Let  $G$  be a graph with  $n$  vertices. Let  $\delta$  and  $\Delta$  be the minimum and maximum degrees of  $G$ , respectively. Then, we have

$$\begin{aligned} \log(GA_1) + \log\left(\frac{\delta}{\Delta}\right) &\leq I(G, GA_1) \\ \log(GO_1) - \log(2\Delta\delta^2) &\geq I(G, ReZ_1) \\ &\geq \log(GO_1) - \log(2\delta\Delta^2). \end{aligned} \quad (15)$$

*Proof.* Since  $G$  is the connected graph with  $n$  number of vertices, for any vertex, the maximum possible degree can be  $n-1$  and the minimum possible degree can be one. Hence, for any edge  $uv$ , the minimum degrees for  $u$  and  $v$  can be 1

and 2 and maximum possible degrees for  $u$  and  $v$  can be  $n-1$  and  $n-1$ , and hence, we have

---


$$\begin{aligned} I(G, GO_1) &= \log(GO_1) - \frac{1}{GO_1} \sum_{uv \in E(G)} [(d_u + d_v) + (d_u \cdot d_v)] \\ &\cdot \log \left[ \sum_{uv \in E(G)} (d_u + d_v) + (d_u \cdot d_v) \right] \\ &= \log(GO_1) - \frac{1}{GO_1} \sum_{uv \in E(G)} (d_u + d_v) + (d_u \cdot d_v) \\ &\cdot [\log(d_u + d_v) + \log(d_u \cdot d_v)] \\ &= \log(GO_1) - \log(2\Delta\delta^2). \end{aligned} \quad (16)$$

Also,

$$I(G, GO_1) \geq \log(GO_1) - \log(2\delta\Delta^2). \quad (17) \quad \square$$

**Theorem 3.** Consider  $G$  is a regular graph having  $n$  vertices with  $n \geq 3$ . Then, we have

$$\log(n) \leq I(G, GA_1) \leq \log\left(\frac{n(n-1)}{2}\right). \quad (18)$$

Note that left inequality turns into equality if  $G$  is a cyclic graph, and the right inequality turns into equality if  $G$  is a complete graph.

**Theorem 4.** Consider  $G$  is a complete bipartite graph having  $n$  vertices. Then, we have

$$\log(n-1) \leq I(G, GA_1) \leq \log\left(\left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil\right). \quad (19)$$

Note that the left inequality turns into equality if  $G$  is a star graph, and  $\log(n-1) = \log(\lfloor n/2 \rfloor \lceil n/2 \rceil)$  if and only if  $G$  is a complete bipartite graph (balanced).

**Theorem 5.** Let  $T$  be a tree of order  $n$  ( $n > 2$ ) with maximum degree vertex  $\Delta$ ; then, we have

$$\log(GA_1) \leq \log\left(\frac{2p\sqrt{\Delta}}{\Delta+1}\right). \quad (20)$$

**Theorem 6.** Let  $G$  be a simple graph having  $n$  vertices and  $m$  edges and let  $\Delta$  and  $\delta$  be the maximum degree and minimum degree of a vertex, respectively; then,

$$\log\left(\frac{2m\sqrt{\delta \cdot \Delta}}{\delta + \Delta}\right) \leq I(G, GA_1), \quad (21)$$

with equality if and only if  $G$  is either a regular graph or a bipartite graph.

In graph theory, the molecular graph is obtained by taking atoms as vertices and bonds as edges. It can be noted that the maximum possible degree for a vertex in a molecular graph is four. Following theorem is about the bounds of weighted entropy for the molecular graph.

**Theorem 7.** Consider a molecular graph  $G$  having  $n$  vertices. Then, we have

$$\log(GA_1) - \log 4 \leq I(G, GA_1) \leq \log(GA_1) + \log\sqrt{2}. \quad (22)$$

### 2.1. Relation of Entropy with Zagreb Indices

**Theorem 8.** Consider  $G$  is a simple connected graph. Then, we have

$$\log\left(\frac{2M_1(G)}{n(n-1)}\right) \leq I(G, GA_1) \leq \log(M_1(G))^2. \quad (23)$$

The equality holds if and only if  $G$  is a union of  $K_2$ .

**Theorem 9.** Consider  $G$  is any graph, then we have

$$I(G, GA_1) \geq \log\left(\frac{2\sqrt{mM_2(G)}}{(n-1)}\right). \quad (24)$$

Equality holds if and only if  $G$  is isomorphic to the regular graph.

**Theorem 10.** Consider  $G$  is any graph, then we have

$$\log\left(\frac{2\sqrt{mM_2(G) + m(m-1)}}{(n-1)}\right) \leq I(G, GA_1). \quad (25)$$

Equality holds if and only if  $G$  is isomorphic to the regular graph.

**Theorem 11.** Consider  $G$  is any graph, then we have

$$I(G, GA_1) \leq \log\left(\frac{4(n-1)^2\nu\sqrt{\Delta}}{\Delta+1} + \frac{2(n-1)}{\Delta} \cdot \sqrt{M_2(G) - \nu\Delta + (m-\nu)\delta_1^2}\right). \quad (26)$$

Equality holds if and only if  $G$  is isomorphic to the regular graph.

**2.2. Numerical Examples.** Here, we compute the weighted entropies introduced in this paper for some chemical structures.

**Example 1.** Consider the porphyrin dendrimers shown in Figure 1. We denote the graph of porphyrin dendrimers by  $G$ , and the edge partition of  $G$  is given in Table 1. Using Table 1 and definition of entropy,

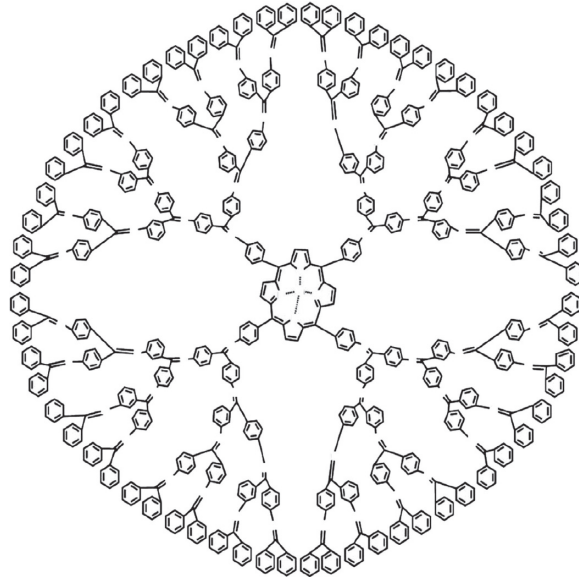


FIGURE 1: Porphyrin dendrimer.

TABLE 1: Edge partition of porphyrin dendrimers.

$(d_u, d_v)$	(1, 3)	(1, 4)	(2, 2)	(2, 3)	(3, 3)	(3, 4)
Number of edges	$2n$	$24n$	$10n - 5$	$48n - 6$	$13n -$	$8n -$

we have the following entropies for porphyrin dendrimers:

$$\begin{aligned}
 I(G, GA_1) &= \log(GA_1) - \frac{1}{GA_1} \sum_{uv \in E(G)} 2 \frac{\sqrt{d_u d_v}}{d_u \cdot d_v} \cdot \log\left(2 \frac{\sqrt{d_u d_v}}{d_u \cdot d_v}\right) \\
 &= \log(98.88020n - 10.87877) - \frac{1}{98.88020n - 10.87877} \left[ |E_1| \left( 2 \frac{\sqrt{3}}{4} \cdot \log\left(2 \frac{\sqrt{3}}{4}\right) \right) \right. \\
 &\quad + |E_2| \left( 2 \frac{\sqrt{4}}{5} \cdot \log\left(2 \frac{\sqrt{4}}{5}\right) \right) + |E_3| \left( 2 \frac{\sqrt{4}}{4} \cdot \log\left(2 \frac{\sqrt{4}}{4}\right) \right) + |E_4| \left( 2 \frac{\sqrt{6}}{4} \cdot \log\left(2 \frac{\sqrt{6}}{4}\right) \right) \\
 &\quad \left. + |E_5| \left( 2 \frac{\sqrt{6}}{6} \cdot \log\left(2 \frac{\sqrt{6}}{6}\right) \right) + |E_7| \left( 2 \frac{\sqrt{12}}{7} \cdot \log\left(2 \frac{\sqrt{12}}{7}\right) \right) \right] \\
 &= \log(98.88020n - 10.87877) - \frac{1}{98.88020n - 10.87877} \left[ (2n) \left( 2 \frac{\sqrt{3}}{4} \cdot \log\left(2 \frac{\sqrt{3}}{4}\right) \right) \right. \\
 &\quad + (24n) \left( 2 \frac{\sqrt{4}}{5} \cdot \log\left(2 \frac{\sqrt{4}}{5}\right) \right) + (10n - 5) \left( 2 \frac{\sqrt{4}}{4} \cdot \log\left(2 \frac{\sqrt{4}}{4}\right) \right) + (48n - 6) \left( 2 \frac{\sqrt{6}}{5} \cdot \log\left(2 \frac{\sqrt{6}}{5}\right) \right) \\
 &\quad \left. + (13n) \left( 2 \frac{\sqrt{6}}{6} \cdot \log\left(2 \frac{\sqrt{6}}{6}\right) \right) + (8n) \left( 2 \frac{\sqrt{12}}{7} \cdot \log\left(2 \frac{\sqrt{12}}{7}\right) \right) \right],
 \end{aligned} \tag{27}$$

$$I(G, GA_1) = \log(GA_1) - \frac{1}{GA_1} (-23.308n - 2.695),$$

$$\begin{aligned}
I(G, GO_1) &= \log(GO_1) - \frac{1}{GO_1} \sum_{uv \in E(G)} [(d_u + d_v) + (d_u \cdot d_v)] \cdot \log \left[ \sum_{uv \in E(G)} (d_u + d_v) + (d_u \cdot d_v) \right] \\
&= \log(1169n - 106) - \frac{1}{1169n - 106} \left[ |E_1| [(1 + 3) + (1.3)] \cdot \log[(1 + 3) + (1.3)] \right. \\
&\quad + |E_2| [(1 + 4) + (1.4)] \cdot \log[(1 + 4) + (1.4)] + |E_3| [(2 + 2) + (2.2)] \cdot \log[(2 + 2) + (2.2)] \\
&\quad + |E_4| [(2 + 3) + (2.3)] \cdot \log[(2 + 3) + (2.3)] + |E_5| [(3 + 3) + (3.3)] \cdot \log[(3 + 3) + (3.3)] \\
&\quad \left. + |E_6| [(3 + 4) + (3.4)] \cdot \log[(3 + 4) + (3.4)] \right] \tag{28} \\
&= \log(1169n - 106) - \frac{1}{1169n - 106} [(14n) \cdot \log[7]] \\
&\quad + |E_2| [(1 + 4) + (1.4)] \cdot \log[(1 + 4) + (1.4)] + |E_3| [(2 + 2) + (2.2)] \cdot \log[(2 + 2) + (2.2)] \\
&\quad + |E_4| [(2 + 3) + (2.3)] \cdot \log[(2 + 3) + (2.3)] + |E_5| [(3 + 3) + (3.3)] \cdot \log[(3 + 3) + (3.3)] \\
&\quad + |E_6| [(3 + 4) + (3.4)] \cdot \log[(3 + 4) + (3.4)], \\
I(G, GO_1) &= \log(GA_1) - \frac{1}{GA_1} (-23.308n - 2.695),
\end{aligned}$$

*Example 2.* The graph  $G$  of zinc-porphyrin dendrimer is shown in Figure 2, and the edge partition for this dendrimer

is given in Table 2. We have the following computations for the entropies of zinc-porphyrin dendrimer.

$$\begin{aligned}
I(G, GA_1) &= \log(GA_1) - \frac{1}{GA_1} \sum_{uv \in E(G)} 2 \frac{\sqrt{d_u d_v}}{d_u \cdot d_v} \cdot \log \left( 2 \frac{\sqrt{d_u d_v}}{d_u \cdot d_v} \right) \\
&= \log(44.9312^n - 22.226) - \frac{1}{2(44.9312^n - 22.226)} \left[ |E_1| \left( \frac{\sqrt{2}}{2} \right) \cdot \log \left( \frac{\sqrt{2}}{2} \right) \right. \\
&\quad + |E_2| \left( \frac{\sqrt{2}}{2} \right) \cdot \log \left( \frac{\sqrt{2}}{2} \right) + |E_3| \left( \frac{2}{3} \right) \cdot \log \left( \frac{2}{3} \right) + |E_4| \left( \frac{\sqrt{15}}{6} \right) \cdot \log \left( \frac{\sqrt{15}}{6} \right) \tag{29} \\
&= \log(44.9312^n - 22.226) - \frac{1}{2(44.9312^n - 22.226)} \left[ (16 \cdot 2^n - 4) \left( \frac{\sqrt{2}}{2} \right) \cdot \log \left( \frac{\sqrt{2}}{2} \right) \right. \\
&\quad \left. + (40 \cdot 2^n - 16) \left( \frac{\sqrt{2}}{2} \right) \cdot \log \left( \frac{\sqrt{2}}{2} \right) + (8 \cdot 2^n - 16) \left( \frac{2}{3} \right) \cdot \log \left( \frac{2}{3} \right) + (4) \left( \frac{\sqrt{15}}{6} \right) \cdot \log \left( \frac{\sqrt{15}}{6} \right) \right].
\end{aligned}$$

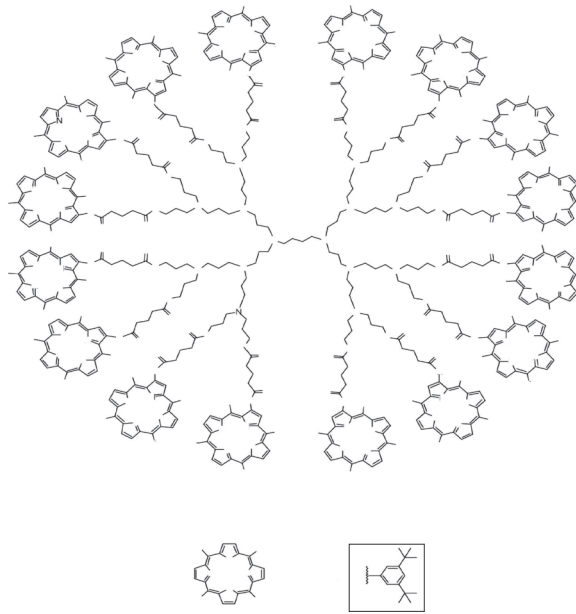


FIGURE 2: Zinc-porphyrin dendrimer.

TABLE 2: Edge partition of zinc-porphyrin dendrimers.

$(d_u, d_v)$	(2, 2)	(2, 3)	(3, 3)	(3, 4)
Number of edges	$16 \cdot 2^n - 4$	$40 \cdot 2^n - 16$	$8 \cdot 2^n - 16$	4

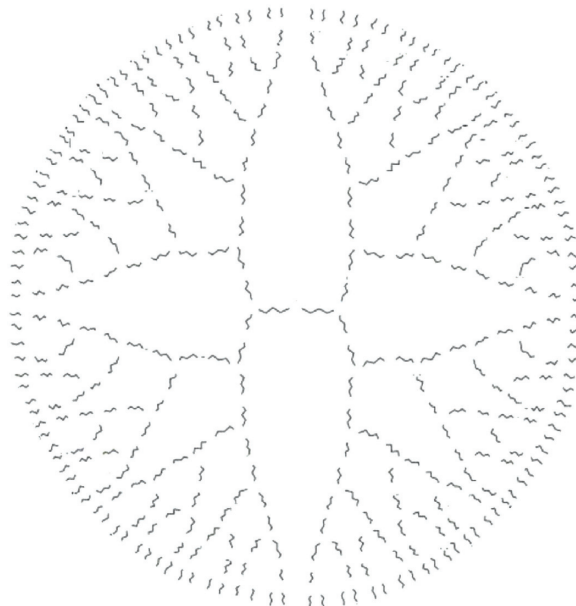


FIGURE 3: Poly(ethylene amidoamine) dendrimer.

TABLE 3: Edge partition of the poly(ethylene amidoamine) dendrimer.

$(d_u, d_v)$	(1, 2)	(1, 3)	(2, 2)	(2, 3)
Number of edges	$4 \cdot 2^n$	$4 \cdot 2^n - 2$	$16 \cdot 2^n$	$20 \cdot 2^n - 9$

*Example 3.* Let  $G$  be the graph of poly(ethylene amido-amine) dendrimers as shown in Figure 3. Then, the edge

partition of this dendrimer is given in Table 3 and we have the following results.

$$\begin{aligned}
 I(G, GA_1) &= \log(GA_1) - \frac{1}{GA_1} \sum_{uv \in E(G)} 2 \frac{\sqrt{d_u d_v}}{d_u \cdot d_v} \cdot \log\left(2 \frac{\sqrt{d_u d_v}}{d_u \cdot d_v}\right) \\
 &= \log(31.5502^n - 13.653) - \frac{1}{2(31.5502^n - 13.653)} \left[ |E_1| \left(\frac{\sqrt{2}}{2}\right) \cdot \log\left(\frac{\sqrt{2}}{2}\right) \right. \\
 &\quad \left. + |E_2| \left(\frac{\sqrt{6}}{3}\right) \cdot \log\left(\frac{\sqrt{6}}{3}\right) + |E_3| \left(\frac{2}{2}\right) \cdot \log\left(\frac{2}{2}\right) + |E_4| \left(\frac{\sqrt{2}}{2}\right) \cdot \log\left(\frac{\sqrt{2}}{2}\right) \right] \quad (30) \\
 &= \log(31.5502^n - 13.653) - \frac{1}{2(31.5502^n - 13.653)} \left[ (4.2^n) \left(\frac{\sqrt{2}}{2}\right) \cdot \log\left(\frac{\sqrt{2}}{2}\right) \right. \\
 &\quad \left. + (4.2^n - 2) \left(\frac{\sqrt{6}}{3}\right) \cdot \log\left(\frac{\sqrt{6}}{3}\right) + (16.2^n) \left(\frac{2}{2}\right) \cdot \log\left(\frac{2}{2}\right) + (20.2^n - 9) \left(\frac{\sqrt{2}}{2}\right) \cdot \log\left(\frac{\sqrt{2}}{2}\right) \right].
 \end{aligned}$$

### 3. Conclusion

In information theory, the graph entropy is a measure of the information rate achievable by communicating symbols over a channel in which certain pairs of values may be confused. This measure, first introduced by Körner in the 1970s, has since also proven itself useful in other settings, including combinatorics. In this paper, we have studied graph entropy with GA and Gaurava indices and justified it by some numerical examples. It would be interesting to work on entropy of weighted graphs with some other degree- and distance-based topological indices. The bounds of degree-based network entropy can also be used in national security, Internet networks, social networks, structural chemistry, ecological networks, computational systems biology, etc. They will play an important role in analyzing structural symmetry and asymmetry in real networks in the future.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they do not have any conflicts of interest.

### Authors' Contributions

Tiejun Wu enhanced the introduction section and improved the reference list. Hafiz Mutee Ur Rehman wrote the original paper. Yu-Ming Chu gave the equality conditions for Theorems 1, 2, and 7 and verified all results. Deeba Afzal supervised the work and Jianfeng Yu prepared the revision and arranged funding for the paper.

### Acknowledgments

The research was supported by the National Natural Science Foundation of China (Grant nos. 11971142, 11871202, 61673169, 11701176, 11626101, and 11601485). This work was also supported by the Scientific Research Project of Department of Education of Guangdong Province (natural science) (no. 2017GKTSCX102), Scientific Research Project of Department of Education of Guangdong Province (innovation) (No. 2017GGXJK095), Teaching Reform Project of Guangdong Higher Vocational Education Machinery Manufacturing Major Teaching Steering Committee (JZ201907), and Curriculum Ideological and Political Demonstration Curriculum Construction Project of Dongguan Polytechnic (KCSZ202002).

### References

- [1] A. Tabassum, M. A. Umar, M. Perveen, and A. Raheem, "Antimagicness of subdivided fans," *Open Journal of Mathematical Sciences*, vol. 4, no. 1, pp. 18–22, 2020.
- [2] H. Fadhil, H. F. M. Salih, and S. M. Mershkhan, "Generalized the Liouville's and Möbius functions of graph," *Open Journal of Mathematical Sciences*, vol. 4, no. 1, pp. 186–194, 2020.
- [3] U. Ali, Y. Ahmad, and M. S. Sardar, "On 3-total edge product cordial labeling of tadpole, book and flower graphs," *Open Journal of Mathematical Sciences*, vol. 4, no. 1, pp. 48–55, 2020.
- [4] M. Munir, W. Nazeer, S. Rafique, and S. Kang, "M-polynomial and degree-based topological indices of polyhex nanotubes," *Symmetry*, vol. 8, no. 12, p. 149, 2016.
- [5] W. Gao, M. Asif, and W. Nazeer, "The study of honey comb derived network via topological indices," *Open Journal of Mathematical Analysis*, vol. 2, no. 2, pp. 10–26, 2018.
- [6] A. Aslam, M. K. Jamil, W. Gao, and W. Nazeer, "Topological aspects of some dendrimer structures," *Nanotechnology Reviews*, vol. 7, no. 2, pp. 123–129, 2018.



- [7] X. Li and Y. Shi, "A survey on the randic index," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 59, no. 1, pp. 127–156, 2008.
- [8] K. C. Das, K. Das, and I. Gutman, "Some properties of the second Zagreb index," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 52, no. 52, pp. 3–1, 2004.
- [9] F. Asif, Z. Zahid, and S. Zafar, "Leap Zagreb and leap hyper-Zagreb indices of Jahangir and Jahangir derived graphs," *Engineering and Applied Science Letter*, vol. 3, no. 2, pp. 1–8, 2020.
- [10] D. Vukićević and B. Furtula, "Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges," *Journal of Mathematical Chemistry*, vol. 46, no. 4, pp. 1369–1376, 2009.
- [11] A. Iranmanesh and M. Zeraatkar, "Computing Ga index of HAC5C7 [p, q] and HAC5C6C7 [p, q] nanotubes," *Optoelectronics and Advanced Materials-Rapid Communications*, vol. 5, no. 7, pp. 790–792, 2011.
- [12] M. Ghorbani and M. Ghazi, "Computing some topological indices of triangular Benzenoid," *Digest Journal of Nanomaterials and Biostructures*, vol. 5, no. 4, pp. 1107–1111, 2010.
- [13] W. Gao, M. Younas, A. Farooq, A. Virk, and W. Nazeer, "Some reverse degree-based topological indices and polynomials of dendrimers," *Mathematics*, vol. 6, no. 10, p. 214, 2018.
- [14] Y. Kwun, A. Virk, W. Nazeer, M. Rehman, and S. Kang, "On the multiplicative degree-based topological indices of silicon-carbon  $\text{Si}_2\text{C}_3\text{-I}[p,q]$  and  $\text{Si}_2\text{C}_3\text{-II}[p,q]$ ," *Symmetry*, vol. 10, no. 8, p. 320, 2018.
- [15] M. S. Ahmad, W. Nazeer, S. M. Kang, M. Imran, and W. Gao, "Calculating degree-based topological indices of dominating David derived networks," *Open Physics*, vol. 15, no. 1, pp. 1015–1021, 2017.
- [16] Y. C. Kwun, M. A. Zahid, W. Nazeer, A. Ali, M. Ahmad, and S. M. Kang, "On the zagreb polynomials of benzenoid systems," *Open Physics*, vol. 16, no. 1, pp. 734–740, 2018.
- [17] M. Mogharrab and G. H. Fath-Tabar, "Some bounds on  $\text{GA}_1$  index of graphs," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 65, pp. 33–38, 2011.
- [18] M. K. Siddiqui, M. Naeem, N. A. Rahman, and M. Imran, "Computing topological indices of certain networks," *Journal of Optoelectronics and Advanced Materials*, vol. 18, pp. 884–892, 2016.
- [19] S. Deng, W. Su, and L. Zhao, "A neural network for predicting normal boiling point of pure refrigerants using molecular groups and a topological index," *International Journal of Refrigeration*, vol. 63, pp. 63–71, 2016.
- [20] S. M. Kang, M. Yousaf, M. A. Zahid, M. Younas, and W. Nazeer, "Zagreb Polynomials and redefined Zagreb indices of nanostar dendrimers," *Open Physics*, vol. 17, no. 1, pp. 31–40, 2019.
- [21] A. Farooq, M. Habib, A. Mahboob, W. Nazeer, and S. M. Kang, "Zagreb polynomials and redefined zagreb indices of dendrimers and polyomino chains," *Open Chemistry*, vol. 17, no. 1, pp. 1374–1381, 2019.
- [22] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," *Information Sciences*, vol. 181, no. 1, pp. 57–78, 2011.
- [23] G. Simonyi, "Graph entropy: a survey," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 20, pp. 399–441, 1995.
- [24] J. Shetty and J. Adibi, "Discovering important nodes through graph entropy the case of enron email database," in *Proceedings of the 3rd International Workshop on Link Discovery—LinkKDD '05*, pp. 74–81, Chicago, IL, USA, August 2005.
- [25] M. Dehmer, "Information processing in complex networks: graph entropy and information functionals," *Applied Mathematics and Computation*, vol. 201, no. 1-2, pp. 82–94, 2008.
- [26] S. Cao, M. Dehmer, and Y. Shi, "Extremality of degree-based graph entropies," *Information Sciences*, vol. 278, pp. 22–33, 2014.
- [27] I. Newman, P. Ragde, and A. Wigderson, "Perfect hashing, graph entropy, and circuit complexity," in *Proceedings of the Fifth Annual Structure in Complexity Theory Conference*, pp. 91–99, Barcelona, Spain, July 1990.
- [28] Z. Chen, M. Dehmer, F. Emmert-Streib, and Y. Shi, "Entropy of weighted graphs with Randić weights," *Entropy*, vol. 17, no. 6, pp. 3710–3723, 2015.
- [29] R. Kazemi, "Entropy of weighted graphs with the degree-based topological indices as weights," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 76, no. 1, pp. 69–80, 2016.
- [30] Y. C. Kwun, H. Mutee Ur Rehman, M. Yousaf, W. Nazeer, and S. M. Kang, "The entropy of weighted graphs with atomic bond connectivity edge weights," *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 8407032, 10 pages, 2018.
- [31] F. Afzal, M. Abdul Razaq, M. Abdul Razaq, D. Afzal, and S. Hameed, "Weighted entropy of penta chains graph," *Eurasian Chemical Communications*, vol. 2, no. 6, pp. 652–662, 2020.

## Research Article

# Computing Edge Weights of Magic Labeling on Rooted Products of Graphs

Jia-Bao Liu <sup>1,2</sup>, Hafiz Usman Afzal,<sup>3</sup> and Muhammad Javaid <sup>3</sup>

<sup>1</sup>School of Finance and Mathematics, Huainan Normal University, Huainan 232038, China

<sup>2</sup>School of Mathematics and Physics, Anhui Jianzhu University, Hefei 230601, China

<sup>3</sup>Department of Mathematics, School of Science, University of Management and Technology, Lahore 54770, Pakistan

Correspondence should be addressed to Muhammad Javaid; [javidmath@gmail.com](mailto:javidmath@gmail.com)

Received 10 May 2020; Revised 5 August 2020; Accepted 25 August 2020; Published 29 September 2020

Academic Editor: Alessandro Palmeri

Copyright © 2020 Jia-Bao Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Labeling of graphs with numbers is being explored nowadays due to its diverse range of applications in the fields of civil, software, electrical, and network engineering. For example, in network engineering, any systems interconnected in a network can be converted into a graph and specific numeric labels assigned to the converted graph under certain rules help us in the regulation of data traffic, connectivity, and bandwidth as well as in coding/decoding of signals. Especially, both antimagic and magic graphs serve as models for surveillance or security systems in urban planning. In 1998, Enomoto et al. introduced the notion of super  $(a, 0)$  edge-antimagic labeling of graphs. In this article, we shall compute super  $(a, 0)$  edge-antimagic labeling of the rooted product of  $P_n$  and the complete bipartite graph  $(K_{2,m})$  combined with the union of path, copies of paths, and the star. We shall also compute a super  $(a, 0)$  edge-antimagic labeling of rooted product of  $P_n$  with a special type of pancyclic graphs. The labeling provided here will also serve as super  $(a', 2)$  edge-antimagic labeling of the aforesaid graphs. All the structures discussed in this article are planar. Moreover, our findings have also been illustrated with examples and summarized in the form of a table and 3D plots.

## 1. Introduction

The antimagic and magic labelings on graphs are designed due to their wide applicability in various branches of engineering. In the literature, many results have appeared regarding numeric labelings on several operations of graphs such as graphs obtained from cartesian, corona, rooted, and strong products of various connected graphs; for instance, see [1–5]. In this article, we will provide super  $(a, 0)$  edge-antimagic labeling of the rooted product of  $P_n$  and the complete bipartite graph  $(K_{2,m})$  taking its disjoint union with path, copies of paths, and star graph. Further, we will provide super  $(a, 0)$  edge-antimagic labeling of the rooted product of path  $P_n$  with specifically designed pancyclic graphs. We shall, in particular, target the planar graphs that are obtained as a result of our rooted products. These planar graphs minimize the possibility of overlapping of various entities in practical purposes, which is a major cause of inefficiency in organizations. The super  $(a, 0)$  edge-antimagic labeling provided in this article on the specified graphs can be used as test-ready

labeling in any engineering, networking, or industrial project where the scheme of design of connections is similar to the graphs obtained in this note.

### 1.1. Applications of Graph Labeling in Engineering

**1.1.1. Software Engineering.** In software engineering, the role of graph labeling is getting improved in the encryption of the security codes in order to halt the attacks of hackers on precious data and also in coding of data to transmit it to different networks and devices alike. Similarly, test-ready labels and reference labels are improving the configurations of software in designing the updated versions. A two-scan algorithm for the connected components in binary images involves labeling and is significantly helpful in making the graphics better and clearer (see [6]). In data mining, the concept of magic labeling is becoming useful with the passage of time. It makes the data collection for deriving new information easier by indicating the data of same weightage

as one entity. In this way, magic labeling is making the task of data mining more easy, error-free, and less time- and effort-consuming in various organizations.

*1.1.2. Networking.* A network consists of nodes (vertices) and links between these nodes (edges). More technically, a network, say  $N$ , is an ordered 2-tuple consisting of two sets, set of nodes  $V(N)$  and links between nodes called set of edges  $E(N)$  such that  $(E(N) \subset V(N) \times V(N))$ . Thus, a graph is directly a representative of a network. In network engineering, the optimization and functioning of the networks are the primary hallmarks that require solid planning, construction, and management of the network at its core. Two basic types of networking are wired networking and wireless networking. One cannot deny the existence, importance, and major usage of wired networking in many principal instances. But, due to more usefulness, the apparent increase in the use of the wireless networks demands the application of robust tools, like graph labeling, to get more accuracy in the engineering of wireless networking (see [7]). We are living in an era of communication, in which radio transmission is playing an extremely important part. These wireless radio networks face a major challenge in the form of interference which makes the task of channel assignment harder. The main reason for this unwanted interruption is constraint-free transmission of the concurrent networks admitting same instance appearance [8, 9]. Such networks are first converted into graphs and then magic labeling helps in assigning constant weights to the concurrent networks. This whole procedure minimizes and even eliminates the interference in networks. Moreover, the radio labeling of graphs is tremendously helpful in the minimization of the problem of interference in wireless networks and has been playing a very vital role in the last few years. The  $(a, 0)$  edge-antimagic labeling is particularly used for automatic routing in a network. A static network is represented first as a specific graph by connecting nodes in some topology to form a connected graph and then magic labeling is applied for automatic routing of data in the network. This labeling is designed with a constant edge weight, which helps routing to automatically detect the next node within that network (see [10]).

*1.1.3. Telecommunication.* The most commercially successful application of graph labeling appears in telecom engineering these days [11]. In telecommunication, a service coverage area is divided into a quadrilateral or a hexagon, termed as a cell in cellular networking. Furthermore, each cell works as a station. The base cell has the ability to communicate with mobile stations using its radio transceiver. The challenge for base cell here is to provide maximum channel reuse without violating the constraints in order to minimize the blocking. To tackle this challenge, a label is assigned to each user and the communication link of this user receives a distinct label. In this way, the numbers assigned to any two communicating terminals automatically specify the link label of the connecting path by simply using magic, antimagic, or graceful labeling. Conversely, the path label uniquely specifies the pair of users which it interconnects (see [12]).

*1.1.4. Civil Engineering and Urban Planning.* As a particular example, consider the wheel  $W_6$ , helm  $H_6$ , and prisms  $D_5$  and  $D_6$  in Figure 1. The edges of the graphs  $W_6$ ,  $H_6$ , and  $D_6$  are labeled with consecutive labels ranging from 1 up to the size of the graph such that the labels appearing on all the vertices are distinct. That is, the vertex antimagic labeling of these graphs is being provided here. Meanwhile, on  $D_5$ , edge-magic labeling with constant edge weight 29 is provided [13, 14]. Now, for instance, in a surveillance design of highly sensitive office, the rooms can be represented by vertices and legitimate or specified passages to reach those rooms can be represented by edges. If someone tries to violate even a single legitimate passage, a complete disruption in the labeling will occur. As in case of structure like  $D_5$ , the magic constant will get disturbed abruptly due to violation of passage. This will promptly indicate to the concerned security through a computer programme that someone has violated the passage and the security team will get alert. Now, these particular magic and antimagic schemes, once designed, can be used wherever they are needed in a similar security pattern. Both magic labeling and antimagic labeling are equally important in this regard. This is a major use of magic labeling and antimagic labeling in urban planning. Thus, these labelings play their part as model for surveillance or security for various types of buildings or areas [15].

Moreover, the functioning and routing of robots installed in restaurants and in factories as production lines and other such units come under light using one of the labeling functions. As in robotics, they help to decide which function to be used and which to be skipped at a certain instance for making the robots or certain robotic components moving or keeping them static. The idea of robotic routing with the help of the tools like labeling functions and distance-based dimensions not only helps in minimizing the time for a robot for taking a decision but also maximizes its accuracy [16]. Such benefits are the cause of massive reduction of cost in industry.

We have organized this article into six main sections. Section 1 presents the introduction, followed by Section 2 of the preliminary definitions. Section 3 contains our main findings. Section 4 discusses the illustration of our findings through examples and proposed open problems. Section 5 focuses on the synopsis and 3D plots of our results. Conclusion is given in the last section.

## 2. Preliminary Definitions

In this section, we will discuss some definitions and results that are useful in the presentation of our theorems in the next sections. Also, salient works previously done in this field shall be mentioned here.

Let  $G = (V(G), E(G))$  be a simple, connected, and nonempty graph with vertex set  $V(G)$  and edge set  $E(G)$  such that  $|V(G)| = p$  and  $|E(G)| = q$ , respectively. We call  $G$  in this case a  $(p, q)$ -graph. Throughout this article, we will use  $(p, q)$ -graphs. For more insight into the graph theoretic terminologies, we refer the reader to [17].

A mapping that maps nonzero positive integers onto the vertices, edges, or both of a graph under certain conditions is called a labeling. It is called total labeling if we include both sets of vertices and edges in its domain. Some labelings carry the

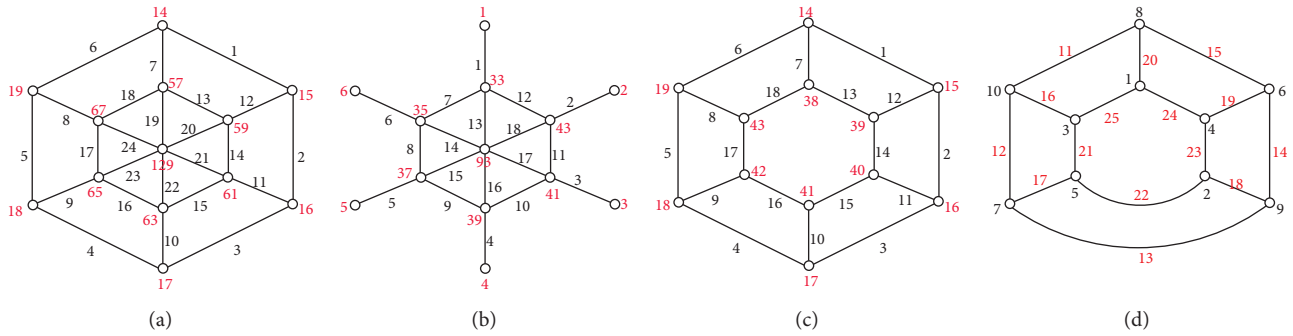


FIGURE 1: Vertex-antimagic labeling of wheel  $W_6$ , prism  $D_6$ , and helm  $H_6$  and edge-magic labeling of prism  $D_5$  [13, 14].

vertex set only or the edge set only in the domain and they are termed as vertex labelings or edge labelings, respectively. Two main types of labeling are magic labeling and antimagic labeling. In simple terms, magic labeling refers to equal vertex/edge weights and antimagic labeling refers to unequal vertex/edge weights.

**Definition 1.** For a  $(p, q)$ - graph  $G = (V(G), E(G))$ , the bijection  $f: V(G) \cup E(G) \rightarrow \{1, 2, \dots, p + q\}$  is called  $(a, d)$  edge-antimagic labeling (or  $(a, d)$  edge-antimagic total labeling) if the edge weights  $(f(x) + f(xy) + f(y))$ , for each  $xy \in E(G)$ , form a sequence of consecutive positive integers with minimum edge weight  $(a > 0)$  and common difference  $d$ . If such a labeling exists, then  $G$  is said to be an  $(a, d)$  edge-antimagic graph.

**Definition 2.** An  $(a, d)$  edge-antimagic labeling  $f$  is called a super  $(a, d)$  edge-antimagic labeling of  $(p, q)$ - graph  $G$  if  $f(V(G)) = \{1, 2, \dots, p\}$ . Thus, a super  $(a, d)$  edge-antimagic graph is one that admits a super  $(a, d)$  edge-antimagic labeling.

In the above definitions, if we have  $d = 0$ , then the minimum edge weight  $a$  becomes constant for all edges  $xy \in E(G)$  known as magic constant or magic sum of the graph  $G$ .

**Definition 3.** A simple graph  $G$  with  $|V(G)| = p$  is called a pancyclic graph if it contains cycle of every order from 3 to  $p$ .

**Definition 4.** The rooted product of two graphs  $G_1$  and  $G_2$  is obtained by taking  $|V(G_1)|$  copies of  $G_2$  and then, for every vertex  $v_i$  of  $G_1$ , identifying  $v_i$  with the root node of the  $i_{th}$  copy of  $G_2$ . It is denoted by  $(G_1 \circ G_2)$ .

In 1963, Sadláček defined the concept of magic labeling of graphs [18]. Later on, Hartsfield and Ringel [19] presented the idea of antimagic labeling for vertex-sums of a graph. The concept of  $(a, 0)$  edge-antimagic labeling of graphs was studied for the first time in [20] by Kotzig and Rosa who identified it by the name of magic valuation. In 1996, Ringel and Llado [21] studied this concept using a different terminology, that is,  $(a, 0)$  edge-antimagic labeling. Motivated by this concept, Enomoto et al., in 1998, defined the notion of super  $(a, 0)$  edge-antimagic labeling of graphs in [22]. They studied this concept with the term super edge-magic labeling of graphs. In the year 2000, Simanjantuk et al. introduced the idea of  $(a, d)$  edge-antimagic labeling of graphs in [23].

The historical background of the  $(a, 0)$  edge-antimagic labeling of graphs includes the following important and interesting conjectures on trees.

**Conjecture 1.** Every tree is  $(a, 0)$  edge-antimagic [20].

**Conjecture 2.** Every tree is super  $(a, 0)$  edge-antimagic [22].

In support of Conjecture 2, many particular classes of trees have been studied by various authors. Lee and Shah [24] verified this conjecture for trees with at most 17 vertices with the help of a programming software. In particular, the results can be found for stars, subdivided stars [25–29],  $W$ -trees [30–32], banana trees [33], caterpillars [34], subdivided caterpillars [35], and disjoint union of stars and books [36]. Further related studies can be seen in [37–39]. However, this conjecture is still open for working. In [22], it is proven that if a nontrivial  $(p, q)$ - graph  $G$  is super  $(a, 0)$  edge-antimagic, then  $(q \leq 2p - 3)$ . In the same article, the authors proved that a complete bipartite graph  $K_{m,n}$  is super  $(a, 0)$  edge-antimagic if and only if  $m = 1$  or  $n = 1$ . In [40], it is proven that  $K_{1,m} \cup K_{1,n}$  is super  $(a, 0)$  edge-antimagic if either  $m$  is a multiple of  $n + 1$  or  $n$  is a multiple of  $m + 1$ . Enomoto et al. [22] proved that  $C_n$  is super  $(a, 0)$  edge-antimagic if and only if  $n$  is odd. In [41], it has been proven that  $C_3 \cup C_n$  is super  $(a, 0)$  edge-antimagic if and only if  $n \geq 6$  and  $n$  is even (also see [42]). In [43], Figueroa-Centeno et al. showed that the generalized prism  $C_n \times P_m$  is super  $(a, 0)$  edge-antimagic for every odd integer  $n$ . In [3], Baig et al. presented the super  $(a, 0)$  edge-antimagic labeling of a class of pancyclic graphs. The following lemma regarding super  $(a, 0)$  edge-antimagic graphs is very useful.

**Lemma 1** (see [43]). A  $(p, q)$ -graph  $G$  is super  $(a, 0)$  edge-antimagic total if and only if there exists a bijective function  $f: V(G) \rightarrow \{1, 2, \dots, p\}$  such that the set  $S = \{f(x) + f(y) \mid xy \in E(G)\}$  consists of  $q$  consecutive integers. In such a case,  $f$  extends to a super  $(a, 0)$  edge-antimagic total labeling of  $G$  with magic constant  $(c = p + q + s)$ , where  $s = \min(S)$  and  $S = \{c - (p + 1), c - (p + 2), \dots, c - (p + q)\}$ .

The sum  $(f(x) + f(y))$  in Lemma 1 is termed as edge sum for each edge  $(xy \in E(G))$ . In fact, the set of all edge sums  $S = \{f(x) + f(y) \mid xy \in E(G)\}$  in Lemma 1 constituting an arithmetic progression  $\{c - (p + 1), c - (p + 2), \dots, c - (p + q)\}$  forms a super  $(a, 1)$  edge-



antimagic vertex labeling of the graph  $G$  [23]. That is,  $f$  constitutes a super  $(a, 1)$  edge-antimagic vertex labeling of  $G$ . It means Lemma 1 states that whenever the set of edge sums forms a super  $(a, 1)$  edge-antimagic vertex labeling of  $G$ , it extends to a super  $(a, 0)$  edge-antimagic labeling of  $G$ .

We will frequently use the above lemma in the proofs of our results. Another very useful relevant result is as follows.

**Theorem 1** (see [45]). *If a  $(p, q)$ -graph  $G$  is super  $(a, 0)$  edge-antimagic, then it is super  $(a - q + 1, 2)$  edge-antimagic always.*

### 3. Main Results

This section consists of two further sections, in which we will present our main results. In Section 3.1, we will study the super  $(a, 0)$  edge-antimagic labeling of the disjoint union of the rooted product of  $P_n$  and the complete bipartite graph  $K_{2,m}$  with path, copies of paths, and the star. Meanwhile, in Section 3.2, we will provide the super  $(a, 0)$  edge-antimagic labeling of rooted product of  $P_n$  with certain pancyclic graphs  $H_1$  and  $H_2$ . It is pertinent to mention here that all our graphs obtained as the result of the rooted products are planar.

*3.1. Super  $(a, 0)$  Edge-Antimagic Labeling of the Disjoint Union of the Rooted Product of  $P_n$  and  $K_{2,m}$  with Path, Copies of Paths, and Star.* The following open problem proposed by Ngurah et al. in [46] is our main motivation to study super

$(a, 0)$  edge-antimagic labeling of the graph containing copies of complete bipartite graph  $K_{2,m}$ .

*3.1.1. Open Problem.* For  $n \geq 2$  and  $m \geq 3$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of  $(nP_n \circ K_{2,m})$ ?

The rooted product  $(P_n \circ K_{2,m})$ , in fact, contains  $n$  copies of the complete bipartite graph  $K_{2,m}$ . We swiftly move to our theorems now.

**Theorem 2.** *For even  $(m \geq 2)$  and odd  $(n \geq 3)$ , the graph  $(P_n \circ K_{2,m}) \cup K_{1,n} \cup (n - 1/2)K_1$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $(3mn + 8n + 2)$ .*

*Proof.* Consider the graph  $G_1 \cong (P_n \circ K_{2,m}) \cup K_{1,n} \cup (n - 1/2)K_1$  with vertex and edge sets as follows:

$$\begin{aligned} V(G_1) &= \{x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{y_i, z_i: 1 \leq i \leq n\} \\ &\cup \{p_i: 1 \leq i \leq n\} \cup \left\{c_i: 1 \leq i \leq \frac{n-1}{2}\right\} \cup \{c\}, \\ E(G_1) &= \{y_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\ &\cup \{z_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\ &\cup \{c p_i: 1 \leq i \leq n\} \cup \{y_i y_{i+1}: 1 \leq i \leq n-1\}, \end{aligned} \quad (1)$$

where  $|V(G_1)| = ((7n + 2mn + 1)/2)$  and  $(|E(G_1)| = 2n(m + 1) - 1)$ . We define a labeling  $(f_1: V(G_1) \rightarrow \{1, 2, \dots, ((7n + 2mn + 1)/2)\})$  as follows:

$$\begin{aligned} f_1(x_j^i) &= \begin{cases} \frac{5n+i}{2}: & j = \frac{m}{2}, 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{4n+i}{2}: & j = \frac{m}{2}, 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \\ n(m+2) - (i-1): & j = \frac{m}{2} + 1, 1 \leq i \leq n, \\ n(j+3) - (i-1): & 1 \leq i \leq n, 1 \leq j \leq \frac{m}{2} - 1, \\ nj - (i-1) + n: & 1 \leq i \leq n, \frac{m}{2} + 2 \leq j \leq m, \end{cases} \\ f_1(y_i) &= \begin{cases} \frac{i+2n+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{i+3n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\ f_1(z_i) &= \begin{cases} n(m+1) + \frac{i+2n+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m+1) + \frac{i+3n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}. \end{cases} \\ f_1(p_i) &= i: 1 \leq i \leq n, \\ f_1(c) &= \frac{7n+2mn+1}{2}, \\ f_1(c_i) &= mn + 3n + i: 1 \leq i \leq \frac{n-1}{2}. \end{aligned} \quad (2)$$

The set of all edge sums generated by the labeling scheme  $f_1$  forms a sequence of consecutive integers  $((5n+5/2), (5n+7/2), \dots, ((4mn+9n+1)/2))$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $G_1$ . Therefore, by Lemma 1,  $f_1$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $G_1$  with magic constant  $(a = 3mn + 8n + 2)$ .

**Theorem 3.** For odd  $m, n \geq 3$ , the graph  $(P_n \circ K_{2,m}) \cup K_{1,n} \cup (n-1/2)K_1$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $(3mn + 8n + 2)$ .

*Proof.* Consider the graph  $(G_2 \cong (P_n \circ K_{2,m}) \cup K_{1,n} \cup (n-1/2)K_1)$  with odd  $m$  as follows:

$$\begin{aligned}
 f_2(x_j^i) &= \begin{cases} \frac{5n+i}{2}: & j = \frac{m+1}{2}, 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{4n+i}{2}: & j = \frac{m+1}{2}, 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \\ n(m+2) - (i-1): & j = \frac{m+3}{2}, 1 \leq i \leq n, \\ n(j+3) - (i-1): & 1 \leq i \leq n, 1 \leq j \leq \frac{m}{2} - 1, \\ nj - (i-1) + n: & 1 \leq i \leq n, \frac{m+5}{2} \leq j \leq m, \end{cases} \\
 f_2(y_i) &= \begin{cases} \frac{i+2n+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{i+3n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_2(z_i) &= \begin{cases} n(m+1) + \frac{i+2n+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m+1) + \frac{i+3n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_2(p_i) &= i: 1 \leq i \leq n, \\
 f_2(c) &= \frac{7n+2mn+1}{2}, \\
 f_2(c_i) &= mn + 3n + i: 1 \leq i \leq \frac{n-1}{2}.
 \end{aligned} \tag{4}$$

The set of all edge sums generated by the above labeling scheme  $f_2$  forms a sequence of consecutive integers  $((5n+5/2), (5n+7/2), \dots, ((4mn+9n+1)/2))$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $G_2$ . Therefore, by Lemma 1,  $f_2$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $G_2$  with magic constant  $(a = 3mn + 8n + 2)$ .

**Theorem 4.** For even  $m \geq 2$  and odd  $n \geq 3$ , the graph  $((P_n \circ K_{2,m}) \cup nP_2)$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $((6mn + 17n + 3)/2)$ .

$$\begin{aligned}
 V(G_2) &= \{x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{y_i, z_i: 1 \leq i \leq n\} \\
 &\cup \{p_i: 1 \leq i \leq n\} \cup \left\{c_i: 1 \leq i \leq \frac{n-1}{2}\right\} \cup \{c\}, \\
 E(G_2) &= \{y_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\
 &\cup \{z_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{cp_i: 1 \leq i \leq n\} \\
 &\cup \{y_i y_{i+1}: 1 \leq i \leq n-1\},
 \end{aligned} \tag{3}$$

where  $|V(G_2)| = ((7n + 2mn + 1)/2)$  and  $(|E(G_2)| = 2n(m+1) - 1)$ . We define a labeling  $(f_2: V(G_2) \rightarrow \{1, 2, \dots, ((7n + 2mn + 1)/2)\})$  as follows:

*Proof.* Consider the graph  $(G_3 \cong ((P_n \circ K_{2,m}) \cup nP_2))$ , for odd  $n \geq 3$ , constructed as

$$\begin{aligned}
 V(G_3) &= \{x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{y_i, z_i: 1 \leq i \leq n\} \\
 &\cup \{p_i, q_i: 1 \leq i \leq n\}, \\
 E(G_3) &= \{y_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\
 &\cup \{z_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{p_i q_i: 1 \leq i \leq n\} \\
 &\cup \{y_i y_{i+1}: 1 \leq i \leq n-1\},
 \end{aligned} \tag{5}$$



where we have  $(|V(G_3)| = n(m + 4))$  and  $(|E(G_3)| = 2n(m + 1) - 1)$ . Now, we define a labeling  $(f_3: V(G_3) \rightarrow \{1, 2, \dots, n(m + 4)\})$  as follows:

$$\begin{aligned}
 f_3(x_j^i) &= \begin{cases} \frac{1}{2}(5n + i): & j = \frac{m}{2}, 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{1}{2}(4n + i): & j = \frac{m}{2}, 2 \leq i \leq n - 1, i \equiv 0 \pmod{2}, \\ n(m + 2) - (i - 1): & j = \frac{m}{2} + 1, 1 \leq i \leq n, \\ n(j + 3) - (i - 1): & 1 \leq i \leq n, 1 \leq j \leq \frac{m}{2} - 1, \\ nj - (i - 1) + n: & 1 \leq i \leq n, \frac{m}{2} + 2 \leq j \leq m, \end{cases} \\
 f_3(y_i) &= \begin{cases} \frac{i + 2n + 1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{i + 3n + 1}{2}: & 2 \leq i \leq n - 1, i \equiv 0 \pmod{2}, \end{cases} \tag{6} \\
 f_3(z_i) &= \begin{cases} n(m + 1) + \frac{i + 2n + 1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m + 1) + \frac{i + 3n + 1}{2}: & 2 \leq i \leq n - 1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_3(q_i) &= \begin{cases} n(m + 1) + \frac{5n - i + 2}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m + 1) + \frac{6n - i + 2}{2}: & 2 \leq i \leq n - 1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_3(p_i) &= i: 1 \leq i \leq n.
 \end{aligned}$$

The set of all edge sums generated by the above labeling scheme  $f_3$  constitutes a sequence of consecutive integers  $((5n + 5/2), (5n + 7/2), \dots, ((4mn + 9n + 1)/2))$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $G_3$ . Therefore, by Lemma 1,  $f_3$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $G_3$  with magic constant  $a = ((6mn + 17n + 3)/2)$ .

**Theorem 5.** For odd  $(m, n \geq 3)$ , the graph  $((P_n \circ K_{2,m}) \cup nP_2)$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $((6mn + 17n + 3)/2)$ .

*Proof.* Consider the graph  $G_4 \cong ((P_n \circ K_{2,m}) \cup nP_2)$ , for odd  $m$ , with the construction:

$$\begin{aligned}
 V(G_4) &= \{x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{y_i, z_i: 1 \leq i \leq n\} \\
 &\quad \cup \{p_i, q_i: 1 \leq i \leq n\}, \\
 E(G_4) &= \{y_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \tag{7} \\
 &\quad \cup \{z_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{p_i, q_i: 1 \leq i \leq n\} \\
 &\quad \cup \{y_i y_{i+1}: 1 \leq i \leq n - 1\},
 \end{aligned}$$

where  $(|V(G_4)| = n(m + 4))$  and  $(|E(G_4)| = 2n(m + 1) - 1)$ . Now, consider a labeling  $(f_4: V(G_4) \rightarrow \{1, 2, \dots, n(m + 4)\})$  as follows:

$$\begin{aligned}
 f_4(x_j) &= \begin{cases} \frac{1}{2}(5n+i): & j = \frac{m+1}{2}, 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{1}{2}(4n+i): & j = \frac{m+1}{2}, 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \\ n(m+2) - (i-1): & j = \frac{m+3}{2}, 1 \leq i \leq n, \\ n(j+3) - (i-1): & 1 \leq i \leq n, 1 \leq j \leq \frac{m-1}{2}, \\ nj - (i-1) + n: & 1 \leq i \leq n, \frac{m+5}{2} \leq j \leq m, \end{cases} \\
 f_4(y_i) &= \begin{cases} \frac{i+2n+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{i+3n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_4(z_i) &= \begin{cases} n(m+1) + \frac{i+2n+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m+1) + \frac{i+3n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_4(q_i) &= \begin{cases} n(m+1) + \frac{5n-i+2}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m+1) + \frac{6n-i+2}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_4(p_i) &= i: 1 \leq i \leq n.
 \end{aligned} \tag{8}$$

The set of all edge sums generated by the above labeling scheme  $f_4$  constitutes a sequence of consecutive integers  $((5n+5/2), (5n+7/2), \dots, ((4mn+9n+1)/2))$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $G_4$ . Therefore, by Lemma 1,  $f_4$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $G_4$  with magic constant  $a = ((6mn+17n+3)/2)$ .

**Theorem 6.** For even  $m \geq 2$  and odd  $n \geq 3$ , the graph  $(P_n \circ K_{2,m}) \cup P_{n+1}$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $((6mn+13n+7)/2)$ .

*Proof.* Consider the graph  $G_5 \cong (P_n \circ K_{2,m}) \cup P_{n+1}$ , for odd  $n \geq 3$ , with the construction:

$$\begin{aligned}
 V(G_5) &= \{x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\
 &\cup \{y_i, z_i: 1 \leq i \leq n\} \\
 &\cup \{p_i: 1 \leq i \leq n+1\}, \\
 E(G_5) &= \{y_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{z_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\
 &\cup \{p_i p_{i+1}: 1 \leq i \leq n\} \cup \{y_i y_{i+1}: 1 \leq i \leq n-1\},
 \end{aligned} \tag{9}$$

where we have  $(|V(G_5)| = mn + 3n + 1)$  and  $(|E(G_5)| = 2n(m+1) - 1)$ . We define a labeling  $(f_5: V(G_5) \rightarrow \{1, 2, \dots, mn + 3n + 1\})$  as follows:

$$\begin{aligned}
f_5(x_j^i) &= \begin{cases} \frac{1}{2}(4n+i+1): & j = \frac{m}{2}, 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{1}{2}(3n+i+1): & j = \frac{m}{2}, 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \\ \frac{1}{2}(2mn+3n-2i+3): & j = \frac{m}{2}+1, 1 \leq i \leq n, \\ \frac{1}{2}(2nj+5n-2i+3): & 1 \leq i \leq n, 1 \leq j \leq \frac{m}{2}-1, \\ \frac{1}{2}(2nj+n-2i+3): & 1 \leq i \leq n, \frac{m}{2}+2 \leq j \leq m, \end{cases} \\
f_5(y_i) &= \begin{cases} \frac{n+i+2}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{2n+i+2}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \tag{10} \\
f_5(z_i) &= \begin{cases} n(m+1) + \frac{n+i+2}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m+1) + \frac{2n+i+2}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
f_5(p_i) &= \begin{cases} \frac{i+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{1}{2}(2mn+5n+i+1): & 2 \leq i \leq n+1, i \equiv 0 \pmod{2}. \end{cases}
\end{aligned}$$

The set of all edge sums generated by the above labeling scheme  $f_5$  forms a sequence of consecutive integers  $((3n+7)/2), (3n+9)/2, \dots, ((4mn+7n+3)/2)$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $G_5$ . Therefore, by Lemma 1,  $f_5$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $G_5$  with magic constant  $a = ((6mn+13n+7)/2)$ .

**Theorem 7.** For odd  $m, n \geq 3$ , the graph  $(P_n \circ K_{2,m}) \cup P_{n+1}$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $((6mn+13n+7)/2)$ .

*Proof.* Consider the graph  $(G_6 \cong (P_n \circ K_{2,m}) \cup P_{n+1})$  for both odd  $m$  and  $n \geq 3$  with vertex and edge sets as follows.

$$\begin{aligned}
V(G_6) &= \{x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \cup \{y_i, z_i: 1 \leq i \leq n\} \\
&\quad \cup \{p_i: 1 \leq i \leq n+1\}, \\
E(G_6) &= \{y_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\
&\quad \cup \{z_i x_j^i: 1 \leq i \leq n, 1 \leq j \leq m\} \\
&\quad \cup \{p_i p_{i+1}: 1 \leq i \leq n\} \cup \{y_i y_{i+1}: 1 \leq i \leq n-1\}, \tag{11}
\end{aligned}$$

where  $(|V(G_6)| = mn + 3n + 1)$  and  $(|E(G_6)| = 2n(m+1) - 1)$ . Now, we define a labeling  $(f_6: V(G_6) \rightarrow \{1, 2, \dots, mn + 3n + 1\})$  as follows:

$$\begin{aligned}
 f_6(x_j) &= \begin{cases} \frac{1}{2}(4n+i+1): & j = \frac{m+1}{2}, 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{1}{2}(3n+i+1): & j = \frac{m+1}{2}, 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \\ \frac{1}{2}(2mn+3n-2i+3): & j = \frac{m+3}{2}, 1 \leq i \leq n, \\ \frac{1}{2}(2nj+5n-2i+3): & 1 \leq i \leq n, 1 \leq j \leq \frac{m-1}{2}, \\ \frac{1}{2}(2nj+n-2i+3): & 1 \leq i \leq n, \frac{m+5}{2} \leq j \leq m, \end{cases} \\
 f_6(y_i) &= \begin{cases} \frac{n+i+2}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{2n+i+2}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_6(z_i) &= \begin{cases} n(m+1) + \frac{n+i+2}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ n(m+1) + \frac{2n+i+2}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases} \\
 f_6(p_i) &= \begin{cases} \frac{i+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{1}{2}(2mn+5n+i+1): & 2 \leq i \leq n+1, i \equiv 0 \pmod{2}. \end{cases}
 \end{aligned} \tag{12}$$

The set of all edge sums generated by the above labeling scheme  $f_6$  forms a sequence of consecutive integers  $((3n+7/2), (3n+9/2), \dots, ((4mn+7n+3)/2))$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $G_6$ . Therefore, by Lemma 1,  $f_6$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $G_6$  with magic constant  $a = ((6mn+13n+7)/2)$ .

### 3.1.2. Observations

- (1) In Theorems 2, 4, and 6, the labels  $(f_k(x_j^i))$  for  $(1 \leq j \leq (m/2) - 1) \& ((m/2) + 2 \leq j \leq m)$  will receive the values only when  $m > 2$ . Similarly, in Theorems 3, 5, and 7, the labels  $(f_k(x_j^i))$  for  $(m+5/2) \leq j \leq m$  will receive the values when  $m > 3$ . This happens because  $K_{2,2} \cong C_4$  and  $K_{2,3}$  has 3 vertices in one partitioned set. However, the labeling of our graphs  $G_1, \dots, G_6$  never gets disturbed anyway (here,  $k \in \{1, \dots, 6\}$ ).
- (2) The following results from Theorems 2–7 are direct consequences of Theorem 1.
- (3) As a cycle  $C_n$  is super  $(a, 0)$  edge-antimagic if and only if  $n$  is odd [22], this means that  $C_4$  is not super  $(a, 0)$  edge-antimagic. We have observed an interesting substructure in Theorems 2, 4, and 6. That is, if we fix  $m = 2$  in these theorems, we obtain a cyclic

family of graphs, in which only cycle  $C_4$  appears  $n$  times ( $n$  is odd). Obviously, these families are also super  $(a, 0)$  edge-antimagic, as per the proofs of our results. See super  $(a, 0)$  edge-antimagic labeling of  $((P_5 \circ C_4) \cup K_{1,5} \cup 2K_1)$ ,  $((P_5 \circ C_4) \cup 5P_2)$ , and  $((P_5 \circ C_4) \cup P_6)$  in Figure 2.

**Theorem 8.** For all  $m \geq 2$  and odd  $n \geq 3$ , the graph  $((P_n \circ K_{2,m}) \cup K_{1,n} \cup (n-1/2)K_1)$  admits a super  $(mn+6n+4, 2)$  edge-antimagic labeling.

**Theorem 9.** For all  $m \geq 2$  and odd  $n \geq 3$ , the graph  $((P_n \circ K_{2,m}) \cup nP_2)$  admits a super  $((2mn+13n+7)/2, 2)$  edge-antimagic labeling.

**Theorem 10.** For all  $m \geq 2$  and odd  $n \geq 3$ , the graph  $((P_n \circ K_{2,m}) \cup P_{n+1})$  admits a super  $((2mn+9n+11)/2, 2)$  edge-antimagic labeling.

**3.2. Super  $(a, 0)$  Edge-Antimagic Labeling of Rooted Product of  $P_n$  with Certain Pancyclic Graphs.** In networking, the networks that are acyclic and networks containing a range of cycles admit same level of importance. In graph theoretic terms, the former corresponds to trees and the latter corresponds to multicyclic graphs. What about a network that

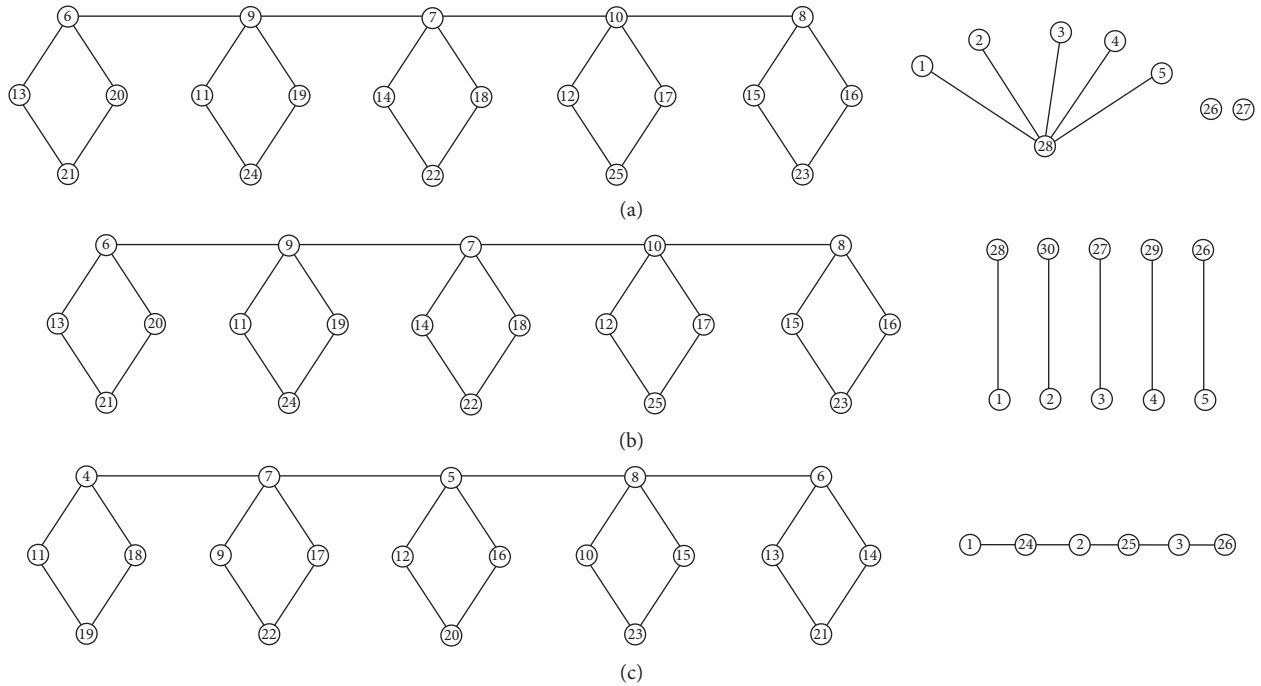


FIGURE 2: A super  $(a, 0)$  edge-antimagic labeling of the cyclic graphs (a)  $((P_5 \circ C_4) \cup K_{1,5} \cup 2K_1)$ , (b)  $((P_5 \circ C_4) \cup 5P_2)$ , and (c)  $((P_5 \circ C_4) \cup P_6)$  with magic constants 72, 74, and 66, respectively.

contains cyclic connection of all orders starting from 3 up to the number of systems connected within that network? In such situations, the task of the programmers gets more tough, as they have to work on encrypting powerful codes to keep their data safe by halting the attacks of hackers. It is because every system in such network is connected within a cycle. This kind of network corresponds to a pancyclic graph (see Definition 3).

We define here a specific pancyclic graph  $H_1$  as follows.

**Definition 5.** Consider a pancyclic graph  $H_1$  with the construction:

$$\begin{aligned} V(H_1) &= \{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{y, z\}, \\ E(H_1) &= \{x_1x_3, x_3x_5, x_2x_4, x_4x_6, x_1x_2, x_3x_4, x_5x_6, x_2x_3, x_4x_5\} \\ &\quad \cup \{yx_1, yx_2, zx_5, zx_6\}. \end{aligned} \tag{13}$$

In the next result, we will show that the rooted product  $(P_n \circ H_1)$  is super  $(a, 0)$  edge-antimagic for odd values of  $n$ .

**Theorem 11.** For odd  $n$ , the rooted product  $(P_n \circ H_1)$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $(a = (45n + 3/2))$ .

*Proof*

- (1) For  $n = 1$ ,  $(P_1 \circ H_1 \cong H_1)$ . The vertex labeling  $\{x_1, x_2, x_3, x_4, x_5, x_6, y, z: 1, 4, 3, 6, 5, 8, 2, 7\}$  extends to a super  $(24, 0)$  edge-antimagic labeling of  $(P_1 \circ H_1)$  by Lemma 1.
- (2) For  $n \geq 3$ , consider the graph  $(P_n \circ H_1)$  with  $(|V(P_n \circ H_1)| = 8n)$  and  $(|E(P_n \circ H_1)| = 14n - 1)$  consisting of the following vertex and edge sets:

$$\begin{aligned} V(P_n \circ H_1) &= \{x_i, v_i: 1 \leq i \leq n\} \cup \{y_i, z_i, w_i: 1 \leq i \leq 2n\}, \\ E(P_n \circ H_1) &= \{y_iy_{i+1}, z_iz_{i+1}, w_iw_{i+1}: 1 \leq i \leq 2n - 1, i \equiv 1 \pmod{2}\} \\ &\quad \cup \{z_iz_{i+1}, w_iz_{i+1}: 1 \leq i \leq 2n - 1, i \equiv 1 \pmod{2}\} \\ &\quad \cup \{x_iy_{2i-1}, x_iy_{2i}, v_iw_{2i-1}, v_iw_{2i}: 1 \leq i \leq n\} \\ &\quad \cup \{x_ix_{i+1}: 1 \leq i \leq n - 1\} \cup \{y_iz_i, z_iw_i: 1 \leq i \leq 2n\}. \end{aligned} \tag{14}$$

Consider a labeling  $(\psi_1: V(P_n \circ H_1) \longrightarrow \{1, 2, \dots, |V(P_n \circ H_1)| = 8n\})$  defined as

$$\psi_1(x_i) = \begin{cases} \frac{i+1}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{i+n+1}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}, \end{cases}$$

$$\psi_1(y_i) = \begin{cases} \frac{1}{2}(6n-i+2): & 2 \leq i \leq 2n, i \equiv 0 \pmod{2}, \\ \frac{1}{4}(6n+i+1): & 1 \leq i \leq 2n-1, i \equiv 1 \pmod{4}, \\ \frac{1}{4}(4n+i+1): & 3 \leq i \leq 2n-3, i \equiv 3 \pmod{4}, \\ \frac{1}{4}(16n+i): & 4 \leq i \leq 2n-2, i \equiv 0 \pmod{4}, \\ \frac{1}{4}(12n+i+3): & 1 \leq i \leq 2n-1, i \equiv 1 \pmod{4}, \\ \frac{1}{4}(18n+i): & 2 \leq i \leq 2n, i \equiv 2 \pmod{4}, \\ \frac{1}{4}(14n+i+3): & 3 \leq i \leq 2n-3, i \equiv 3 \pmod{4}, \end{cases}$$

$$\psi_1(z_i) = \begin{cases} \frac{1}{2}(12n-i+1): & 1 \leq i \leq 2n-1, i \equiv 1 \pmod{2}, \\ \frac{1}{4}(26n+i+2): & 4 \leq i \leq 2n-2, i \equiv 0 \pmod{4}, \\ \frac{1}{4}(24n+i+2): & 2 \leq i \leq 2n, i \equiv 2 \pmod{4}, \end{cases}$$

$$\psi_1(w_i) = \begin{cases} \frac{15n+i}{2}: & 1 \leq i \leq n, i \equiv 1 \pmod{2}, \\ \frac{14n+i}{2}: & 2 \leq i \leq n-1, i \equiv 0 \pmod{2}. \end{cases} \tag{15}$$

The set of all edge sums generated by the above labeling scheme  $\psi_1$  forms a sequence of consecutive integers  $((n+5/2), (n+7/2), \dots, (29n+1/2))$ , which is a super  $(a, 1)$  edge-antimagic vertex labeling of  $(P_n \circ H_1)$ . Therefore, by Lemma 1,  $\psi_1$  extends to a super  $(a, 0)$  edge-antimagic labeling of the graph  $(P_n \circ H_1)$  with magic constant  $(45n+3/2)$ .

*Definition 6.* Consider a pancyclic graph  $H_2$  (non-isomorphic to  $H_1$ ) with the vertex-edge connection as follows:

$$V(H_2) = \{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{y, z\},$$

$$E(H_2) = \{x_1x_3, x_3x_5, x_2x_4, x_4x_6, x_1x_2, x_5x_6, x_2x_3, x_4x_5\} \cup \{yx_1, yx_2, zx_5, zx_6, yz\}. \tag{16}$$

**Theorem 12.** For odd  $n$ , the rooted product  $(P_n \circ H_2)$  admits a super  $(a, 0)$  edge-antimagic labeling with magic constant  $(a = (45n+3/2))$ .

*Proof.* Consider the rooted product  $(P_n \circ H_2)$  with  $(|V(P_n \circ H_2)| = 8n)$  and  $(|E(P_n \circ H_2)| = 14n-1)$  with vertex and edge sets as follows:

$$V(P_n \circ H_2) = \{x_i, v_i: 1 \leq i \leq n\} \cup \{y_i, z_i, w_i: 1 \leq i \leq 2n\},$$

$$E(P_n \circ H_2) = \{y_iy_{i+1}, z_iz_{i+1}, w_iw_{i+1}: 1 \leq i \leq 2n-1, i \equiv 1 \pmod{2}\} \cup \{z_iz_{i+1}, w_iz_{i+1}: 1 \leq i \leq 2n-1, i \equiv 1 \pmod{2}\} \cup \{x_iy_{2i-1}, x_iy_{2i}, v_iw_{2i-1}, v_iw_{2i}: 1 \leq i \leq n\} \cup \{x_ix_{i+1}: 1 \leq i \leq n-1\} \cup \{y_iz_i, z_iz_i: 1 \leq i \leq 2n\}. \tag{17}$$

And the labeling scheme is the same as  $\psi_1$  designed in Theorem 11.

The following result is a direct consequence of Theorem 1.

**Theorem 13.** For odd  $n$ , the rooted products  $(P_n \circ H_1)$  and  $(P_n \circ H_2)$  are super  $((17n+7/2), 2)$  edge-antimagic.

### 4. Illustration through Examples and Proposed Open Problems

*4.1. Examples.* As examples of Theorems 2 and 3, the super  $(a, 0)$  edge-antimagic labeling of  $((P_5 \circ K_{2,6}) \cup K_{1,5} \cup 2K_1)$  and super  $(a, 0)$  edge-antimagic labeling of  $((P_5 \circ K_{2,7}) \cup K_{1,5} \cup 2K_1)$  are presented in Figures 3 and 4, respectively. It is evident that the magic constant of the labeling in Figure 3 is 132 (corresponding to the parameters  $n = 5$  and  $m = 6$ ) and the magic constant of the labeling in Figure 4 is 147 (corresponding to the parameters  $n = 5$  and  $m = 7$ ). These are as per the magic constants depicted in the proofs of Theorems 2 and 3.

Similarly, Figures 5–8 illustrate Theorems 4–7, respectively, for fixed parameters mentioned in the caption of each. Here, the magic constants are as per our depiction as well.

Figures 9 and 10 illustrate super  $(114, 0)$  edge-antimagic labeling of the rooted products appearing in Theorems 11 and 12 with parameter  $n = 5$ .

- (1) In all of the above figures, Lemma 1 has helped us in the way that we do not require to label the edges. We only have observed that the edge sums are consecutive positive integers. So, applying the remaining labels  $\{p+1, p+2, \dots, q\}$  to the edges of the graph in descending or ascending sequence of edge sums will give rise to super  $(a, 0)$  or  $(a', 2)$  edge-antimagic labeling of the graph, respectively, for appropriate values of  $a$  and  $a'$ .



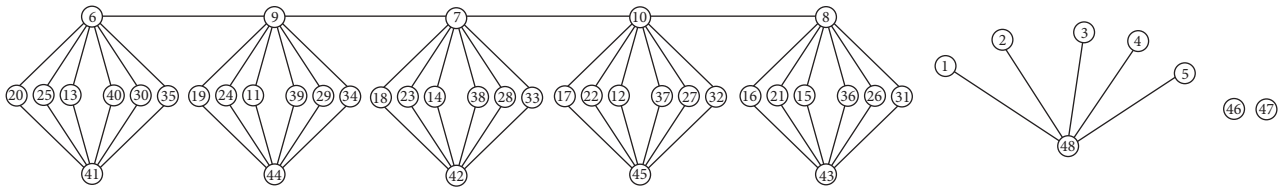


FIGURE 3: A super (132, 0) edge-antimagic labeling of  $((P_5 \circ K_{2,6}) \cup K_{1,5} \cup 2K_1)$ .

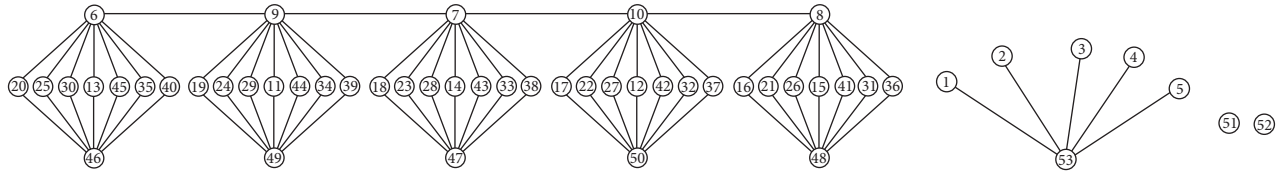


FIGURE 4: A super (147, 0) edge-antimagic labeling of  $((P_5 \circ K_{2,6}) \cup K_{1,5} \cup 2K_1)$ .

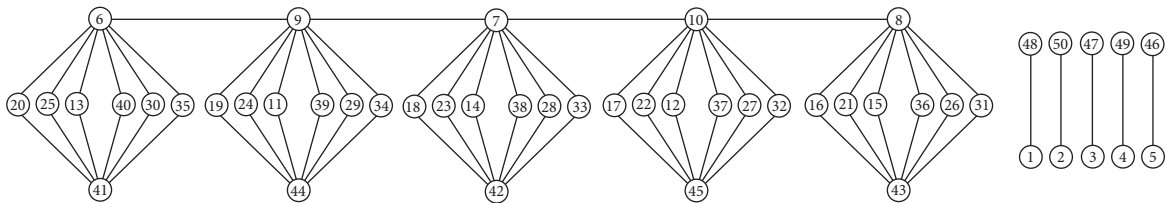


FIGURE 5: A super (134, 0) edge-antimagic labeling of  $((P_5 \circ K_{2,6}) \cup 5P_2)$ .

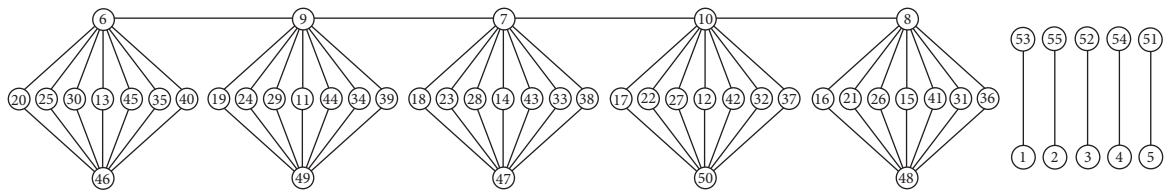


FIGURE 6: A super (149, 0) edge-antimagic labeling of  $((P_5 \circ K_{2,7}) \cup 5P_2)$ .

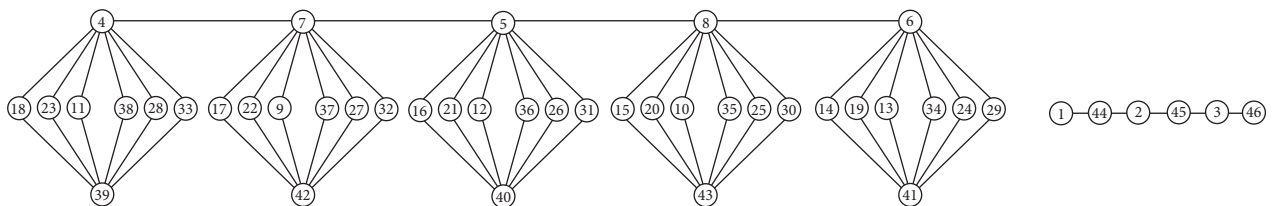


FIGURE 7: A super (126, 0) edge-antimagic labeling of  $((P_5 \circ K_{2,6}) \cup P_6)$ .

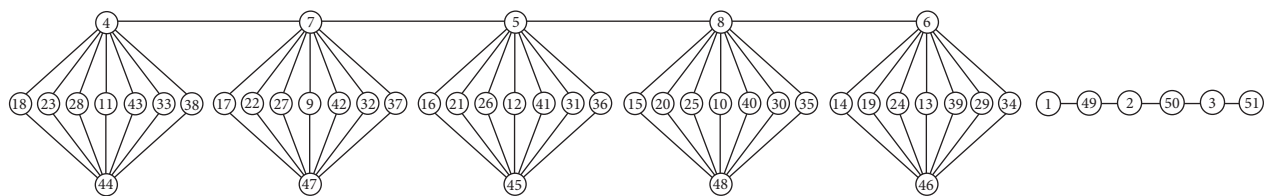


FIGURE 8: A super (141, 0) edge-antimagic labeling of  $((P_5 \circ K_{2,7}) \cup P_6)$ .

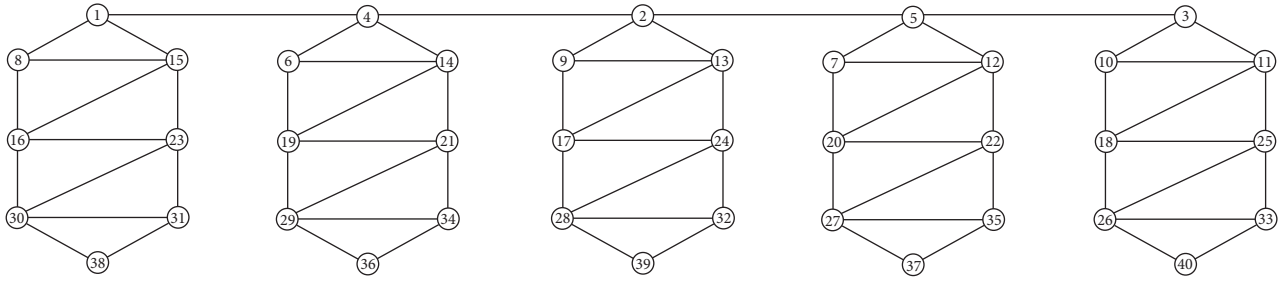


FIGURE 9: A super  $(a, 0)$  edge-antimagic labeling of  $(P_5 \circ H_1)$  with magic constant  $a = 114$ .

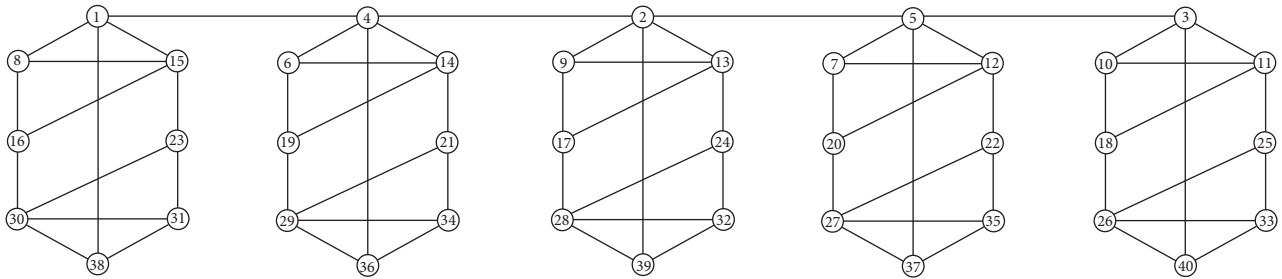


FIGURE 10: A super  $(a, 0)$  edge-antimagic labeling of  $P_5 \circ H_2$  with magic constant  $a = 114$ .

TABLE 1: Summary of the main findings.

Graph	Parameters	Magic constant ( $a(d = 0)$ )	Minimum edge weight ( $a'(d = 2)$ )	Structure
$(P_n \circ K_{2,m}) \cup K_{1,n} \cup (n - 1/2)K_1$	$(m \geq 2, 3 \leq n \equiv 1 \pmod{2})$	$(3mn + 8n + 2)$	$(mn + 6n + 4)$	Planar
$((P_n \circ K_{2,m}) \cup nP_2)$	$(m \geq 2, 3 \leq n \equiv 1 \pmod{2})$	$((6mn + 17n + 3)/2)$	$((2mn + 13n + 7)/2)$	Planar
$((P_n \circ K_{2,m}) \cup P_{n+1})$	$(m \geq 2, 3 \leq n \equiv 1 \pmod{2})$	$((6mn + 13n + 7)/2)$	$((2mn + 9n + 11)/2)$	Planar
$(P_n \circ H_1)$	$(n \equiv 1 \pmod{2})$	$((45n + 3)/2)$	$((17n + 7)/2)$	Planar
$(P_n \circ H_2)$	$(n \equiv 1 \pmod{2})$	$((45n + 3)/2)$	$((17n + 7)/2)$	Planar

4.2. Open Problems. Here, we propose few open problems related to the results (Theorems 2–7) presented in Section 3.1 as follows:

- (1) For even  $n \geq 2$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of  $((P_n \circ K_{2,m}) \cup K_{1,n} \cup (n - 1/2)K_1)$ ?
- (2) For even  $n \geq 2$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of  $((P_n \circ K_{2,m}) \cup nP_2)$ ?
- (3) For even  $n \geq 2$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of  $((P_n \circ K_{2,m}) \cup P_{n+1})$ ?
- (4) For odd  $n \geq 3$ , can you determine super  $(a, 0)$  edge-antimagic labeling of the graphs  $G_1, G_2, G_3, G_4, G_5$ , and  $G_6$  for a magic constant (i.e., for any other value of  $a$ ) other than those obtained in this article?
- (5) For  $l, m, n \in \mathbb{N}$ , find any super  $(a, 0)$  edge-antimagic labeling of the following graphs:
  - (1)  $((P_n \circ K_{2,m}) \cup K_{1,l})$
  - (2)  $((P_n \circ K_{2,m}) \cup lP_2)$
  - (3)  $((P_n \circ K_{2,m}) \cup P_l)$

The open problems related to the results (Theorems 11 and 12) presented in Section 3.2 are as follows:

- (1) For even  $n \geq 2$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of the rooted product  $(P_n \circ H_1)$ ?
- (2) For even  $n \geq 2$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of the rooted product  $(P_n \circ H_2)$ ?
- (3) For odd  $n \geq 3$ , can you determine any super  $(a, 0)$  edge-antimagic labeling of the rooted products  $(P_n \circ H_1)$  and  $(P_n \circ H_2)$  with a magic constant that is different from those obtained here, that is, for any other value of  $a$

### 5. Synopsis and 3D Graphical Trends and Relative Comparison of the Magic Constants

5.1. Synopsis. The computational results of our findings have been summarized in Table 1. The ‘‘Parameters’’ column represents those parameters for which we have been able to find super  $(a, 0)$  and  $(a', 2)$  edge-antimagic labeling of the corresponding graph.

5.2. 3D Graphical Comparison. The following are the 3D plot representations of the magic constants obtained in our

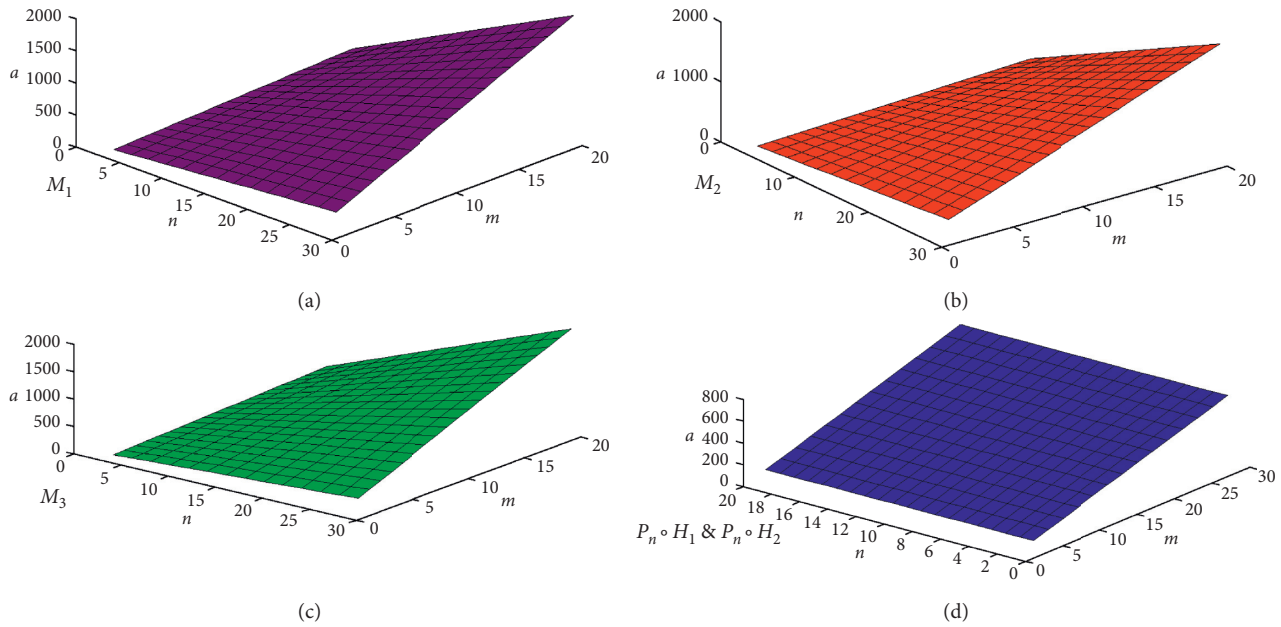


FIGURE 11: The 3D plots of the magic constants of  $M_1 \cong (P_n \circ K_{2,m}) \cup K_{1,n} \cup (n - 1/2)K_1$ ,  $M_2 \cong (P_n \circ K_{2,m}) \cup nP_2$ ,  $M_3 \cong (P_n \circ K_{2,m}) \cup P_{n+1}$ , and  $(P_n \circ H_1)$  and  $(P_n \circ H_2)$ .

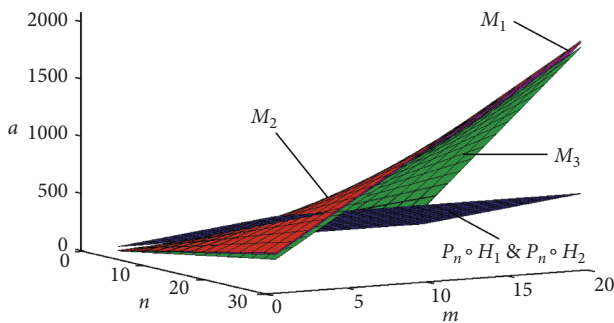


FIGURE 12: The comparison of the values of magic constants of  $M_1$ ,  $M_2$ , and  $M_3$ , as well as  $(P_n \circ H_1)$  and  $(P_n \circ H_2)$ .

results against the parameters  $n \leq 30$  and  $m \leq 20$ , shown in Figure 11 ( $a$  represents magic constant in the plots). Meanwhile, Figure 12 shows the comparison of these magic constants against the said values of the parameters. Among these magic constants, the most dominant magic constant, especially in the comparison of  $M_1$ ,  $M_2$ , and  $M_3$ , is of the graph  $M_2$  (as red layer is the dominant one). It means that, among the magic constants of the networks discussed here, the disjoint union of the rooted product  $(P_n \circ K_{2,m})$  and  $nP_2$  is having the highest values.

### 6. Conclusion

In this note,

- (1) we have obtained a super  $(a, 0)$  edge-antimagic labeling of the rooted product of  $P_n$  with the complete bipartite graph  $(K_{2,m})$  taking its union with path, copies of paths, and star for possible values of  $n$ . The

- obtained results partially point towards the problem proposed in [46] regarding  $nK_{2,m}$ ;
- (2) we have presented a super  $(a, 0)$  edge-antimagic labeling of the rooted product of  $P_n$  with special pancyclic graphs  $H_1$  and  $H_2$ ;
- (3) importantly, all the resultant graphs of the rooted products here are planar graphs;
- (4) several open problems have also been proposed in this article for further working in this area;
- (5) The obtained schemes are now all set to serve as test-ready labelings for engineers, programmers, and networking professionals to utilize them where they find them suitable.

### Data Availability

The whole data are included within this article. However, more details on the data can be obtained from the corresponding author upon request.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Acknowledgments

This work was supported in part by the China Postdoctoral Science Foundation under Grant 2017M621579, in part by the Postdoctoral Science Foundation of Jiangsu Province under Grant 1701081B, and in part by the Project of Anhui Jianzhu University under Grant 2016QD116 and Grant 2017dc03.

## References

- [1] Z. Shao, " $L(2, 1)$ -Labeling of the strong product of paths and cycles," *The Scientific World Journal*, vol. 2014, Article ID 741932, 12 pages, 2014.
- [2] T.-M. Wang and C.-C. Hsiao, "On anti-magic labeling for graph products," *Discrete Mathematics*, vol. 308, no. 16, pp. 3624–3633, 2008.
- [3] J. W. Daykin, C. S. Iliopoulos, M. Miller, and O. Phanalasy, "Antimagicness of generalized corona and snowflake graphs," *Mathematics in Computer Science*, vol. 9, no. 1, pp. 105–111, 2015.
- [4] B. M. Stewart, "Magic graphs," *Canadian Journal of Mathematics*, vol. 18, pp. 1031–1056, 1996.
- [5] I. W. Sudarsana, E. T. Baskoro, D. Ismailmuza, and H. Assiyatun, "Creating new super edge-magic total labeling from old ones," in *Proceedings of the Proceedings of the Second International Workshop on Graph Labeling (IWOGL)*, pp. 77–84, Batu, Indonesia, December 2004.
- [6] N. L. Prasanna, K. Sravanthi, and N. Sudhakar, "Applications of graph labeling in major areas of computer science," *International Journal of Research in Computer and Communication Technology*, vol. 3, pp. 819–823, 2014.
- [7] N. L. Prasanna, "Applications of graph labeling in communication networks," *Oriental Journal of Computer Science and Technology*, vol. 7, pp. 893–897, 2014.
- [8] M. Andreas, *Wireless Communications*, Wiley, New York, NY, USA, 2005.
- [9] R. Wakefield, "Radio broadcasting," 1959.
- [10] Z. Zhou, S. Das, and H. Gupta, "Connected  $K$ -coverage problem in sensor networks," in *Proceedings of the 13th International Conference on Computer Communications and Networks*, pp. 373–378, IEEE Cat. No.04EX969, Chicago, IL, 2004.
- [11] M. S. Vinutha and P. Arathi, "Applications of graph coloring and labeling in computer science," *International Journal on Future Revolution in Computer Science and Communication Engineering*, vol. 3, pp. 14–15, 2017.
- [12] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," *IEEE Personal Communications*, vol. 3, no. 3, pp. 10–31, 1996.
- [13] A. Krishnaa, "On antimagic labellings of some cycle related graphs," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 15, no. 4-5, pp. 225–235, 2012.
- [14] A. Krishnaa, "Formulas and algorithms of antimagic labelings of some helm related graphs," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 19, no. 2, pp. 435–445, 2016.
- [15] A. Krishnaa, "Some applications of labelled graphs," *International Journal of Mathematics Trends and Technology*, vol. 37, no. 3, pp. 19–23, 2016.
- [16] J. P. Desai, "A graph theoretic approach for modeling mobile robot team formations," *Journal of Robotic Systems*, vol. 19, no. 11, pp. 511–525, 2002.
- [17] D. B. West, *Introduction to Graph Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 2001.
- [18] J. Sedlacek, "Problem 27 in the theory of graphs and its applications," in *Proceedings of the Symposium Smolenice*, pp. 163–167, Smolenice, Slovakia, June 1963.
- [19] N. Hartsfield and G. Ringel, *Pearls in Graph Theory*, Academia Press, Boston, MA, USA, 1990.
- [20] A. Kotzig and A. Rosa, "Magic valuations of finite graphs," *Canadian Mathematical Bulletin*, vol. 13, no. 4, pp. 451–461, 1970.
- [21] G. Ringel and A. S. Llado, "Another tree conjecture," *Bulletin of the ICA*, vol. 18, pp. 83–85, 1996.
- [22] H. Enomoto, A. Llado, T. Nakamigawa, and G. Ringel, "Super edge-magic graphs," *SUT Journal of Mathematics*, vol. 34, pp. 105–109, 1998.
- [23] R. Simanjuntak, F. Bertault, and M. Miller, "Two new  $(a, d)$ -EAT graph labelings," in *Proceedings of the Eleventh Australian Workshop of Combinatorial Algorithm*, pp. 179–189, Hunter Valley, Australia, July 2000.
- [24] S. M. Lee and Q. X. Shan, "All trees with at most 17 vertices are super edgemagic," in *Proceedings of the 16th MCCC Conference*, University Southern Illinois., Carbondale, IL, USA, July 2002.
- [25] M. Javaid, A. A. Bhatti, and M. K. Aslam, "Super  $(a, d)$ -edge antimagic total labeling of a subclass of trees," *AKCE International Journal of Graphs and Combinatorics*, vol. 14, no. 2, pp. 158–164, 2017.
- [26] M. Javaid, "On super edge-antimagic total labeling of subdivided stars," *Discussiones Mathematicae Graph Theory*, vol. 34, no. 4, pp. 691–705, 2014.
- [27] M. Javaid and A. A. Bhatti, "On super  $(a, d)$  edge- antimagic total labeling of subdivided stars," *Ars Combinatoria*, vol. 105, pp. 503–512, 2012.
- [28] A. Raheem, A. Q. Baig, and M. Javaid, "On  $(a, d)$  -EAT labeling of subdivision of  $K_1, r$ ," *Journal of Information and Optimization Sciences*, vol. 39, no. 3, pp. 643–650, 2018.
- [29] A. N. M. Salman, A. A. G. Ngurah, and N. Izzati, "On super edge- magic total labeling of a subdivision of a star  $S_n$ ," *Utilitas Mathematica*, vol. 81, pp. 275–284, 2010.
- [30] M. Javaid and A. A. Bhatti, "On super  $(a, d)$  edge- antimagic total labeling of generalized extended  $-W$  trees," *AKCE International Journal of Graphs and Combinatorics*, vol. 11, pp. 115–126, 2014.
- [31] M. Javaid, A. A. Bhatti, and M. Hussain, "On  $(a, d)$  edge-antimagic total labeling of extended  $-W$  trees," *Utilitas Mathematica*, vol. 87, pp. 293–303, 2012.
- [32] M. Javaid, M. Hussain, K. Ali, and K. H. Dar, "Super edge-magic total labeling on  $W$ -trees," *Utilitas Mathematica*, vol. 86, pp. 183–191, 2011.
- [33] M. Hussain and E. T. Boskoro, "Slamin, on Super edge magic total labeling of banana trees," *Utilitas Mathematica*, vol. 79, pp. 243–251, 2009.
- [34] K. A. Sugeng, M. Miller, Slamin, and M. Baa, " $(a, d)$ edge-antimagic total labelings of caterpillars," *Lecture Notes in Computer Science*, vol. 3330, Article ID 169aAS180, 2005.
- [35] M. Javaid, A. A. Bhatti, and M. Hussain, "On super  $(a, d)$  edge- antimagic total labeling of subdivided caterpillar," *Utilitas Mathematica*, vol. 98, pp. 227–241, 2015.
- [36] R. M. Figueroa-Centeno, R. Ichishima, and F. A. Muntaner-Batle, "On super edge- magic graph," *Ars Combinatoria*, vol. 64, pp. 81–95, 2002.
- [37] A. A. Bhatti, A. Nisar, and M. Kanwal, "Radio number of wheel like graphs," *International Journal on Applications of Graph Theory In Wireless Ad Hoc Networks And Sensor Networks*, vol. 3, no. 4, pp. 39–57, 2011.
- [38] J. A. Gallian, "A dynamic survey of graph labeling," *Electronic Journal of Combinatorics*, 2019.
- [39] R. Khennoufa and O. Togni, "The radio antipodal and radio numbers of the hypercube," *Ars Combinatoria*, vol. 102, pp. 447–461, 2011.

- [40] R. Figueroa-Centeno, R. Ichishima, and F. Muntaner-Batle, "On edge-magic labelings of certain disjoint unions of graphs," *Australasian Journal of Combinatorics*, vol. 32, pp. 225–242, 2005.
- [41] R. M. Figueroa-Centeno, R. Ichishima, F. A. Muntaner-Batle, and A. Oshima, "A magical approach to some labeling conjectures," *Discussiones Mathematicae Graph Theory*, vol. 31, no. 1, pp. 79–113, 2011.
- [42] I. Gray, *New construction methods for vertex-magic total labeling of graphs*, Ph.D thesis, University of Newcastle, Callaghan, Australia, 2006.
- [43] R. M. Figueroa-Centeno, R. Ichishima, and F. A. Muntaner-Batle, "The place of super edge-magic labelings among other classes of labelings," *Discrete Mathematics*, vol. 231, no. 1–3, pp. 153–168, 2001.
- [44] A. Q. Baig, H. U. Afzal, M. Imran, and I. Javaid, "Super edge-magic labeling of volvox and pancyclic graphs," *Utilitas Mathematica*, vol. 93, pp. 49–56, 2014.
- [45] P. R. L. Pushpam and A. Saibulla, "Super  $(a, d)$  edge-anti-magic total labeling of some classes of graphs," *SUT Journal of Mathematics*, vol. 48, no. 1, pp. 1–12, 2012.
- [46] A. A. G. Ngurah, E. T. Baskoro, R. Simanjuntak, and S. Uttunggadewa, "On the super edge-magic strength and deficiency of graphs," *Kyoto CGGT*, vol. 4535, pp. 144–154, 2007.



## Research Article

# Hermite–Hadamard Inequalities for Harmonic $(s, m)$ -Convex Functions

Jian Zhong Xu,<sup>1</sup> Umar Raza,<sup>2</sup> Muhammad Waqas Javed,<sup>3</sup> and Zaryab Hussain <sup>4,5</sup>

<sup>1</sup>Department of Electronics and Information Engineering, Bozhou University, Bozhou 236800, China

<sup>2</sup>Department of Mathematics, University of Jhang, Jhang 35200, Pakistan

<sup>3</sup>Department of Sciences and Humanities, National University of Computer and Emerging Sciences (FAST), Chiniot 35400, Pakistan

<sup>4</sup>Department of Mathematics, Punjab College of Commerce New Campus, Faisalabad 38000, Pakistan

<sup>5</sup>Department of Mathematics, Government College University Faisalabad, Faisalabad 38000, Pakistan

Correspondence should be addressed to Zaryab Hussain; zaryabhussain2139@gmail.com

Received 11 July 2020; Accepted 25 August 2020; Published 26 September 2020

Guest Editor: Shaohui Wang

Copyright © 2020 Jian Zhong Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The objective of this article is to establish some Hermite–Hadamard-type inequalities via harmonic  $(s, m)$ -convex functions in the framework of conformal fractional integral.

## 1. Introduction and Preliminaries

Let  $\mathcal{C} \subset \mathbb{R}$  be an interval. Then,  $\mathcal{C}$  is said to be convex, if

$$(1-t)u + tv \in \mathcal{C}, \quad (1)$$

holds  $\forall u, v \in \mathcal{C}$  and  $t \in [0, 1]$ .

Let  $\mathcal{C} \subset \mathbb{R}$  be an interval. Then, a function  $f: \mathcal{C} \rightarrow \mathbb{R}$  is said to be convex (concave), if

$$f((1-t)u + tv) \leq (\geq) (1-t)f(u) + tf(v), \quad (2)$$

holds  $\forall u, v \in \mathcal{C}$  and  $t \in [0, 1]$ .

It can be easily seen in [1–7] that the convex (concave) functions have extensive applications in pure and applied mathematics, and in the literature [8–15], many eminent inequalities and other properties can be found in the framework of convexity. One of the renowned inequalities in the literature of Hermite–Hadamard Integral Inequality is given below:

$$f\left(\frac{u+v}{2}\right) \leq \frac{1}{v-u} \int_u^v \frac{f(x)}{x^2} dx \leq \frac{f(u) + f(v)}{2}. \quad (3)$$

These both inequalities hold in reverse if the function is concave. Now, the harmonic convex set is defined as follows.

*Definition 1.* Let  $\mathcal{C} \subset \mathbb{R}$  be an interval. Then,  $\mathcal{C}$  is said to be harmonic convex, if

$$\frac{uv}{(1-t)u + tv} \in \mathcal{C}, \quad (4)$$

holds  $\forall u, v \in \mathcal{C}$  with  $(u, v) \neq (0, 0)$  and  $t \in [0, 1]$ .

Iscan [8] introduced the concept of harmonic convex function.

*Definition 2* (see [8]). Let  $\mathcal{C} \subset \mathbb{R}$  be an interval. Then, a function  $f: \mathcal{C} \rightarrow \mathbb{R}$  is called harmonic convex (concave), if

$$f\left(\frac{uv}{(1-t)u + tv}\right) \leq (\geq) tf(u) + (1-t)f(v), \quad (5)$$

holds for all  $u, v \in \mathcal{C}$  with  $(u, v) \neq (0, 0)$  and  $t \in [0, 1]$ .

In [8], Iscan by using the concept of harmonic convex function gave a new refinement of Hermite–Hadamard inequality as



$$f\left(\frac{2uv}{u+v}\right) \leq \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \leq \frac{f(u)+f(v)}{2}. \quad (6)$$

*Definition 3.* Let  $\mathcal{C} \subset \mathbb{R}$  be an interval and  $s, m \in (0, 1]$ . Then, a function  $f: \mathcal{C} \rightarrow \mathbb{R}$  is called harmonic  $(s, m)$ -convex (concave), if

$$f\left(\frac{muv}{(1-t)u+mtv}\right) \leq (\geq) t^s f(u) + m(1-t)^s f(v), \quad (7)$$

holds  $\forall u, v \in \mathcal{C}$  with  $(u, v) \neq (0, 0)$  and  $t \in [0, 1]$ .

If  $s = m = 1$ , then harmonic  $(s, m)$ -convex function becomes the classical harmonic convex function. So harmonic convex function is a special case of harmonic  $(s, m)$ -convex function.

The main purpose of this article is to establish some conformable fractional estimates of Hermite–Hadamard-type inequalities via harmonic  $(s, m)$ -convex functions. Before going further towards our main results, let us have a brief review of the previously well known concepts and results. These preliminaries will be highly helpful in acquiring the main results.

The eminent gamma and beta functions are defined as

$$\begin{aligned} \Gamma(u) &= \int_0^\infty e^{-t} t^{u-1} dt, \quad \text{for } u > 0, \\ \beta(u, v) &= \int_0^1 t^{u-1} (1-t)^{v-1} dt = \frac{\Gamma u \Gamma v}{\Gamma(u+v)}, \quad \text{for } u, v > 0. \end{aligned} \quad (8)$$

The integral form of hypergeometric function is defined as

$${}_2F_1(u, v; w; z) = \frac{1}{\beta(v, w-v)} \int_0^1 t^{v-1} (1-t)^{w-v-1} (1-zt)^{-u} dt, \quad (9)$$

for  $|z| \geq 1$ .

Now, if  $f \in L_1[u, v]$  with  $u \geq 0$ , then Riemann–Liouville integrals  $I_{u^+}^\alpha f$  and  $I_{v^-}^\alpha f$  of any positive order  $\alpha$  are defined as

$$\begin{aligned} I_{u^+}^\alpha f(a) &= \frac{1}{\Gamma \alpha} \int_0^a (a-t)^{\alpha-1} f(t) dt, \quad a > u, \\ I_{v^-}^\alpha f(a) &= \frac{1}{\Gamma \alpha} \int_a^v (t-a)^{\alpha-1} f(t) dt, \quad a < v. \end{aligned} \quad (10)$$

For more details, see [11].

Recently, Abdeljawad [16] introduced the notation of right and left conformable fractional integrals for any positive order  $\alpha$  as follows.

*Definition 4* (see [16]). Let  $\alpha \in (n, n+1]$ . Then, the left and right conformable fractional integrals starting from  $u$  of any positive order  $\alpha$  is given as

$$I_{\alpha}^u f(t) = \frac{1}{n!} \int_u^t (t-a)^n (a-u)^{\alpha-n-1} f(a) da, \quad (11)$$

$$I_{\alpha}^v f(t) = \frac{1}{n!} \int_t^v (a-t)^n (v-a)^{\alpha-n-1} f(a) da.$$

**Lemma 1** (see [5]). Let  $f: \mathcal{C} = [u, v] \subset \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  be differentiable on  $\mathcal{C}^o$ ,  $u, v \in \mathcal{C}$ , and  $f' \in L[u, v]$ . Then,

$$\begin{aligned} \frac{f(u)+f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx &\leq \frac{uv(v-u)}{2} \\ &\cdot \int_0^1 \frac{1-2t}{(tv+(1-t)u)^2} f'\left(\frac{uv}{tv+(1-t)u}\right) dt. \end{aligned} \quad (12)$$

## 2. Main Results

In this section, we will present our main results.

**Theorem 1.** Let  $f: \mathcal{C} = [a, b] \subset \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  be a harmonic  $(s, m)$ -convex function such that  $f \in L_1[u, v]$  and  $s, m \in (0, 1]$ . Then,

$$\begin{aligned} f\left(\frac{2uv}{u+v}\right) &\leq \left(\frac{1}{2}\right)^s \frac{uv}{v-u} \left\{ \int_u^v \frac{f(x)}{x^2} dx + m \int_u^v \frac{f(x/m)}{x^2} dx \right\} \\ &\leq \left(\frac{1}{2}\right)^s \left\{ f(x) + mf\left(\frac{y}{m}\right) \right\}. \end{aligned} \quad (13)$$

*Proof.* By applying the definition of harmonic  $(s, m)$ -convex function for  $t = 1/2$ , we have

$$f\left(\frac{xy}{x+y}\right) \leq \left(\frac{1}{2}\right)^s \left\{ f(x) + mf\left(\frac{y}{m}\right) \right\}. \quad (14)$$

Put  $x = uv/tu + (1-t)v$  and  $y = uv/tv + (1-t)u$ . Then,

$$\begin{aligned} f\left(\frac{2uv}{u+v}\right) &\leq \left(\frac{1}{2}\right)^s \left\{ f\left(\frac{uv}{tu+(1-t)v}\right) + mf\left(\frac{uv}{m(tv+(1-t)u)}\right) \right\}, \\ &\leq \left(\frac{1}{2}\right)^s \left\{ \int_0^1 f\left(\frac{uv}{tu+(1-t)v}\right) dt + m \int_0^1 f\left(\frac{uv}{m(tv+(1-t)u)}\right) dt \right\}. \end{aligned} \quad (15)$$

We know that

$$\int_0^1 \left(\frac{1}{2}\right)^s f\left(\frac{uv}{tu+(1-t)v}\right) dt = \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx.$$

$\Rightarrow$

$$f\left(\frac{2uv}{u+v}\right) \leq \left(\frac{1}{2}\right)^s \frac{uv}{v-u} \left\{ \int_u^v \frac{f(x)}{x^2} dx + m \int_u^v \frac{f(x/m)}{x^2} dx \right\}. \quad (16)$$

Now, consider a function  $f: R \rightarrow R$  such that  $f(x) = 0$ . Then,

$$f\left(\frac{mxy}{mty + (1-t)x}\right) = 0. \tag{17}$$

Also,

$$t^s f(x) + m(1-t)^s f(y) = 0. \tag{18}$$

So  $f$  is harmonic  $(s, m)$ -convex, and also, we have

$$\begin{aligned} f\left(\frac{2uv}{u+v}\right) &= 0, \\ \left(\frac{1}{2}\right)^s \frac{uv}{v-u} \left\{ \int_u^v \frac{f(x)}{x^2} dx + m \int_u^v \frac{f(x/m)}{x^2} dx \right\} &= 0, \end{aligned} \tag{19}$$

which implies that the inequality holds. □

**Theorem 2.** Let  $f: \mathcal{C} = [u, v] \subset R \setminus \{0\} \rightarrow R$  be a harmonic  $(s, m)$ -convex function such that  $f \in L_1[u, v]$ , where  $s, m \in (0, 1]$ . Then,

$$\begin{aligned} \frac{\Gamma\alpha - n}{\Gamma\alpha + 1} f\left(\frac{2uv}{u+v}\right) &\leq \left(\frac{1}{2}\right)^s \left(\frac{uv}{v-u}\right)^\alpha \left[ I_\alpha^{1/v}(\text{fog})\left(\frac{1}{u}\right) \right. \\ &\quad \left. + m I_\alpha^{1/u}(\text{fog})\left(\frac{1}{v}\right) \right]. \end{aligned} \tag{20}$$

Here, fog is the composition function.

*Proof.* From inequality 1, we have

$$f\left(\frac{2xy}{x+y}\right) \leq \left(\frac{1}{2}\right)^s \left\{ f(x) + mf\left(\frac{y}{m}\right) \right\}. \tag{21}$$

Put  $x = uv/tu + (1-t)v$  and  $y = uv/tv + (1-t)u$ . Then,

$$f\left(\frac{2uv}{u+v}\right) \leq \left(\frac{1}{2}\right)^s \left\{ f\left(\frac{uv}{tu + (1-t)v}\right) + mf\left(\frac{uv}{m(tv + (1-t)u)}\right) \right\}, \tag{22}$$

$$\begin{aligned} \frac{1}{n!} f\left(\frac{2uv}{u+v}\right) \int_0^1 t^n (1-t)^{\alpha-n-1} dt &\leq \frac{1}{n!} \int_0^1 t^n (1-t)^{\alpha-n-1} f\left(\frac{uv}{tu + (1-t)v}\right) dt, \\ &+ \frac{m}{n!} \int_0^1 t^n (1-t)^{\alpha-n-1} f\left(\frac{uv}{m(tv + (1-t)u)}\right) dt = J_1 + J_2. \end{aligned} \tag{23}$$

By using change of variable technique of integration, we have

$$\begin{aligned} J_1 &= \frac{1}{n!} \int_0^1 t^n (1-t)^{\alpha-n-1} f\left(\frac{uv}{tu + (1-t)v}\right) dt \\ &= \left(\frac{uv}{v-u}\right)^\alpha I_\alpha^{1/v}(\text{fog})\left(\frac{1}{u}\right), \end{aligned} \tag{24}$$

where  $g(x) = 1/x$ , and

$$\begin{aligned} J_2 &= \frac{m}{n!} \int_0^1 t^n (1-t)^{\alpha-n-1} f\left(\frac{uv}{m(tv + (1-t)u)}\right) dt \\ &= m \left(\frac{uv}{v-u}\right)^\alpha I_\alpha^{1/u}(\text{fog})\left(\frac{1}{v}\right), \end{aligned} \tag{25}$$

where  $g(x) = 1/x$ .

Now, from (23), we have

$$\begin{aligned} \frac{\Gamma\alpha - n}{\Gamma\alpha + 1} f\left(\frac{2uv}{u+v}\right) &\leq \left(\frac{1}{2}\right)^s \left(\frac{uv}{v-u}\right)^\alpha \left[ I_\alpha^{1/v}(\text{fog})\left(\frac{1}{u}\right) \right. \\ &\quad \left. + m I_\alpha^{1/u}(\text{fog})\left(\frac{1}{v}\right) \right]. \end{aligned} \tag{26}$$

So the inequality holds. □

**Theorem 3.** Let  $f: \mathcal{C} = [u, v] \subset R \setminus \{0\} \rightarrow R$  be differentiable on  $\mathcal{C}^\circ$ ,  $u, v/m \in \mathcal{C}$ ,  $m \in [0, 1]$ , and  $f' \in L[u, v]$ . If  $|f'|^q$  for  $q \geq 1$  is harmonic  $(s, m)$ -convex on  $[u, v/m]$ , then

$$\left| \frac{f(u) + f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \right| \leq \frac{uv(v-u)}{2^{2-(1/q)}} (\Omega_1 |f'(u)|^q + m\Omega_2 |f'(v/m)|^q)^{1/q}, \tag{27}$$

where

$$\Omega_1 = \int_0^1 \frac{t^s}{(tv + (1-t)u)^2} dt = \frac{u^{-2}}{s+1} {}_2F_1\left(2, s+1; s+2; 1 - \frac{v}{u}\right),$$

$$\Omega_2 = \int_0^1 \frac{(1-t)^s}{(tv + (1-t)u)^2} dt = \frac{-u^{-2}}{s+1} {}_2F_1\left(2, 1; s+2; 1 - \frac{v}{u}\right). \tag{28}$$

*Proof.* Hölder's inequality and Lemma 1 implies that

$$\begin{aligned} \left| \frac{f(u) + f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \right| &\leq \frac{uv(v-u)}{2} \int_0^1 \left| \frac{1-2t}{(tv + (1-t)u)^2} f' \left( \frac{uv}{tv + (1-t)u} \right) \right| dt \\ &\leq \frac{uv(v-u)}{2} \left( \int_0^1 |1-2t| dt \right)^{1-1/q} \times \left( \int_0^1 \frac{1}{(tv + (1-t)u)^2} \left| f' \left( \frac{uv}{tv + (1-t)u} \right) \right|^q dt \right)^{1/q}. \end{aligned} \tag{29}$$

Since  $|f'|^q$  is harmonically  $(s, m)$ -convex on  $[u, v/m]$ , we have

$$\begin{aligned} &\leq \frac{uv(v-u)}{2} \left(\frac{1}{2}\right)^{1-1/q} \left( \int_0^1 \frac{1}{(tv + (1-t)u)^2} \left( t^s |f'(u)|^q + m \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q} dt \right), \\ &\leq \frac{uv}{2^{2-1/q}} \left( \int_0^1 \frac{t^s}{(tv + (1-t)u)^2} dt |f'(a)|^q + m \int_0^1 \frac{(1-t)^s}{(tv + (1-t)u)^2} dt \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}, \\ &= \frac{uv(v-u)}{2^{2-1/q}} \left( \Omega_1 |f'(u)|^q + m \Omega_2 \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}. \end{aligned} \tag{30}$$

Here,

$$\Omega_1 = \int_0^1 \frac{t^s}{(tv + (1-t)u)^2} dt = \frac{u^{-2}}{s+1} {}_2F_1(2, s+1; s+2; 1 - v/u), \Omega_2 = \int_0^1 \frac{(1-t)^s}{(tv + (1-t)u)^2} dt = \frac{-u^{-2}}{s+1} {}_2F_1(2, 1; s+2; 1 - v/u). \tag{31}$$

□ **Theorem 4.** Let  $f: \mathcal{C} = [u, v] \subset \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  be differentiable on  $\mathcal{C}^\circ$ ,  $u, v/m \in \mathcal{C}$ ,  $m \in [0, 1]$ , and  $f' \in L[u, v]$ . If  $|f'|^q$  for  $q \geq 1$  is harmonic  $(s, m)$ -convex on  $[u, v/m]$ , then

$$\left| \frac{f(u) + f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \right| \leq \frac{uv(v-u)}{2} u^{-2} {}_2F_1\left(2, 1; 2; 1 - \frac{v}{u}\right)^{1-1/q} \times \left( -\Psi |f'(u)|^q + m\Psi \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}, \tag{32}$$

where  $\Psi = s/(s+1)(s+2)$ .

*Proof.* Hölder's inequality and Lemma 1 implies that

$$\begin{aligned} \left| \frac{f(u) + f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \right| &\leq \frac{uv(v-u)}{2} \int_0^1 \left| \frac{1-2t}{(tv+(1-t)u)^2} f' \left( \frac{uv}{tv+(1-t)u} \right) \right| dt \\ &\leq \frac{uv(v-u)}{2} \left( \int_0^1 \frac{dt}{(tv+(1-t)u)^2} \right)^{1-1/q} \times \left( \int_0^1 (1-2t) \left| f' \left( \frac{uv}{tv+(1-t)u} \right) \right|^q dt \right)^{1/q}. \end{aligned} \tag{33}$$

Since  $|f'|^q$  is harmonically  $(s, m)$ -convex on  $[u, v/m]$ , we have

$$\begin{aligned} &\leq \frac{uv(v-u)}{2} u^{-2} {}_2F_1 \left( 2, 1; 2; 1 - \frac{v}{u} \right)^{1-1/q} \times \left( \int_0^1 (1-2t) \left( t^s |f'(u)|^q + m(1-t)^s \left| f' \left( \frac{v}{m} \right) \right|^q \right) dt \right)^{1/q}, \\ &\leq \frac{uv(v-u)}{2} v^{-2} {}_2F_1 \left( 2, 1; 2; 1 - \frac{v}{u} \right)^{1-1/q} \times \left( \int_0^1 (1-2t) t^s dt |f'(u)|^q + m \int_0^1 (1-2t) (1-t)^s dt \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}, \\ &\leq \frac{uv(v-u)}{2} u^{-2} {}_2F_1 \left( 2, 1; 2; 1 - \frac{v}{u} \right)^{1-1/q} \times \left( \frac{-s}{(s+1)(s+2)} |f'(u)|^q + m \frac{s}{(s+1)(s+2)} \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}, \\ &\leq \frac{uv(v-u)}{2} u^{-2} {}_2F_1 \left( 2, 1; 2; 1 - \frac{v}{u} \right)^{1-1/q} \times \left( -\Psi |f'(a)|^q + m \Psi \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}, \end{aligned} \tag{34}$$

where  $\Psi = s/(s+1)(s+2)$ .

□ **Theorem 5.** Let  $f: I = [u, v] \subset \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  be differentiable on  $\mathcal{C}^o$ ,  $u, v/m \in \mathcal{C}$ ,  $m \in [0, 1]$ , and  $f' \in L[u, v]$ . If  $|f'|^q$  for  $q \geq 1$  is harmonic  $(s, m)$ -convex on  $[u, v/m]$ , then

$$\left| \frac{f(u) + f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \right| \leq \frac{uv(v-u)}{2} \left( \frac{1}{p+1} \right)^{1/p} \times \left( \Phi_1 |f'(u)|^q + m \Phi_2 \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q}, \tag{35}$$

where

$$\begin{aligned} \Phi_1 &= \int_0^1 \frac{t^s}{(tv+(1-t)u)^{2q}} dt = \frac{u^{-2q}}{s+1} {}_2F_1 \left( 2q, s+1; s+2; 1 - \frac{v}{u} \right), \\ \Phi_2 &= \int_0^1 \frac{(1-t)^s}{(tv+(1-t)u)^{2q}} dt = \frac{-u^{-2q}}{s+1} {}_2F_1 \left( 2q, 1; s+2; 1 - \frac{v}{u} \right). \end{aligned} \tag{36}$$

*Proof.* Hölder's inequality and Lemma 1 implies that

$$\begin{aligned} \left| \frac{f(u) + f(v)}{2} - \frac{uv}{v-u} \int_u^v \frac{f(x)}{x^2} dx \right| &\leq \frac{uv(v-u)}{2} \int_0^1 \left| \frac{1-2t}{(tv+(1-t)u)^2} f' \left( \frac{uv}{tv+(1-t)u} \right) \right| dt, \\ &\leq \frac{uv(v-u)}{2} \left\{ \left( \int_0^1 |1-2t|^p \right)^{1/p} \times \left( \int_0^1 \frac{1}{(tv+(1-t)u)^{2q}} \left| f' \left( \frac{uv}{tv+(1-t)u} \right) \right|^q dt \right)^{1/q} \right\}. \end{aligned} \tag{37}$$

Since  $|f'|^q$  is harmonically  $(s, m)$ -convex on  $[u, v/m]$ , we have

$$\begin{aligned} &\leq \frac{uv(v-u)}{2} \left\{ \left( \frac{1}{p+1} \right)^{1/p} \times \left( t^s |f'(u)|^q + m(1-t)^s \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/q} dt \right\}, \\ &\leq \frac{uv(v-u)}{2} \left\{ \left( \frac{1}{p+1} \right)^{1/p} \times \left( \int_0^1 \frac{t^s}{(tv+(1-t)u)^{2q}} dt |f'(u)|^q + m \int_0^1 \frac{(1-t)^s}{(tv+(1-t)u)^{2q}} dt \left| f' \left( \frac{v}{m} \right) \right|^q \right)^{1/p} \right\}, \quad (38) \\ &\leq \frac{uv(v-u)}{2} \left\{ \left( \frac{1}{p+1} \right)^{1/p} \times \left( \Phi_1 |f'(u)|^q + m \left| \Phi_2 f' \left( \frac{v}{m} \right) \right|^q \right)^{1/p} \right\}, \end{aligned}$$

where

$$\begin{aligned} \Phi_1 &= \int_0^1 \frac{t^s}{(tv+(1-t)u)^{2q}} dt = \frac{u^{-2q}}{s+1} {}_2F_1 \left( 2q, s+1; s+2; 1-\frac{v}{u} \right), \\ \Phi_2 &= \int_0^1 \frac{(1-t)^s}{(tv+(1-t)u)^{2q}} dt = \frac{-u^{-2q}}{s+1} {}_2F_1 \left( 2q, 1; s+2; 1-\frac{v}{u} \right). \end{aligned} \quad (39)$$

This completes our arguments.  $\square$

### 3. Conclusion

In this article, we have established new Hermite–Hadamard-type inequalities via harmonic  $(s, m)$ -convex functions in the framework of fractional integrals. This article may be useful for the researchers to derive some other type of inequalities.

### Data Availability

No data were used to support this study.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This work was supported by the Key Natural Science Foundation of the Education Department of Anhui Province (KJ2019A1303 and KJ2019A1300); the Key Project for Teaching Research of Bozhou University (2018zdjy01 and 2018zdjy02); and the Key Natural Science Foundation of the Education Department of Bozhou University (BYZ2018B03 and BYZ2019B03). Moreover, the second author wants to acknowledge the efforts of the vice chancellor, University of Jhang, for providing such a good environment for research.

### References

- [1] T. H. Zhao, Y. M. Chu, and H. Wang, "Logarithmically complete monotonicity properties relating to the gamma

function," *Abstract and Applied Analysis*, vol. 13, p. 2011, 2011.

- [2] Y. Jiang and X. Xu, "A monotone finite volume method for time fractional Fokker-Planck equations," *Science China Mathematics*, vol. 62, no. 4, pp. 783–794, 2019.
- [3] Z. Cai, J. Huang, and L. Huang, "Periodic orbit analysis for the delayed Filippov system," *Proceedings of the American Mathematical Society*, vol. 146, no. 11, pp. 4667–4682, 2018.
- [4] Y. Tan, C. Huang, B. Sun, and T. Wang, "Dynamics of a class of delayed reaction-diffusion systems with Neumann boundary condition," *Journal of Mathematical Analysis and Applications*, vol. 458, no. 2, pp. 1115–1130, 2018.
- [5] L. Duan, X. Fang, and C. Huang, "Global exponential convergence in a delayed almost periodic Nicholson's blowflies model with discontinuous harvesting," *Mathematical Methods in the Applied Sciences*, vol. 41, no. 5, pp. 1954–1965, 2018.
- [6] W. Wang and Y. Chen, "Fast numerical valuation of options with jump under Merton's model," *Journal of Computational and Applied Mathematics*, vol. 318, pp. 79–92, 2017.
- [7] L. Duan and C. Huang, "Existence and global attractivity of almost periodic solutions for a delayed differential neo-classical growth model," *Mathematical Methods in the Applied Sciences*, vol. 40, no. 3, pp. 814–822, 2017.
- [8] I. Iscan, "Hermite-Hadamard type inequalities for harmonically convex functions," *Hacetatepe Journal of Mathematics and Statistics*, vol. 43, no. 6, pp. 935–942, 2014.
- [9] M. U. Awan, M. A. Noor, M. V. Mihai, and K. I. Noor, "Inequalities via harmonic convex functions: conformable fractional calculus approach," *Journal of Mathematical Inequalities*, vol. 12, no. 1, pp. 143–153, 2018.
- [10] M. U. Noor, S. Talib, Y. M. Chu, M. A. Noor, and K. I. Noor, "Some new refinements of Hermite-Hadamard-type inequalities involving-Riemann-Liouville fractional integrals and applications," *Mathematical Problems in Engineering*, vol. 2020, Article ID 3051920, 2020.
- [11] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John-Wily and Sons. Inc., Hoboken, NJ, USA, 1993.
- [12] M. A. Noor, K. I. Noor, and M. U. Awan, "Integral inequalities for coordinated harmonically convex functions," *Complex Variables and Elliptic Equations*, vol. 60, no. 6, pp. 776–786, 2015.
- [13] M. Z. Awan, E. Set, H. Yaldiz, and N. Basak, "Hermite-Hadamard's inequalities for fractional integrals and related

- fractional inequalities,” *Mathematical and Computer Modelling*, vol. 57, no. 9-10, pp. 2403–2407, 2013.
- [14] S. Talib, M. U. Awan, M. A. Noor, and K. I. Noor, “Approximately h-preinvex functions, associated Hermite-Hadamard-like inequality, new q-identity, and estimation of its bounds with applications,” *AIP Advances*, vol. 10, no. 4, Article ID 045209, 2020.
- [15] S. Wu, M. U. Awan, M. V. Mihai, M. A. Noor, and S. Talib, “Estimates of upper bound for a kth order differentiable functions involving Riemann-Liouville integrals via higher order strongly h-preinvex functions,” *Journal of Inequalities and Applications*, vol. 227, no. 1, 2019.
- [16] T. Abdeljawad, “On conformable fractional calculus,” *Journal of Computational and Applied Mathematics*, vol. 279, pp. 57–66, 2015.



## Research Article

# Super-Efficiency Infeasibility in the Presence of Nonradial Measurement

Dong Guo<sup>1</sup> and Zheng-Qun Cai<sup>2</sup> 

<sup>1</sup>*School of Mathematics and Physics, Anhui Jianzhu University, Hefei 230601, China*

<sup>2</sup>*School of Foreign Studies, Anhui Jianzhu University, Hefei 230601, China*

Correspondence should be addressed to Zheng-Qun Cai; [caizhengqun1983@163.com](mailto:caizhengqun1983@163.com)

Received 11 August 2020; Revised 6 September 2020; Accepted 10 September 2020; Published 19 September 2020

Academic Editor: Shaohui Wang

Copyright © 2020 Dong Guo and Zheng-Qun Cai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The original radial VRS super-efficiency model in DEA excludes the DMU under evaluation from the reference set. However, it must lead to the problem of infeasibility specifically when the DMU under consideration is at the extremity of the frontier. In this paper, a modified nonradial VRS super-efficiency model is established. The super-efficiency model in the presence of nonradial measurement still maintains some good properties, and the original radial VRS super-efficiency model infeasibility can also be detected through it. Our model with nonradial measurement can help decision makers allocate input resources and arrange production activities because it finds an efficient benchmark DMU, which is different from the reference DMU under radial measurement.

## 1. Introduction

Data envelopment analysis (DEA), first introduced by Charnes et al. [1] and extended by Banker et al. [2], is an effective nonparametric technique for measuring the relative efficiency of peer decision-making units (DMUs) with multiple inputs and outputs. Their initial models based upon the constant returns to scale (CRS) and variable returns to scale (VRS) are commonly referred to as the CCR model and the BCC model, respectively. They compute scalar efficiency scores with a range of zero to unity which indicate how efficient each DMU has performed as compared to other DMUs in converting inputs to outputs and determine efficient level or position for each DMU under evaluation among all DMUs. DMUs that obtain a score of unity are deemed as efficient and on the DEA (best-practice) frontier, while other DMUs are treated as inefficient. Nowadays, DEA has become a popular method without any prior complicated weight assumptions since its advent and has been rapidly applied in improving the performances of different kinds of entities engaged in different activities and contexts [3, 4], such as human resources planning [5, 6], fixed cost allocation [7, 8], and resource sharing [9, 10].

However, when some DMUs all get scores of unity, these efficient DMUs cannot be distinguished further through the CCR model alone. To break the tie of efficient DMUs and further enhance the discrimination power of DEA, Andersen and Petersen [11] proposed a new model according to the CCR model, which is called super-efficiency model where the DMU under evaluation is excluded from the reference set. It allows efficient DMUs to have efficiency scores larger than or equal to unity (under input-oriented super-efficiency model), and for inefficient DMUs, the super-efficiency model yields scores that are identical to those received from the CCR model. Analogously, based upon the variable returns to scale (VRS) model of Banker et al. [12], the VRS super-efficiency model can be obtained. But under the condition of VRS, the super-efficiency model may be infeasible when some efficient DMUs are under evaluation, while the super-efficiency model under CRS does not suffer the problem of infeasibility. In face of this trouble, much effort has been focused on solving the problem of VRS super-efficiency model's infeasibility.

Seiford and Zhu [13] indicate that infeasibility must occur in the case of the variable returns to scale (VRS) super-efficiency model and further provide the necessary and sufficient conditions for infeasibility of super-efficiency models. Lovell and

Rouse [14] assign a user-defined scaling factor to find a feasible solution for those efficient DMUs for which feasible solutions are unavailable in the VRS super-efficiency model. Chen [15] shows that in order to fully characterize the super-efficiency, both input-oriented and output-oriented super-efficiency DEA models are needed when infeasibility occurs. Chen [15] further points out that super-efficiency can be regarded as input saving/output surplus achieved by an efficient DMU. Cook et al. [16] develop a modified VRS super-efficiency model that yields optimal solutions and super-efficiency scores that characterize the extent of super-efficiency in both inputs and outputs. Lee et al. [17] develop a two-stage process to address the VRS infeasibility issue. In the first stage, they test whether a VRS super-efficiency model is infeasible by investigating the existence of output surplus (input saving) when infeasibility occurs in the input-oriented (output-oriented) VRS super-efficiency model. In the second stage, they proposed a modified VRS super-efficiency model to yield a super-efficiency score that characterizes both the radial efficiency and input saving/output surplus. Chen and Liang [18] further prove that the two-stage process can be solved in a single linear program. However, when a DMU has zero data, these models may still be infeasible. Thrall [19] and Zhu [20] point out that the CRS super-efficiency model can also be infeasible when an efficient DMU has zero input values. The same conclusion can be applied to non-CRS super-efficiency models. Lee et al. [21] first point out that zero output data will not lead to infeasibility of the output-oriented super-efficiency models developed in the studies of Cook et al. [16], Lee et al. [17], and Chen and Liang [18]. This is because the output side of the constraints can always be satisfied. Therefore, they only assume that some inputs are zero for some efficient DMUs. Then, they revise the model of Lee et al. [17], and the revised model will be feasible when zero data exist in inputs.

Nevertheless, these models mentioned above so far only consider radial efficiency. The current paper extends the work of Lee and Zhu [21] to nonradial measurement. And we find that the super-efficiency model in the presence of nonradial measurement still maintains some good properties. In fact, super-efficiency with nonradial measurement can also help decision makers more allocate input resources according to underlying preferences and resource availability.

The remainder of this paper is organized as follows: Section 2 looks back upon several radial super-efficiency models that had been established and the problem of super-efficiency infeasibility. In Section 3, we develop our super-efficiency model with nonradial measurement and demonstrate that the model has some good properties. Meanwhile, we point out the difference between radial and nonradial measurement. In Section 4, the newly developed approach is applied to a data set on the 15 Illinois strip mines. In the end, main conclusions are given in Section 5.

## 2. Radial Super-Efficiency Models

Suppose there are  $n$  DMUs  $\{DMU_j (j = 1, 2, \dots, n)\}$ . Each DMU <sub>$j$</sub>  consumes a set of  $m$  inputs,  $x_{ij} (i = 1, \dots, m)$ , in the production of a set of  $s$  outputs,  $y_{rj} (r = 1, \dots, s)$ . Based upon the VRS model of Banker et al. [12], the input-oriented

VRS super-efficiency model for efficient DMU <sub>$k$</sub>  can be expressed as

$$\begin{aligned}
 & \min \quad \theta \\
 & \text{s.t.} \quad \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} \leq \theta x_{ik}, \quad i = 1, \dots, m \\
 & \quad \quad \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq y_{rk}, \quad r = 1, \dots, s \\
 & \quad \quad \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1 \\
 & \quad \quad \lambda_j \geq 0, \quad j \neq k,
 \end{aligned} \tag{1}$$

where the DMU <sub>$k$</sub>  under consideration is excluded from the reference set. Obviously, one of the circumstances where model (1) is infeasible if the DMU under consideration has the largest outputs, regardless of the input values. In fact, as pointed out by Lee et al. [17], when the outputs of the evaluated DMU is outside the production possibility set spanned by the outputs of the remaining DMUs, the infeasibility of input-oriented super-efficiency will occur.

Consider the same simple numerical example given in the study of Lee et al. [17] in Table 1. Owing that DMU<sub>D</sub> has the largest outputs and DMU<sub>E</sub> has zero input data, they are both infeasible under model (1). Lee et al. [17] develop a new super-efficiency model and two-stage process to investigate the issue of the VRS super-efficiency infeasibility. Chen and Liang [18] establish a single model and integrate the two-stage process into an equivalent ‘‘one model’’ approach:

$$\begin{aligned}
 & \min \quad \tau + M \times \sum_{r=1}^s \beta_r \\
 & \text{s.t.} \quad \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} \leq (1 + \tau) x_{ik}, \quad i = 1, \dots, m \\
 & \quad \quad \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq (1 - \beta_r) y_{rk}, \quad r = 1, \dots, s \\
 & \quad \quad \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1 \\
 & \quad \quad \lambda_j \geq 0, \quad j \neq k, \beta_r \geq 0,
 \end{aligned} \tag{2}$$

where  $M$  is a user-defined large positive number, and in the study of Cook et al. [16],  $M$  is set equal to  $10^5$ . When all inputs are positive, model (1) is feasible if and only if all  $\beta_r^* = 0$  in model (2). The super-efficiency score is defined as  $1 + \tau^* + 1/|R| \sum_{r \in R} 1/1 - \beta_r^*$ , where  $R = \{r | \beta_r^* > 0\}$ . However, owing to zero data  $X_2 = 0$  in DMU<sub>E</sub>, as depicted in Figure 1, DMU<sub>E</sub> remains infeasible under model (2).

TABLE 1: Data set of a simple example.

DMU	$X_1$	$X_2$	$Y_1$	Super-efficiency model (1)	Super-efficiency model (2)
A	2	1	1	1	1
B	1	2	1	1.4	1.4
C	1	4	2	Infeasible	3
D	2	3	1	0.6	0.6
E	3	0	1	Infeasible	Infeasible

Subsequently, Lee et al. [21] consider the situation when zero input data exist. They propose model (3) to address this issue about zero input data and define the super-efficiency score as  $1 + \tau^* + 1/|R| \sum_{r \in R} 1/1 - \beta_r^* + 1/|I| \sum_{r \in R} 1 + t_i^*/1$  when neither one of the sets  $R$  and  $I$  is empty, in which  $R = \{r | \beta_r^* > 0\}$  and  $I = \{i | t_i^* > 0\}$ . They demonstrate that models (2) and (3) yield the same results when data are positive:

$$\begin{aligned}
 \min \quad & \tau + M \times \left( \sum_{r=1}^s \beta_r + \sum_{i=1}^m t_i \right) \\
 \text{s.t.} \quad & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} - t_i x_i^{\max} \leq (1 + \tau) x_{ik}, \quad i = 1, \dots, m \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq (1 - \beta_r) y_{rk}, \quad r = 1, \dots, s \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1 \\
 & \lambda_j \geq 0, j \neq k, t_i \geq 0, \beta_r \geq 0, \tau \text{ is unrestricted,}
 \end{aligned} \tag{3}$$

where  $x_i^{\max} = \max_{1 \leq k \leq n} \{x_{ik}\}$ . The radial scaling factor  $1 + \tau^*$ , super-efficiency score, and more details of results by using model (3) that are computed by Lee et al. [21] are listed in Table 2. Here,  $DMU_E$  can be feasible under model (3).

### 3. A Modified Nonradial Super-Efficiency Model

In the real production processes, decision makers tend to arrange production plan reasonably according to possession of the resources and production capability. Investment of input resources must be consistent with resource availability and underlying preferences (for example, decision makers may be more willing to use cheaper to help lower costs). Radial scaling with inputs, i.e., decreasing inputs, proportionally is not always met in the actual production processes. Therefore, it is necessary and meaningful to incorporate nonradial measurement into super-efficiency model at some point. So, we modify model (3) to the situation of nonradial measurement. Consider the following model for  $DMU_k$ :

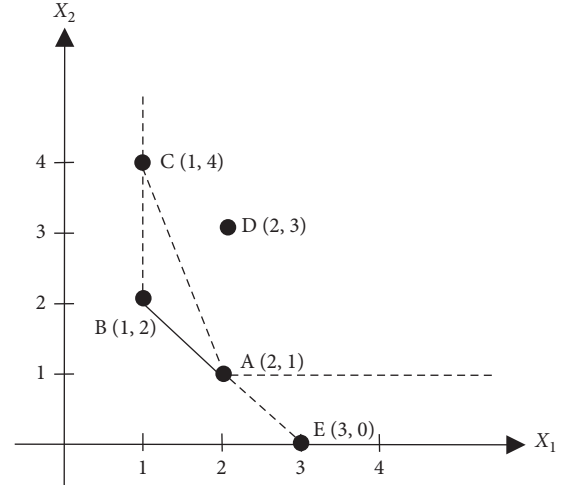


FIGURE 1: A simple example.

$$\begin{aligned}
 \min \quad & \frac{1}{m} \sum_{i=1}^m \tau_i + M \times \left( \sum_{r=1}^s \beta_r + \sum_{i=1}^m t_i \right) \\
 \text{s.t.} \quad & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} - t_i x_i^{\max} \leq (1 + \tau_i) x_{ik}, \quad i = 1, \dots, m \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq (1 - \beta_r) y_{rk}, \quad r = 1, \dots, s \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1 \\
 & \lambda_j \geq 0, j \neq k, 0 \leq t_i \leq 1, \beta_r \geq 0, \tau_i \text{ is unrestricted.}
 \end{aligned} \tag{4}$$

Under nonradial measurement, we permit that each input element could be scaled down by an exclusive scaling factor,  $1 + \tau_i$  ( $i = 1, \dots, m$ ), instead of the common proportional scaling factor  $1 + \tau$  for all inputs in model (3). Obviously, model (4) still works when zero input data exist. Based on model (4), some good properties can be derived as follows.

**Theorem 1.** In model (4),  $\tau_i^* \geq -1$  and  $1/m \sum_{i=1}^m \tau_i^* \geq -1$ .

*Proof.* Note that  $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1$ . Therefore,  $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} - t_i x_i^{\max} \leq \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_i^{\max} - t_i x_i^{\max} = (1 - t_i) x_i^{\max}$ . Also, there exists  $0 \leq t_i \leq 1$ ; thus,  $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} - t_i x_i^{\max} \geq 0$ . So, when  $\tau_i$  is large enough,  $1 + \tau_i^* \geq 0$  must exist; then,  $\tau_i^* \geq -1$ . Furthermore,  $1/m \sum_{i=1}^m \tau_i^* \geq -1$ .  $\square$

TABLE 2: Results under model (3).

DMU	$1 + \tau^*$	Input saving	Output surplus	Super-efficiency	$t_1^*$	$t_2^*$	$\beta_1^*$
A	1	0	0	1	0	0	0
B	1.4	0	0	1.4	0	0	0
C	1	0	2	3	0	0	0.5
D	0.6	0	0	0.6	0	0	0
E	0.666667	1.25	0	1.916667	0	0.25	0

**Theorem 2.** Model (4) is identical with model (3) when  $\tau_1 = \tau_2 = \dots = \tau_m$ .

In other words, when  $\tau_1 = \tau_2 = \dots = \tau_m$ , the objective function values of model (3) and model (4) are identical.

**Theorem 3.** Model (4) is always feasible.

*Proof.* According to the proof for feasibility of model (3) of Lee et al. [21], model (3) is always feasible. Suppose  $\bar{\lambda}_j, \bar{t}_i, \bar{\beta}_r, \bar{\tau}$  is a feasible solution of model (3), let  $\tau_1 = \tau_2 = \dots = \tau_m = \bar{\tau}$ , then,  $\bar{\lambda}_j, \bar{t}_i, \bar{\beta}_r, \tau_1, \tau_2, \dots, \tau_m$  is just a feasible solution to model (4). In other words, the feasible region of model (4) is not empty. Thus, model (4) is always feasible.  $\square$

**Theorem 4.** Model (1) is infeasible if and only if

- (i) Some  $\beta_r^* > 0$  ( $r = 1, \dots, s$ ) when  $x_i^{\max} = x_{ik}$ , where  $\beta_r^*$  is the optimal solution in model (4)
- (ii) Some  $\beta_r^* > 0$  ( $r = 1, \dots, s$ ) or some  $t_i^* > 0$  ( $i = 1, \dots, m$ ) when  $x_i^{\max} \neq x_{ik}$ , where  $\beta_r^*, t_i^*$  is the optimal solution in model (4)

*Proof*

- (i) When  $x_i^{\max} = x_{ik}$ , the first constraint condition in model (4) becomes  $\sum_{j \neq k} \lambda_j x_{ij} \leq (1 + \tau_i + t_i) x_{ik}$ ,  $i = 1, \dots, m$ . Let  $\tau'_i = \tau_i + t_i$ , then model (4) could turn out to be

$$\begin{aligned}
 \min \quad & \frac{1}{m} \sum_{i=1}^m \tau'_i + M \times \sum_{r=1}^s \beta_r \\
 \text{s.t.} \quad & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} \leq (1 + \tau'_i) x_{ik}, \quad i = 1, \dots, m \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq (1 - \beta_r) y_{rk}, \quad r = 1, \dots, s \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1 \\
 & \lambda_j \geq 0, \quad j \neq k, 0 \leq t_i \leq 1, \beta_r \geq 0, \tau_i \text{ is unrestricted.}
 \end{aligned} \tag{5}$$

Based upon the studies of Cook et al. [16], Lee et al. [17] and Chen and Liang [18], the proposition is obviously true, i.e., model (1) is infeasible if and only if some  $\beta_r^* > 0$  ( $r = 1, \dots, s$ ) in model (4).

- (ii) When  $x_i^{\max} \neq x_{ik}$ ,  $\beta_r^* = 0$  ( $r = 1, \dots, s$ ) and  $t_i^* = 0$  ( $i = 1, \dots, m$ ) in model (4) indicate that model (1) is feasible. So, model (1) is infeasible, and there must be some  $\beta_r^* > 0$  ( $r = 1, \dots, s$ ) or some  $t_i^* > 0$  ( $i = 1, \dots, m$ ) in model (4). On the contrary, if model (1) is feasible, this means  $\beta_r^* = 0$  ( $r = 1, \dots, s$ ) and  $t_i^* = 0$  ( $i = 1, \dots, m$ ) is a feasible solution to model (4), which directly contradict the fact that some  $\beta_r^* > 0$  or  $t_i^* > 0$  is the optimal solution in model (4).  $\square$

**Theorem 5.** Results of model (4) is not larger than that of model (3).

*Proof.* Adding the constraint condition of  $\tau_1 = \tau_2 = \dots = \tau_m$  to model (4), model (4) has become model (3). So, the feasible region of model (3) is smaller than model (4). A relaxation in one of the constraints may yield a better optimal solution inside the new polyhedron in space. Therefore, the result of model (4) is not larger than that of model (3).  $\square$

**Theorem 6.** Models (4) and (3) yield the same amount of input saving and output surplus when zero input data exist.

*Proof.* For any  $i' \in (1, 2, \dots, m)$ , suppose DMU <sub>$k$</sub>  has zero input data  $x_{i'k} = 0$ , we have  $(1 + \tau_i) x_{ik} = 0$ . Due to  $\tau_i \geq -1$  and the minimization of objective function, there must exist  $\tau_i^* = -1$ . So, the first constraint condition  $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} - t_i x_i^{\max} \leq (1 + \tau_i) x_{ik}$  in model (4) will be  $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} - t_i x_i^{\max} \leq 0$ , which is identical with that in model (3) when  $x_{i'k} = 0$ .

Theorem 6 also indicates the way where the DMU under evaluation in model (4) is projected to the frontier formed by other DMUs is the same as that in model (3) when the DMU has zero input values.

The super-efficiency model (4) with nonradial measurement of ours also determines an efficient referent (benchmark) DMU that is on the frontier formed by the remaining DMUs. In order to fully characterize the super-efficiency by input saving/output surplus achieved by an efficient DMU if it exists when the VRS super-efficiency feasibility is present, and based upon Cook et al. [16] and Lee et al. [17, 21], we also denote  $R = \{r | \beta_r^* > 0\}$  and  $I = \{i | t_i^* > 0\}$ . Then input savings index and output savings index can be defined in the following manner:

TABLE 3: Results under model (4).

DMU	$1 + 1/m \sum_{i=1}^m \tau_i^*$	Input saving	Output surplus	Super-efficiency	$\tau_1^*$	$\tau_2^*$	$t_1^*$	$t_2^*$	$\beta_1^*$
A	0.75	0	0	0.75	0.5	-1	0	0	0
B	1.25	0	0	1.25	1	-0.5	0	0	0
C	0.75	0	2	2.75	0	-0.5	0	0	0.5
D	0.58335	0	0	0.58335	-0.5	-0.3333	0	0	0
E	0.33349	1.25	0	1.58349	-0.3333	-1	0	0.25	0

$$\hat{i} = \begin{cases} 0, & \text{if } I = \phi, \\ \frac{\sum_{i \in I} (x_i^{\max} + t_i^* x_i^{\max} / x_i^{\max})}{|I|} = \frac{\sum_{i \in I} (1 + t_i^*)}{|I|} & \text{if } I \neq \phi, \end{cases}$$

$$\hat{o} = \begin{cases} 0, & \text{if } R = \phi, \\ \frac{\sum_{r \in R} (1/1 - \beta_r^*)}{|R|}, & \text{if } R \neq \phi, \end{cases} \quad (6)$$

where  $|R|$  and  $|I|$  are the cardinality of the sets  $R$  and  $I$ , respectively.

Unlike super-efficiency score of Lee et al. [21], we put  $1 + 1/m \sum_{i=1}^m \tau_i^* \geq 0$  as the nonradial efficiency. Then, the super-efficiency score which could also characterize the super-efficiency in both inputs and outputs can be defined as

$$\tilde{\theta} = 1 + \frac{1}{m} \sum_{i=1}^m \tau_i^* + \hat{i} + \hat{o}. \quad (7)$$

The efficiency measure  $\tilde{\theta}$  consists of three parts: the nonradial efficiency  $1 + 1/m \sum_{i=1}^m \tau_i^*$ , the input saving index  $\hat{i}$ , and the output surplus index  $\hat{o}$ . As described by Lee et al. [21], when the set  $I = \{i \mid t_i^* > 0\}$  is not empty, the input savings index  $1/|I| \sum_{i \in I} (1 + t_i^*/1)$  reflects how far the  $DMU_k$  is below the dashed horizontal efficient boundary (see the dashed line through A in Figure 1). And the output surplus index  $1/|R| \sum_{r \in R} (1/1 - \beta_r^*)$  reflects how far the  $DMU_k$  is above the dashed efficient boundary.

We use Figure 1 to illustrate the difference from the method of Lee et al. [21], mainly including the different projection mode to the (best-practice) frontier formed by other DMUs and the different nonradial super-efficiency result obtained. For instance, for  $DMU_B$ , the efficient frontier is made up of broken lines CA and AE.  $DMU_B$  is on the left of the frontier which implies that its super-efficiency should be greater than 1.  $DMU_B$  achieves radial efficiency of 1.4 and gets its projective point (efficient referent DMU) at  $(1.4, 2.8) (= 1.4 \times (1, 2))$  on the frontier through scaling down proportionally under radial measurement. While under nonradial measurement,  $DMU_B$  gets its projective point (efficient referent DMU) at  $(2, 1) (= (1 + 1) \times 1, (1 - 0.5) \times 2)$ ; that is to say,  $DMU_B$  takes  $DMU_A$  as its benchmark under nonradial measurement. Meanwhile,  $DMU_B$  gains its nonradial efficiency of  $1.25 (= 1 + 1/2(1 - 0.5))$ . For  $DMU_C$ , on account that it has the largest input  $X_2 = 4$  and  $\beta_1^* = 0.5 > 0$  in model (4), it is infeasible under model (1), but it is on the frontier and gets

radial efficiency of 1 under model (3). It gets its projective point on the frontier at  $(1, 2) (= (1 + 0) \times 1, (1 - 0.5) \times 4)$  under nonradial measurement, i.e.,  $DMU_C$  puts  $DMU_B$  as its best-practice benchmark. Now,  $DMU_C$  earns its nonradial efficiency of  $0.75 (= 1 + 1/2(0 - 0.5))$ .  $DMU_C$  has the same amount of output surplus of 2 under both measurements which equals the distance from  $DMU_B$  to  $DMU_C$ . Finally, for  $DMU_E$ , it has zero input data  $X_2 = 0 \neq X_2^{\max}$  and  $t_2^* = 0.25 > 0$  in model (4), it is also infeasible under model (1). However, it reaches the efficient referent (benchmark)  $DMU_A$  in the same type of way and derives the same input saving of  $1.25 (= 1 + 0.25/1)$  under both measurements. Concretely,  $DMU_E$  goes through point  $(2, 0)$  on  $X$ -axis by radial or nonradial scaling and then moves upwards by 1 ( $t_i^* x_i^{\max} = 1$ ). But it has different radial efficiency of 0.666667 and nonradial efficiency of 0.33349. Table 3 shows more specific details and final super-efficiency result of model (4) for each DMU under nonradial measurement.  $\square$

#### 4. An Empirical Example

We apply our approach to the data set which contains a set of 15 Illinois strip mines in Lee et al. [21]. The data set is shown in Table 4, which has only one output of tons of coal produced and eight inputs. Of the inputs, one is labor in thousand miner days, and the other variables include three capital variables (K1, K2, and K3) and four geological variables (T1, D1, T2, and D2). In Table 4, K1 = bucket capacity of draglines, K2 = dipper capacity of power shovels, K3 = earth moving capacity of wheel excavators, T1 = thickness of first (upper) seam mined, D1 = depth to first seam mined, T2 = thickness of second (lower) seam mined, and D2 = depth to second seam mined. In this particular data set, there are many zero inputs.

Table 5 shows super-efficiency score and solutions details of each DMU under four models, respectively. But the very first point which needs to be made is that we set  $M = 10^4$ , here, to seek its optimal solutions of all DMUs under model (4) with nonradial measurement. Because under nonradial measurement, it should adjust the value of user-defined large positive number  $M$  to suit for the specific data set to find its optimal solution. From Table 5, we discover that there are 5 DMUs that are infeasible under the original super-efficiency model (1). From Theorem 4 and the solution details of model (4), we know that  $DMU_1$ ,  $DMU_2$ , and  $DMU_4$  have nonzero output surplus because of the existence of  $\beta_1^* > 0$  and  $DMU_{10}$  and  $DMU_{13}$  have nonzero input saving because of the existence of  $t_2^* > 0$  and  $t_3^* > 0$ , respectively. For example,  $DMU_{10}$  is infeasible under model (1) because the second



TABLE 4: Illinois strip mines.

DMU	Labor (1000 miner days)	K1 (cubic yards)	K2 (cubic yards)	K3 (cubic yards per hour)	T1 (feet)	1/D1 (D1 in feet)	T2 (feet)	1/D2 (D2 in feet)	Output (1000 tons)
1	98.5	142	245	0	6	0.016	4.3	0.012	3264
2	96.5	30	215	0	6	0.016	0	0	3065
3	57.6	18	105	0	5.6	0.026	4.2	0.016	2275
4	59.2	160	0	0	5.9	0.025	3.7	0.011	1978
5	57.6	200	0	0	8	0.022	3.5	0.011	1833
6	49.9	27	85	0	4.5	0.019	0	0	1218
7	53.5	143	65	0	6	0.01	0	0	928
8	34	70	65	12	6	0.02	5	0.01	919
9	39.6	67.5	40	0	6.5	0.013	0	0	777
10	51.3	0	145	0	3.2	0.019	0	0	745
11	74.2	110	65	0	2.1	0.014	0	0	742
12	24	25	65	0	4.4	0.012	0	0	488
13	26.5	58	0	0	3	0.014	0	0	407
14	43.1	70	0	0	6.5	0.012	0	0	402
15	20.7	236	0	0	5.7	0.01	0	0	396

TABLE 5: Illinois strip mines super-efficiency score (SE score).

DMU	Model (1)	Model (2)		Model (3)		Model (4)		Solution details	
		SE score	Solution details	$1 + \tau^*$	SE score	Solution details	$1 + 1/m \sum_{i=1}^m \tau_i^*$		SE score
1	Infeasible	2.064927	$\beta_1^* = 0.060968$	1	2.064927	$\beta_1^* = 0.060968$	0.508564	1.573490	$\beta_1^* = 0.060968$
2	Infeasible	3.70392	$\beta_1^* = 0.60261$	1.1875	3.70392	$\beta_1^* = 0.60261$	0.468745	2.985165	$\beta_1^* = 0.60261$
3	1.547346	1.547346		1.547346	1.547346		0.735467	0.735467	
4	Infeasible	2.435037	$\beta_1^* = 0.073306$	1.35593	2.435037	$\beta_1^* = 0.073306$	0.800606	1.879711	$\beta_1^* = 0.073306$
5	1.073871	1.073871		1.073871	1.073871		0.67373	0.67373	
6	1.159819	1.159819		1.159819	1.159819		0.613154	0.613154	
7	1.14681	1.14681		1.14681	1.14681		0.47557	0.47557	
8	0.937467	0.937467		0.937467	0.937467		0.485029	0.485029	
9	1.001438	1.001438		1.001438	1.001438		0.51114	0.51114	
10	Infeasible	Infeasible		0.42488	2.532909	$t_2^* = 0.10805$	0.404706	1.512756	$t_2^* = 0.10805$
11	1.554315	1.554315		1.554315	1.554315		0.501842	0.501842	
12	1.478827	1.478827		1.478827	1.478827		0.681564	0.681564	
13	Infeasible	3.1794104	$\beta_1^* = 0.012285$	1.1664	3.168001	$t_3^* = 0.001648$	0.732485	1.734133	$t_3^* = 0.001648$
14	1.127483	1.127483		1.127483	1.127483		0.383954	0.383954	
15	1.376085	1.376085		1.376085	1.376085		0.431534	0.431534	

input gets input savings of 25.49864 ( $= t_2^* x_2^{\max} = 0.108045 \times 236$ ) compared with the factual second input amount of zero. That is the reason why DMU<sub>10</sub> is infeasible under model (1) and model (2). For DMU<sub>13</sub>, it gets output surplus of 5 ( $= \beta_1^* y_1^1 = 0.012285 \times 407$ ) under model (2), so it is infeasible under model (1). However, under model (3) and model (4), output surplus is eliminated and input saving of the third input with the value of 0.40368 ( $= t_3^* x_3^{\max} = 0.001648 \times 245$ ) comes into being. The result under model (4) with nonradial measurement is smaller than that of model (3) for each DMU. Using model (3) and model (4), the same amount of input saving or output surplus derived under both radial and nonradial measurement can also be seen in Table 5.

### 5. Conclusions

Based upon the previous work, the current paper extends radial super-efficiency model to the circumstances of nonradial measurement and establishes a modified

nonradial super-efficiency model. We find that the super-efficiency model with nonradial measurement still maintains some good properties and the original radial VRS super-efficiency model can also be detected whether it is infeasible through our nonradial super-efficiency model. In the actual production processes, due to underlying preferences and resource availability, super-efficiency model with nonradial measurement can help decision makers seek an efficient benchmark different from the reference under radial measurement. It is beneficial for decision makers to allocate input resources, make production plan, and arrange production activities. This paper only discusses the input-oriented nonradial model, for output-oriented situation, and it can be discussed in the same way.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.



## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work was partially supported by the Humanities and Social Science Youth Foundation of the Ministry of Education of China (Grant No. 20YJC630029); the Natural Science Foundation of Anhui Province of China (Grant No. 2008085MG228); and the Scientific Research Starting Foundation of Anhui Jianzhu University (Grant No. 2019QDZ09).

## References

- [1] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.
- [2] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in DEA," *Management Science*, vol. 30, no. 9, pp. 1078–1092, 1984.
- [3] W. W. Cooper, L. M. Seiford, and K. Tone, *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Kluwer Academic Publishers, Boston, MA, USA, 2000.
- [4] W. W. Cooper, L. M. Seiford, and J. Zhu, *Handbook on Data Envelopment Analysis*, Springer, Berlin, Germany, 2011.
- [5] M.-M. Yu, C.-C. Chern, and B. Hsiao, "Human resource rightsizing using centralized data envelopment analysis: evidence from Taiwan's Airports," *Omega*, vol. 41, no. 1, pp. 119–130, 2013.
- [6] A. Varmaz, A. Varwig, and T. Poddig, "Centralized resource planning and yardstick competition," *Omega*, vol. 41, no. 1, pp. 112–118, 2013.
- [7] F. Li, Q. Zhu, and Z. Chen, "Allocating a fixed cost across the decision making units with two-stage network structures," *Omega*, vol. 83, pp. 139–154, 2019.
- [8] Q. An, P. Wang, A. Emrouznejad, and J. Hu, "Fixed cost allocation based on the principle of efficiency invariance in two-stage systems," *European Journal of Operational Research*, vol. 283, no. 2, pp. 662–675, 2020.
- [9] D. Gong, S. Liu, and X. Lu, "Modelling the impacts of resource sharing on supply chain efficiency," *International Journal of Simulation Modelling*, vol. 14, no. 4, pp. 744–755, 2015.
- [10] Q. An, Y. Wen, T. Ding, and Y. Li, "Resource sharing and payoff allocation in a three-stage system: integrating network DEA with the Shapley value method," *Omega*, vol. 85, pp. 16–25, 2019.
- [11] P. Andersen and N. C. Petersen, "A procedure for ranking efficient units in data envelopment analysis," *Management Science*, vol. 39, no. 10, pp. 1261–1264, 1993.
- [12] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management Science*, vol. 30, no. 9, pp. 1078–1092, 1984.
- [13] L. M. Seiford and J. Zhu, "Infeasibility of super-efficiency data envelopment analysis models," *INFOR*, vol. 37, pp. 174–187, 1999.
- [14] C. A. K. Lovell and A. P. B. Rouse, "Equivalent standard DEA models to provide super-efficiency scores," *Journal of the Operational Research Society*, vol. 54, no. 1, pp. 101–108, 2003.
- [15] Y. Chen, "Measuring super-efficiency in DEA in the presence of infeasibility," *European Journal of Operational Research*, vol. 161, no. 2, pp. 545–551, 2005.
- [16] W. D. Cook, L. Liang, Y. Zha, and J. Zhu, "A modified super-efficiency DEA model for infeasibility," *Journal of the Operational Research Society*, vol. 60, no. 2, pp. 276–281, 2009.
- [17] H.-S. Lee, C.-W. Chu, and J. Zhu, "Super-efficiency DEA in the presence of infeasibility," *European Journal of Operational Research*, vol. 212, no. 1, pp. 141–147, 2011.
- [18] Y. Chen and L. Liang, "Super-efficiency DEA in the presence of infeasibility: one model approach," *European Journal of Operational Research*, vol. 213, no. 1, pp. 359–360, 2011.
- [19] R. M. Thrall, "Chapter 5 duality, classification and slacks in DEA," *Annals of Operations Research*, vol. 66, no. 2, pp. 109–138, 1996.
- [20] J. Zhu, "Robustness of the efficient DMUs in data envelopment analysis," *European Journal of Operational Research*, vol. 90, no. 3, pp. 451–460, 1996.
- [21] H.-S. Lee and J. Zhu, "Super-efficiency infeasibility and zero data in DEA," *European Journal of Operational Research*, vol. 216, no. 2, pp. 429–433, 2012.

## Research Article

# Research on the Identification of Key Nodes in the Process of WeChat Epidemic Information Dissemination: A Supernetwork Perspective

Peng Wu  and Di Zhao

*Institute of Management Science and Engineering, Henan University, Kaifeng 475004, China*

Correspondence should be addressed to Peng Wu; akasfd@126.com

Received 24 June 2020; Accepted 25 July 2020; Published 27 August 2020

Guest Editor: Shaohui Wang

Copyright © 2020 Peng Wu and Di Zhao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is of great significance to comprehensively and reasonably identify the key nodes in the WeChat epidemic information dissemination system. First, the recognition results can be used to guide the spread of epidemic-related information in WeChat, such as accelerating the spread of valuable information or monitoring the spread of rumors. Secondly, the analysis of key nodes helps us understand the evolution of the epidemic information dissemination network in WeChat, and the analysis of key nodes also helps us understand the modes and methods of epidemic-related information dissemination in WeChat. Finally, the results of these studies may be generalized to other fields of social life. This paper analyzes the composition of and relationship between epidemic-related information dissemination systems in WeChat and proposes a Supernetwork model for WeChat epidemic information dissemination on this basis. In this study, a comprehensive identification method of key nodes of the WeChat epidemic information dissemination system under the Supernetwork vision was constructed, and the method is analyzed and verified through examples in this paper.

## 1. Introduction

With the rapid development of information technology and the increasingly complex social system, the current information dissemination related to the epidemic situation is also becoming more and more rapid. If the spread of relevant public opinion or rumors in a network cannot be promptly understood and effectively guided and managed, the spread of its outbreak will have catastrophic consequences. Therefore, it is necessary to conduct research and analysis on the spread of information related to the epidemic situation [1].

Tencent released the “2019 WeChat Annual Data Report” in January 2020. The report stated that in 2019, the number of WeChat monthly active accounts exceeded 11 billion [2]. With such a huge monthly active user scale, WeChat’s role in the spread of epidemic information cannot be underestimated.

After the outbreak, WeChat, as a highly active and highly connected medium, has become an important channel for the public to understand and participate in the epidemic. Scholar Liu [3] used the framework analysis theory to systematically analyze the top 200 WeChat posts using the Qingbo Index over a period of time and found that the top posts on WeChat during the epidemic mainly showed the following characteristics. In terms of media report types, both self-media and traditional media commonly guide the network public opinion pattern. In terms of the evolution of public opinion, online public opinion events frequently occur and advance in waves. In the news framework, the response framework and the emotional framework have the highest public acceptance and recognition. Through the above research, Liu proposed that the guidance of public opinion on the epidemic should strengthen the “strength” of rational dissemination, pay attention to the “temperature” of humanistic care, and improve the “validity” of public

opinion guidance. Based on a full-sample empirical analysis of 2966 official WeChat public account articles from 42 “double first-class” universities, Deng et al. [4] found that article topics, information sources, and emotional tendencies have a significant impact on communication effects. Public opinion propaganda with a richer connotation is more popular. The original works of universities’ WeChat public accounts are more influential. Public opinion propaganda with obvious “positive” or “negative” sentimental tendencies has a significant impact on the communication effect but in different ways. Readers who read short and concise articles are more likely to keep reading until the end of the article and like the article. Chen [5] found that rumors showed a rapid growth trend in the early stages of the outbreak. The formation and spread of rumors mainly occurs on WeChat and Weibo, and the spread of rumors is directly related to the degree of government information disclosure. Due to the strong relationship and semienclosed space feature of WeChat, WeChat rumors are mainly spread through public accounts, group chats, private chats, and circles of friends. They have both interpersonal and group transmission characteristics and are extremely infectious. As the number of confirmed diagnoses across the country began to decline, and some cities began to resume work, the government continued to disclose information about the epidemic in an open and transparent manner, and related rumors began to decline. During this period, the government responded quickly to rumors, and there gradually became fewer “unverified” rumors. At the same time, the spread of strong relationships in WeChat groups allowed the public to receive a large amount of factual splicing from relatives and friends in different spaces, which eliminated rumors to a great extent. The ambiguity in the communication model improved the public’s judgment, and the rumors lost their foundation for spreading. Wan et al. [6] used the case study method to establish the status quo of the “fitness anti-epidemic tweets” pushed by the 117 WeChat subscription accounts they followed and found that their main features are the time and frequency of article posting, which are consistent with the changes in the number of confirmed cases, and the overall national prevention and control deployment. The online dissemination of content is comprehensive, and the amount of information is gradually decreasing with the effective relief of the epidemic. By analyzing popular words in WeChat, Lu and Zhang [7] found that the development of the epidemic had a certain pattern. First of all, people paid attention to which pathogen the disease comes from, the location of the outbreak, and what symptoms appear. Then, the public began to pay attention to the source of pathogens, the behavior and self-protection of medical staff, and the needs of daily life. On the other hand, people paid more attention to the epidemic situation in other places and the results of local control. Through the analysis of public concern, we reviewed the development trend of the epidemic, which will set an example for future epidemics. It is believed that public health authorities will rely more on these social media to monitor the development of epidemics or pandemics in the future. Judging from the existing related results, there are few studies in this field, and

the use of Supernetwork theory to study the spread of epidemic information in WeChat is still rare.

This article studies and analyzes the composition and relationship between the epidemic-related information dissemination systems in WeChat. On this basis, three types of active and representative elements are established, namely, ordinary WeChat users, WeChat groups, and WeChat public accounts. The three epidemic information dissemination network relationships are combined and integrated into a Supernetwork.

## 2. Supernetwork Theory and WeChat Supernetwork Model Construction

Scholars previously used the term “Supernetwork” in the fields of computer systems, genetics, and transportation systems to generally refer to complex systems with complicated nodes, complicated relationships, and networks within them. In the American scholar Nagurney’s work on interwoven networks, the network that is higher than and surpasses the existing network is referred to as the “Supernetwork,” which helps to clearly define this term [8, 9]. Chinese scholars, Academician Wang Zhongtuo and Professor Wang Zhiping, further proposed the relevant characteristics of the Supernetwork. They believe that the Supernetwork has one or more of the following characteristics: (1) The network contains the network, or it can be said that the network is nested in the network. (2) It has the characteristics of multiple layers, and there are connections within and between layers. For example, the protocol of the information network is multilayered, and the transportation network has a management layer, a business layer, and a physical layer, which are also multilayered. (3) It has the characteristics of multiple levels, and there is a connection between these levels. For example, the information network of an enterprise has multiple levels such as headquarters, companies, and departments. (4) Its flow can have multidimensional characteristics. For example, aviation, roads, water transportation, and railways all have both freight and passenger transportation. (5) It has the characteristics of multiple attributes or multiple criteria. For example, the transportation network needs to consider factors such as cost, safety, time, and comfort at the same time, and when traveling in a city, there are not only travel modes (such as self-driving, public transportation, or walking) but also route options. (6) It has the characteristics of congestion. For example, the transportation network has the problem of congestion, as does the information network. (7) Sometimes, there is a conflict between individual optimization and global optimization, which requires coordination [9].

Recent research on Supernetworks can be divided into three directions: based on variational inequalities [10–12], the work mainly intends to transform the multilayered, multistandard Supernetwork balance model into an optimization problem and then use evolutionary variation inequality to solve the problem [9]. The system science [13–16] research studies the Supernetwork as a whole as well as individual aspects, including the study of the relationship between the network and the network in the Supernetwork,

the study of the network using the relationship between the outside and the network, and the study of overall performance. The modeling work of this article is mainly inspired and helped by some research results in this area. Research based on hypergraphs [17] is considered to be another more important research direction. It is mostly used to solve the problem of different textures between edges and nodes in network research. Hypergraphs are used to describe the system and study the statistical characteristics of the network. The research on Supernetworks has developed rapidly, and its application fields are also relatively wide, such as knowledge management, supply chains, transportation, financial management, advertising, e-commerce, biology, and organization and cooperation [18]. However, Supernetwork theory and methods are rarely applied to the field of WeChat epidemic information dissemination.

WeChat public accounts, common WeChat users, and WeChat groups are the main outlets through which epidemic information is spread in WeChat. The public account mainly disseminates information through the weak relationship between the public account and the fans. The public account and the fans form a star network. Relatively speaking, the communication speed is fast and the spread is wide, but the audience trusts the public account to release information, though this is not the most trusted source of information.

The ordinary users of WeChat mainly spread information through the strong relationship between ordinary users' friends. The information publisher and information receiver in the circle of friends also form a star network. Because of the relationships between acquaintances, the information posted in the circle of friends is related to the circle of friends. Members tend to believe this information more, but the information in the circle of friends is limited by the border barriers of the circle of friends. On the whole, there is a large gap between the speed of propagation and the range of communication compared with the public account.

The WeChat group has special properties. The WeChat group can take into account both strong and weak relationships for propagation; that is, members in the WeChat group may not be known or trusted, and members of the WeChat group may or may not be influential. This can break the barriers of the circle of friends based on strong relationships, so that the spread of epidemic-related information is wider and faster. The degree to which the recipient of the information believes the information is based on these factors.

In summary, this article integrates these three elements into a "Supernetwork" of WeChat epidemic information dissemination systems based on the association between the public accounts, common WeChat users and WeChat groups, referred to as the WeChat epidemic information dissemination Supernetwork.

**2.1. Supernetwork Modeling of WeChat Epidemic Information Dissemination.** From the previous analysis, we can see that there are three types of epidemic information dissemination networks in Figure 1:

- (1) The P-P common WeChat user relationship network's nodes are common WeChat users. The network can be expressed as  $G_p = (P, E_{p-p})$ , where  $P = \{p_1, p_2, \dots, p_b\}$  is a set of common WeChat user relationship network members;  $E_{p-p} = \{(p_i, p_j) \mid p_i, p_j \in P\}$  is a set of edges; and  $(p_i, p_j)$  is a common WeChat user. There is an association between  $p_i$  and  $p_j$ .
- (2) The M-M WeChat group network is the network constructed by the WeChat group. The nodes are WeChat groups, and the edges represent the association between WeChat groups. Its model is  $G_m = (M, E_{m-m})$ ;  $M = \{m_1, m_2, \dots, m_a\}$  is a collection of WeChat groups;  $E_{m-m} = \{(m_i, m_j) \mid m_i, m_j \in M\}$  is a collection of association relationships between WeChat groups; and edge  $(m_i, m_j)$  indicates that there is an association between WeChat groups  $m_i$  and  $m_j$ .
- (3) The I-I WeChat public account network is formed between public accounts. The node is a WeChat public account. The network can be described as  $G_i = (I, E_{i-i})$ ;  $I = \{i_1, i_2, \dots, i_c\}$  is a collection of public accounts; and  $E_{i-i} = \{(i_i, i_j) \mid i_i, i_j \in I\}$  is a collection of associations between public accounts.  $(i_i, i_j)$  indicates that there is an association between  $i_i$  and  $i_j$ .

The following mapping relationships exist between the three networks:

- (1) The mapping between WeChat general users and WeChat public accounts refers to which ordinary WeChat users pay attention to which public accounts and which common WeChat users follow which public accounts.
- (2) The mapping between WeChat groups and public accounts refers to which public accounts are associated with which WeChat groups.
- (3) The mapping between ordinary WeChat users and WeChat groups refers to which ordinary WeChat users join or are added to which WeChat groups, which WeChat groups are associated, and which WeChat groups include which ordinary WeChat users.

According to the mapping relationship described above, a mapping relationship can be established between  $G_i, G_p, G_m$  nodes so that they are integrated together to construct a Supernetwork with three types of nodes. The specific approach is based on the three networks that already exist, and the mapping relationship between different types of nodes is also added as an edge, thereby integrating three different types of networks together. This process can be expressed as follows.

Let  $p_i \in P$ ,  $i_j \in I$ , and  $m_l \in M$ , respectively, represent any node of the WeChat general user association network, WeChat public account network, and WeChat group association network. The Boolean variables  $\phi(p_i, i_j)$ ,  $\varphi(p_i, m_l)$ , and  $\gamma(i_j, m_l)$  indicate whether the different types of nodes

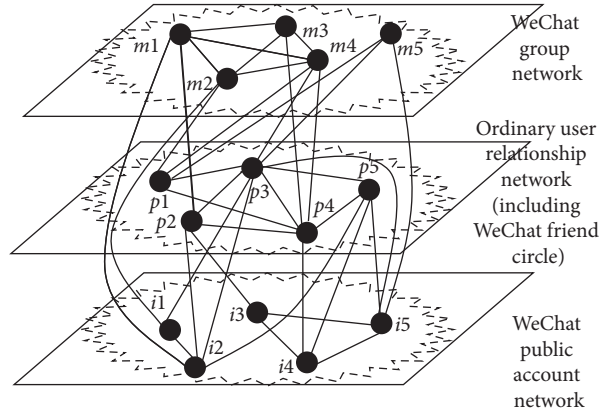


FIGURE 1: WeChat epidemic information dissemination Supernetwork.

are related, where the value 1 is time-dependent, and 0 is not relevant. Thus,

$$\begin{aligned}
 \text{WEIDSN} &= f(G_i, G_p, G_m) \\
 &= G_i + G_p + G_m + E_{p-i} + E_{p-m} + E_{m-i} \\
 &= (I, P, M, E_{I-I}, E_{P-P}, E_{M-M}, E_{p-i}, E_{p-m}, E_{m-i}),
 \end{aligned} \tag{1}$$

here,  $E_{p-i} = \{(p_i, i_j) | \phi(p_i, i_j) = 1\}$  represents a set of association relationships between P and I nodes;  $E_{p-m} = \{(p_i, m_l) | \varphi(p_i, m_l) = 1\}$  represents a collection of association relationships between P and M nodes; and  $E_{m-i} = \{(i_j, m_l) | \gamma(i_j, m_l) = 1\}$  represents a collection of associations between I and M nodes.

It can be seen that the WeChat Supernetwork model is an integration of three epidemic information dissemination networks, which reveals the complex composition and structural form of epidemic information dissemination in WeChat.

**2.2. Construction of Adjacency Matrix and Network Intuitive Graph of WeChat Epidemic Information Supernetwork Model.** At this stage, we can use network analysis tools to measure various parameters of network analysis, and network relationship characterization software can draw network graphs for intuitive comparison.

Figure 1 is a schematic diagram of a three-network integration Supernetwork of a WeChat epidemic information dissemination system. The figure is divided into three layers. The top layer is the WeChat group network. The network members are WeChat groups. There may be associations between the network members, and the network members may also be associated with the members of the other two layers.  $m_i$  represents the nodes in this layer network. The middle layer of the network is the ordinary WeChat user relationship network. The network members are ordinary WeChat users. There may be associations between the network members, and the network members may also be associated with members of the other two layers.  $p_i$  represents nodes in this layer of the network. The bottom layer is the WeChat official account network. The network

members are WeChat official accounts. There may be associations between the network members, and the network members may also be associated with members of the other two layers of networks.  $i_i$  represents nodes in this layer network. Figure 1 is a schematic diagram of a WeChat epidemic information dissemination Supernetwork, in which ordinary WeChat user  $p_2$  joins WeChat group  $m_1$  and ordinary WeChat users  $p_3$  and  $p_4$  join WeChat groups  $m_4$  and  $m_5$ . After joining the WeChat groups  $m_3$  and  $m_4$ , the ordinary WeChat user  $p_1$  is added to the three WeChat groups  $m_2$ ,  $m_4$ , and  $m_5$  at the same time, and the ordinary WeChat user  $p_5$  is temporarily not added to any WeChat group. WeChat groups  $m_1$ ,  $m_2$ , and  $m_5$  are associated with the public account network; ordinary WeChat user  $p_2$  pays attention to the public accounts  $i_2$  and  $i_3$ , ordinary user  $p_3$  pays attention to the public accounts  $i_1$ ,  $i_2$ , and  $i_5$  at the same time, and ordinary user  $p_4$  pays attention to the public account  $i_4$ , along with ordinary user  $p_5$ . After paying attention to the public accounts  $i_2$ ,  $i_4$ , and  $i_5$ , the ordinary WeChat user  $p_1$  does not follow any WeChat public account for the time being. Taking the WeChat epidemic information dissemination Supernetwork as an example, the boundary of the epidemic information dissemination subnetwork of the WeChat epidemic information dissemination system is determined, and its many network nodes and their associations are established. The rules for constructing the association relationship of network nodes are as follows. If the two nodes in the schematic diagram of the WeChat epidemic information dissemination Supernetwork are connected, they are marked as 1; otherwise, they are marked as 0. Using network analysis software, the adjacency matrix of the WeChat epidemic information dissemination Supernetwork can be constructed as shown in Figure 2.

At the same time, the corresponding network analysis standard file can be generated and imported into the specified network relationship characterization software, which can be processed according to specific needs and purposes and used to then construct a network intuitive map of the WeChat epidemic information dissemination Supernetwork, as shown in Figure 3.

The representative WeChat epidemic information dissemination Supernetwork model outlined in this paper is a

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	<i>m1</i>	<i>m2</i>	<i>m3</i>	<i>m4</i>	<i>m5</i>	<i>p1</i>	<i>p2</i>	<i>p3</i>	<i>p4</i>	<i>p5</i>	<i>i1</i>	<i>i2</i>	<i>i3</i>	<i>i4</i>	<i>i5</i>
1 <i>m1</i>	0.000	1.000	1.000	1.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
2 <i>m2</i>	1.000	0.000	1.000	1.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
3 <i>m3</i>	1.000	1.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
4 <i>m4</i>	1.000	1.000	1.000	0.000	0.000	1.000	0.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
5 <i>m5</i>	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
6 <i>p1</i>	0.000	1.000	0.000	1.000	1.000	0.000	0.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
7 <i>p2</i>	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	0.000	1.000	1.000	0.000
8 <i>p3</i>	0.000	0.000	0.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000	1.000	0.000	0.000	1.000
9 <i>p4</i>	0.000	0.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000	1.000	0.000	0.000	0.000	1.000	0.000
10 <i>p5</i>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	0.000	1.000	0.000	1.000	1.000
11 <i>i1</i>	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
12 <i>i2</i>	1.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	1.000	1.000	0.000	0.000	0.000	0.000
13 <i>i3</i>	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
14 <i>i4</i>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	0.000	1.000	0.000	1.000
15 <i>i5</i>	0.000	0.000	0.000	0.000	1.000	0.000	0.000	1.000	0.000	1.000	0.000	0.000	1.000	1.000	0.000

FIGURE 2: The adjacency matrix of the WeChat epidemic information Supernetwork.

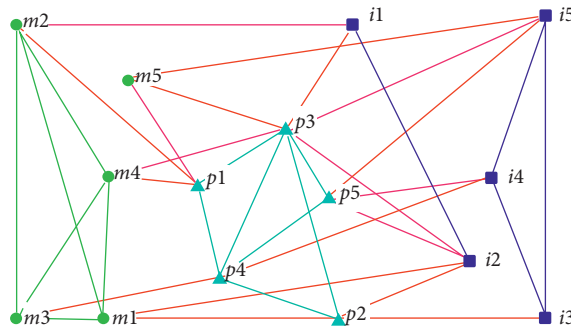


FIGURE 3: Intuitive map of the WeChat epidemic information dissemination Supernetwork.

network of three different types of epidemic information dissemination networks, including the public account network, WeChat group network, and common WeChat user relationship network. The Supernetwork model contains three relatively active and representative information carrier elements in the WeChat epidemic information dissemination system, such as WeChat public accounts, WeChat groups, common WeChat users, and their complex associations. Compared with the epidemic information dissemination network composed of similar nodes and their associations, the Supernetwork model can better reflect the complex composition and structural form of the WeChat epidemic information dissemination network system, so it can be more accurate, comprehensive and provide more in-depth WeChat epidemic information. The communication network system conducts research and analysis, lays a foundation, and provides help for the management and guidance of subsequent WeChat epidemic information dissemination.

### 3. Identification of Key Nodes in WeChat Epidemic Information Dissemination Supernetwork

An important part of guiding the spread of epidemic-related information in WeChat is to first identify the key nodes in the spread of WeChat epidemic information. These

important related nodes largely affect the spread of epidemic-related information in WeChat.

3.1. *Traditional Key Node Identification Method.* The degree centrality method in network analysis theory is the main method for identifying key nodes in a network. The formula is as follows:

$$C_D(n_i) = d(n_i) = \sum_j X_{ij} = \sum_j X_{ji},$$

$$C'_D(n_i) = \frac{d(n_i)}{g-1}. \tag{2}$$

The value of  $X_{ji}$  is 1 or 0, indicating whether node  $j$  is associated with node  $i$ , and  $g$  is the number of nodes in the network. In the network, because the number of nodes in each network is different, and the number of large network nodes is large, the standardization process must be divided by the maximum possible number of relationships of a node in the network, that is,  $g - 1$  relationships [19, 20]. Of course, in addition to the degree centrality method, there are many methods for identifying network central nodes, but the degree centrality method is relatively simple and straightforward, and the application scenarios of this method are also broader, with fewer restrictions on implementation [21]. The network analysis tool was used to calculate the degree centrality of all nodes in the Supernetwork, as shown in Figure 3, and the system calculation results are as follows.



From the results in the above figure, we can see that the degree of centrality of node  $p3$  is 9, the degree of centrality of  $p4$  is 7, and the standardization of the two nodes is 0.64286 and 0.5, respectively. It can be seen that the  $p3$  node and the  $p4$  node are relatively important. During the spread of WeChat epidemic-related information, monitoring and guiding these two nodes has a great impact on the overall information spread.

**3.2. Key Node Recognition Method Based on WeChat Supernetwork.** Degree centrality is currently one of the most commonly used key node identification methods, but its application in a Supernetwork has deficiencies. The degree centrality calculation here does not distinguish which subnet each node belongs to in the Supernetwork, there is no distinction between the relationships between nodes, and there is no difference between the relationships between subnets. Accordingly, this method cannot reflect the connection status of a node in the subnet to which it belongs, nor can it reflect the propagation ability of the node in the subnet to which it belongs; similarly, in the current situation, this method also fails to reflect the connection status of a node between subnetworks, and it fails to measure the propagation ability of the node between subnetworks. Referring to Figures 3 and 4, for example, the  $i1$  node has a degree of centrality of

$$C_D(i1) = d(i1) = \sum_j X_{ij} = \sum_j X_{ji} = 3. \quad (3)$$

And the  $m3$  node has a degree of centrality of

$$C_D(m3) = d(m3) = \sum_j X_{ij} = \sum_j X_{ji} = 4. \quad (4)$$

According to the formula,

$$C'_D(n_i) = \frac{d(n_i)}{g-1}. \quad (5)$$

The standardized degree of centrality of  $i1$  can be obtained as

$$C'_D(i1) = \frac{d(i1)}{g-1} = 0.21429. \quad (6)$$

And the standardized degree of centrality of node  $m3$  is

$$C'_D(m3) = \frac{d(m3)}{g-1} = 0.28571. \quad (7)$$

Thus, node  $i1$  is not as important as  $m3$ , but in the process of epidemic information propagation in WeChat, in the Supernetwork  $i1$  is associated with three unary subnets, while  $m3$  is only associated with the M-M network and P-P network and is not directly connected to the I-I network. From this perspective,  $i1$  is more important than  $m3$ . The spread of epidemic information in WeChat has unique features. For example, the ability of ordinary WeChat users to spread information in the circle of friends is related to the number of their friends, the ability of the public account to spread information in WeChat is related to the number of

fans, and the ability to spread information in WeChat groups and the number of WeChat group members is related. The nodes with strong propagation ability are more important, and the identification methods of these key nodes in WeChat need to be rebuilt.

In order to avoid the above deficiencies, this paper proposes a new comprehensive method for identifying key nodes in WeChat epidemic information dissemination based on a Supernetwork, part of which is based on an improvement of the decision method proposed previously [13], and the other part is mainly constructed based on the characteristics of the spread of epidemic information in WeChat. The new method is briefly described as follows.

Let  $n_i$  be any node in the Supernetwork;  $QC_D(n_i)$  represents the degree of centrality of  $n_i$  in the entire Supernetwork, and  $ZC_D(n_i)$  is the degree of centrality of  $n_i$  in the unary subnet to which it belongs:

- (1) *Step 1.* Calculate and compare the degree of centrality of each node in the entire Supernetwork.
- (2) *Step 2.* Calculate and compare the degree centrality of all nodes in the M-P binary subnetwork and the I-P binary subnetwork in the M-P binary subnetwork and I-P binary subnetwork to which they belong.
- (3) *Step 3.* Calculate and compare the degree centrality of all the individuals in the entire Supernetwork system in the subunit networks to which they belong.
- (4) *Step 4.* Calculate and compare the internetwork propagation capabilities of each node in the subnet of a unitary network in the Supernetwork. Let  $LZS(n_i)$  be the number of unitary subnetworks connected by node  $n_i$  and  $ZJBS(n_i)$  the edges of node  $n_i$  between the unitary subnetworks. The number can be obtained from the above analysis.  $ZJBS(n_i) = QC_D(n_i) - ZC_D(n_i)$ , the network propagation ability of the node  $n_i$  in the Supernetwork is expressed by  $ZJCN(n_i)$ , and the calculation and comparison methods are as follows:

```

Select Case LZS(A)
  Case IS > LZS(B)
    ZJCN(A) > ZJCN(B)
  Case IS < LZS(B)
    ZJCN(A) < ZJCN(B)
  Case IS = LZS(B)
    Select Case ZJBS(A)
      Case IS > ZJBS(B)
        ZJCN(A) > ZJCN(B)
      Case IS < ZJBS(B)
        ZJCN(A) < ZJCN(B)
      Case IS = ZJBS(B)
        ZJCN(A) = ZJCN(B)
    End Select
End Select

```

		1	2	3
		Degree	NrmDegree	Share
8	<i>p3</i>	9.000	64.286	0.122
9	<i>p4</i>	7.000	50.000	0.095
4	<i>m4</i>	6.000	42.857	0.081
10	<i>p5</i>	5.000	35.714	0.068
12	<i>i2</i>	5.000	35.714	0.068
6	<i>p1</i>	5.000	35.714	0.068
7	<i>p2</i>	5.000	35.714	0.068
1	<i>m1</i>	5.000	35.714	0.068
2	<i>m2</i>	5.000	35.714	0.068
15	<i>i5</i>	5.000	35.714	0.068
3	<i>m3</i>	4.000	28.571	0.054
14	<i>i4</i>	4.000	28.571	0.054
13	<i>i3</i>	3.000	21.429	0.041
11	<i>i1</i>	3.000	21.429	0.041
5	<i>m5</i>	3.000	21.429	0.041

FIGURE 4: The degree of centrality of all nodes in the WeChat epidemic information Supernetwork.

- (5) *Step 5.* Calculate and compare the internetwork propagation capability  $ZJCN_J(n_i)$  between the M-type nodes in the M-P binary subnetwork and the I-type nodes in the I-P binary subnetwork under the corresponding binary subnetwork view. Let  $J_{M-P}C_D(n_i)$  represent the degree of centrality of  $n_i$  in the M-P binary subnetwork to which it belongs and let  $J_{I-P}C_D(n_i)$  represent the degree of centrality of  $n_i$  in the I-P binary subnetwork to which it belongs, and the calculation method of  $ZJCN_J(n_i)$  is as follows:

Select Case  $n_i$

Case  $IS \in M$

$$ZJCN_J(n_i) = J_{M-P}C_D(n_i) - ZC_D(n_i)$$

Case  $IS \in I$

$$ZJCN_J(n_i) = J_{I-P}C_D(n_i) - ZC_D(n_i)$$

End Select

- (6) Combining the results of the above steps, key nodes are identified.

In the above steps, Step 2 prepares for the calculation and comparison of the number of fans of the public account, and the number of members in the WeChat group is determined in Step 5. Step 3, the calculation result, can also obtain the number of friends of ordinary WeChat users in the circle of friends. The circle of friends is one of the main channels for ordinary users to disseminate information in WeChat. The number of friends in the circle of friends is an important criterion for identifying the ability of ordinary WeChat users to disseminate information. Step 4 compares the unary subnet internetwork propagation capabilities of each node in the Supernetwork view by the number of unary subnetworks connected by the nodes and the number of nodes connected between the unary subnetworks. A ranking table was created showing the strengths and weaknesses of the unary subnet's internetwork spreading ability under the Supernetwork vision. Among them,  $QC_D(n_i)$  can be obtained by Step 1,  $ZC_D(n_i)$  can be obtained by Step 3, and the individual's  $ZJBS(n_i)$  can be obtained by the difference between  $QC_D(n_i)$  and  $ZC_D(n_i)$  of an individual. Firstly, it is judged based on the number of unary subnetworks connected to the node.

The number of unary subnetworks connected to node A is greater than the number of unary subnetworks connected to node B, which indicates that the internetwork communication capability of node A is stronger than that of node B. If the number of unary subnetworks connected to node A is less than the number of unary subnetworks connected to node B, it indicates that the internetwork propagation capability of node A between unary subnetworks is weaker than that of node B. If the number of unary subnetworks connected to node A is the same as the number of unary subnetworks connected to node B, further judgments need to be made based on the number of edges between the unary subnetworks of nodes. Step 5 is mainly based on the calculation results of Step 2 and Step 3 and calculates and compares the number of fans of the public account and the number of members of the WeChat group in the WeChat Supernetwork. The information dissemination ability of nodes is measured to provide a reference for the identification of key nodes. This step calculates the internetwork propagation capability between the M-type nodes in the M-P binary subnetwork and the I-type nodes in the I-P binary subnetwork from the binary subnetwork perspective.

Taking the example shown in Figure 3 for verification and analysis:

- (1) *Step 1.* The calculation result is shown in Figure 4.
- (2) *Step 2.* The screenshots of the system calculation results are shown in Figures 5 and 6.
- (3) *Step 3.* The result is shown in Figure 7.
- (4) *Step 4.* The results are shown in Table 1.
- (5) Combining the comparison and calculation results of Step 2 and Figures 5 and 6 with the comparison and calculation results of Step 3 and Figure 7, the calculation and comparison results of this step are shown in Tables 2 and 3.
- (6) From the perspective of intranetwork propagation, Figure 7 shows that the degree of centrality of *i3*, *i4*, and *i5* in the I-I network is the same and the highest, and the propagation ability of *i3*, *i4*, and *i5* in the network is also the strongest. Nodes *m1*, *m2*, *m3*, and

		1	2	3
		Degree	NrmDegree	Share
3	<i>p3</i>	7.000	77.778	0.175
4	<i>p4</i>	5.000	55.556	0.125
5	<i>p5</i>	5.000	55.556	0.125
2	<i>p2</i>	4.000	44.444	0.100
7	<i>i2</i>	4.000	44.444	0.100
10	<i>i5</i>	4.000	44.444	0.100
9	<i>i4</i>	4.000	44.444	0.100
8	<i>i3</i>	3.000	33.333	0.075
1	<i>p1</i>	2.000	22.222	0.050
6	<i>i1</i>	2.000	22.222	0.050

FIGURE 5: The degree of centrality of all nodes in the I-P network.

		1	2	3
		Degree	NrmDegree	Share
8	<i>p3</i>	6.000	66.667	0.143
4	<i>m4</i>	6.000	66.667	0.143
9	<i>p4</i>	6.000	66.667	0.143
6	<i>p1</i>	5.000	55.556	0.119
2	<i>m2</i>	4.000	44.444	0.095
1	<i>m1</i>	4.000	44.444	0.095
3	<i>m3</i>	4.000	44.444	0.095
7	<i>p2</i>	3.000	33.333	0.071
5	<i>m5</i>	2.000	22.222	0.048
10	<i>p5</i>	2.000	22.222	0.048

FIGURE 6: The degree of centrality of all nodes in the M-P network.

		1	2	3
		Degree	NrmDegree	Share
1	<i>m1</i>	3.000	75.000	0.250
2	<i>m2</i>	3.000	75.000	0.250
3	<i>m3</i>	3.000	75.000	0.250
4	<i>m4</i>	3.000	75.000	0.250
5	<i>m5</i>	0.000	0.000	0.000

		1	2	3
		Degree	NrmDegree	Share
3	<i>p3</i>	4.000	100.000	0.286
4	<i>p4</i>	4.000	100.000	0.286
1	<i>p1</i>	2.000	50.000	0.143
2	<i>p2</i>	2.000	50.000	0.143
5	<i>p5</i>	2.000	50.000	0.143

		1	2	3
		Degree	NrmDegree	Share
3	<i>i3</i>	2.000	50.000	0.250
4	<i>i4</i>	2.000	50.000	0.250
5	<i>i5</i>	2.000	50.000	0.250
2	<i>i2</i>	1.000	25.000	0.125
1	<i>i1</i>	1.000	25.000	0.125

FIGURE 7: The degree of centrality of each node in the unary subnet to which it belongs.

*m4* have strong propagation capabilities in the M-M network. In the P-P network, *p3* and *p4* are the strongest nodes in the network. From the perspective of internetwork propagation, Table 1 shows that the number of subnetworks connected by nodes *p3* and *i2* is larger, and the number of interconnected edges between networks is also larger, so their internetwork propagation capabilities are stronger. The degree centrality value of such nodes may not be

large, but its influence on the propagation effect is often large, which is worthy of attention. It can also be seen from Table 1 that although *m3* has the strongest propagation capability in the M-M network, and *i1* has the weakest propagation capability in the I-I network in Figure 7, from the perspective of internetwork propagation, *i1* connects three subnets, there are 2 internetwork edges, *m3* only connects 2 subnets, and there is only 1 internetwork edge, so *i1*

TABLE 1: Sequence table of node's internetwork epidemic information dissemination ability.

Node	Number of unary subnets associated	Number of relationships
<i>p3</i>	3	5
<i>i2</i>	3	4
<i>p2</i>	3	3
<i>p4</i>	3	3
<i>i5</i>	3	3
<i>m1</i>	3	2
<i>i1</i>	3	2
<i>m2</i>	3	2
<i>m5</i>	2	3
<i>p1</i>	2	3
<i>m4</i>	2	3
<i>p5</i>	2	3
<i>i4</i>	2	2
<i>m3</i>	2	1
<i>i3</i>	2	1

TABLE 2: Sorting table of M-type nodes' epidemic information dissemination capability in WeChat.

Node	Degree of centrality within the M-P network	Degree centrality within the M-M network	Important indicators of WeChat epidemic information dissemination capability
<i>m4</i>	6	3	3
<i>m5</i>	2	0	2
<i>m1</i>	4	3	1
<i>m2</i>	4	3	1
<i>m3</i>	4	3	1

TABLE 3: Sorting table of the spreading capacity of epidemic information in WeChat for Class I nodes.

Node	Degree of centrality within the I-P network	Degree of centrality within the I-I network	Important indicators of WeChat epidemic information dissemination capability
<i>i2</i>	4	1	3
<i>i4</i>	4	2	2
<i>i5</i>	4	2	2
<i>i3</i>	3	2	1
<i>i1</i>	2	1	1

is slightly better than *m3* in terms of internetwork communication capability. From the perspective of the spread of epidemic-related information in WeChat, Figure 7 shows that *p3* and *p4* have a strong ability to spread epidemic-related information in the WeChat circle of friends. Under the same conditions, the spread of epidemic-related information of these two ordinary user nodes is faster. The spreading range is wide; Table 2 shows that *m4* has a strong spreading capability in WeChat group nodes. Under the same conditions, the information spreading speed of this WeChat group node is faster, and the spreading range is wider. Table 3 shows *i2* in the public account category. The dissemination capability of the node is strong. Under the same conditions, the public information node's epidemic-related information spreads faster more widely. To guide the dissemination of epidemic-related information in WeChat, it is best to guide the

abovementioned WeChat nodes with a strong ability to disseminate epidemic-related information.

The new method also identifies the importance of ordinary WeChat users, WeChat groups, and official accounts from the perspective of information diffusion based on the particularity of the spread of WeChat epidemic-related information, in order to be more comprehensive and reasonable.

#### 4. Conclusions

This article mainly concerns the model construction and key node identification of the WeChat epidemic-related information dissemination Supernetwork. The identification method can be roughly divided into two parts. One part is based on the improvement of the decision method proposed previously [13], and the other part of it is mainly based on the characteristics of the spread of information related to the

epidemic in WeChat. The Supernetwork model of WeChat epidemic information dissemination and the corresponding key node identification method constructed in this study are the initial attempts to research the information dissemination of epidemic-related information in WeChat.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

### Authors' Contributions

Peng Wu and Di Zhao contributed to the design of the study. Peng Wu gave the main idea of the manuscript, wrote the first draft, and performed the design of figures. All authors contributed to the manuscript revision and read and approved the submitted version.

### Acknowledgments

This work was supported by the National Social Science Fund Project (18BGL238), China; the China Postdoctoral Science Foundation Special Funding Project (2018T110724), China; the Chinese Postdoctoral Science Fund General Support Project (2016M602236), China; and the Henan Philosophy and Social Science Planning Project (2017BJJ021); and Key Research Project of Teacher Education Curriculum Reform in Henan Province (2018-JSJYZD-008), China.

### References

- [1] Science Net, "Mining social network opinion leaders," 2020, <http://blog.sciencenet.cn/home.php?mod=space&uid=3075&do=blog&id=466991>.
- [2] Surging News, "The 2019 WeChat annual data report is out!," 2020, [https://www.thepaper.cn/newsDetail\\_forward\\_5480482](https://www.thepaper.cn/newsDetail_forward_5480482).
- [3] L. L. Liu, "Research on new coronary pneumonia epidemic reports from the perspective of the framework," *Journalism Lover*, vol. 5, pp. 17–21, 2020.
- [4] Z. Deng, X. J. Liu, and Z. D. Kang, "Research on the public opinion propaganda and ideological education function of university new media in the prevention and control of major epidemics," *Studies in Ideological Education*, no. 3, pp. 38–43, 2020.
- [5] Y. S. Chen, "Research on spreading and governance of internet rumors on public health emergency," *E-Government*, no. 6, pp. 2–11, 2020.
- [6] B. J. Wan, C. Guo, Z. Zhao et al., "Characteristics and risks reflections of "anti-epidemic fitness messages" delivered by WeChat during the COVID-19 epidemic," *Journal of Xi'an Physical Education University*, vol. 37, no. 3, pp. 267–275, 2020.
- [7] Y. Lu and L. Zhang, "Social media WeChat infers the development trend of COVID-19," *Journal of Infection*, no. 81, pp. 82–83, 2020.
- [8] A. Nagurney and J. Dong, *Supernetworks: Decision-Making for the Information Age*, Edward Elgar Publishing, Cheltenham, UK, 2002.
- [9] Z. P. Wang and Z. T. Wang, *Supernetwork Theory and its Application*, Science Press, Beijing, China, 2008.
- [10] A. Nagurney, "Optimal supply chain network design and redesign at minimal total cost and with demand satisfaction," *International Journal of Production Economics*, vol. 128, no. 1, pp. 200–208, 2010.
- [11] A. Nagurney, M. Yu, and Q. Qiang, "Supply chain network design for critical needs with outsourcing," *Papers in Regional Science*, vol. 90, no. 1, pp. 123–142, 2011.
- [12] L. Zhu and J. Cao, "Supernetwork optimization of emergency resources allocation under disaster risk," *Chinese Journal of Management Science*, vol. 20, no. 6, pp. 141–148, 2012.
- [13] P. Wu and H. S. Wang, "Modeling of information dissemination supernetwork in the emergencies and determination of the important nodes in the supernetwork," *Journal of the China Society for Scientific and Technical Information*, vol. 37, no. 7, pp. 722–729, 2012.
- [14] Y. J. Xi, Y. Z. Dang, and K. J. Liao, "Knowledge supernetwork model and its application in organizational knowledge systems," *Journal of Management Sciences in China*, vol. 12, no. 3, pp. 12–21, 2009.
- [15] P. Wu, H. S. Wang, Y. Li et al., "Judgment of important control nodes in the information dissemination supernetwork of emergency," *Library and Information Service*, vol. 57, no. 1, pp. 112–116, 2013.
- [16] Y. Yu, *Research on Knowledge Supernetwork in Organizational Knowledge Management*, Dalian University of Technology, Dalian, China, 2009.
- [17] J. L. Segovia-juarez, S. Colombano, and D. Kirschner, "Identifying DNA splice sites using hypernetworks with artificial molecular evolution," *Biosystems*, vol. 87, no. 2–3, pp. 117–124, 2007.
- [18] Z. T. Wang, "Reflection on supernetwork," *Journal of University of Shanghai for Science and Technology*, vol. 33, no. 3, pp. 229–237, 2011.
- [19] J.-B. Liu, J. Cao, A. Alofi, A. AL-Mazrooei, and A. Elaiw, "Applications of laplacian spectra for n-prism networks," *Neurocomputing*, vol. 198, pp. 69–73, 2016.
- [20] J. D. Luo, *Lecture Notes on Social Network Analysis*, Social Sciences Academic Press, Beijing, China, 2nd edition, 2010.
- [21] S. Hayat, S. Wang, and J.-B. Liu, "Valency-based topological descriptors of chemical networks and their applications," *Applied Mathematical Modelling*, vol. 60, pp. 164–178, 2018.

## Research Article

# Irregularity Measures for Benzene Ring Embedded in P-Type Surface

Yun Liu,<sup>1</sup> Aysha Siddiqa,<sup>2</sup> Yu-Ming Chu ,<sup>3,4</sup> Muhammad Azam,<sup>2</sup> Muhammad Asim Raza Basra,<sup>2</sup> and Abaid Ur Rehman Virk <sup>5</sup>

<sup>1</sup>School of Electronic Engineering, Huainan Normal University, Huainan 232038, China

<sup>2</sup>Institute of Chemistry, University of the Punjab, Lahore, Pakistan

<sup>3</sup>Department of Mathematics, Huzhou University, Huzhou 313000, China

<sup>4</sup>Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science & Technology, Changsha 410114, China

<sup>5</sup>Department of Mathematics, University of Management and Technology, Lahore, Pakistan

Correspondence should be addressed to Yu-Ming Chu; [chuyuming@zjhu.edu.cn](mailto:chuyuming@zjhu.edu.cn)

Received 22 June 2020; Accepted 31 July 2020; Published 26 August 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Yun Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A topological index is an important tool in predicting physicochemical properties of a chemical compound. Topological indices help us to assign a single number to a chemical compound. Drugs and other chemical compounds are frequently demonstrated as different polygonal shapes, trees, graphs, etc. In this paper, we will compute irregularity indices for the benzene ring embedded in a P-type surface ( $BR_p$ ) and the simple bounded dual of the benzene ring embedded in a P-type surface ( $SBR_p$ ).

## 1. Introduction

ÓKeeffe et al. [1] have dispersed around a quarter century ago a letter executing two 3D classification of benzene. From which, one is called 6.82P (polybenzene) and has a place with the space group  $Im\bar{3}m$ , contrast to the P-type surface, and this is due to an insertion of the hexagon fix in the surface of negative ebb and flow  $P$ . The P-type surface is facilitated to the Cartesian organizes in the Euclidean space. For further detail about this recurring surface, the author is referred to [2, 3]. This structure needed to be joined as 3D carbon solids; be that as it may, according to our knowledge, no such sequence was assumed before. The goal was to provoke the devotion of scientists to the atomic acknowledgment of such amiable thoughts in carbon nanoscience, as the graphenes took up a moment Nobel prize after C60, and also the immediate union of fullerenes is presently a reality, see for detail [4, 5].

Graph theory provides an interesting appliance in mathematical chemistry where it is used to compute the various kinds of chemical compounds and predict their various properties. One of the most important tools in the chemical graph theory is the topological index, which is useful in predicting the chemical and physical properties of the underlying chemical compound, such as boiling point, strain energy, rigidity, heat of evaporation, and tension [6, 7]. A graph having no loop or multiple edge is known as a simple graph. A molecular graph is a simple graph in which atoms and bonds are represented by vertices and edges, respectively. The degree of the vertex is the number of edges attached with that vertex. These properties of various objects are of primary interest. Wiener, in 1947, introduced the concept of the first topological index while finding the boiling point. In 1975, Gutman gave a remarkable identity [8] about Zagreb indices. Hence, these two indices are



among the oldest degree-based descriptors, and their properties are extensively investigated. The mathematical formulae of these indices are

$$M_1(G) = \sum_{uv \in E(G)} (d_u + d_v), M_2(G) = \sum_{uv \in E(G)} (d_u \times d_v). \quad (1)$$

A topological index is known as an irregularity index [9] if the value of the topological index of the graph is greater than or equal to zero, and the topological index of the graph is equal to zero if and only if the graph is regular. The irregularity indices are given in Table 1. Most of the irregularity indices are from the family of degree-based topological indices and are used in quantitative structure activity relationship modeling.

For more about topological indices, one can read [10–33].

## 2. Irregularity Indices for $BR_p$

This section is about irregularity indices of  $BR_p$ . The molecular graph of  $BR_p$  is given in Figure 1. We can observe from Figure 1 that there are two types of vertices present in the molecular graph of  $BR_p$  i.e., 2 and 3. The cardinality of the edge set is  $32pq - 2p - 2q$ .

The edge partition of  $BR_p$  is given in Table 2.

**Theorem 1.** Let  $G$  be  $BR_p$ . The irregularity indices are

- (1)  $VAR(G) = 8p^2q^2 + 2p^2q + 2pq^2 - p^2 - 2pq - q^2 / (6pq)^2$
- (2)  $AL(G) = 16pq$
- (3)  $IR1(G) = 10(8p^2q^2 + 2p^2q + 2pq^2 - p^2 - 2pq - q^2) / 3pq$
- (4)  $IR2(G) = \sqrt{(240pq - 38p - 38q / 32pq - 2p - 2q) - 32pq - 2p - 2q} / 12pq$
- (5)  $IRF(G) = 16pq$
- (6)  $IRFW(G) = 16pq / 240q - 38p - 38q$
- (7)  $IRA(G) = 8/3(5 - 2\sqrt{6})pq$
- (8)  $IRB(G) = 16(5 - 2\sqrt{6})pq$
- (9)  $IRC(G) = 48\sqrt{(6)p^2q^2 - 112p^2q^2 + 2p^2q + 2pq^2 - p^2 - 2pq - q^2} / 6pq(16pq - p - q)$
- (10)  $IRDIF(G) = 13.3328pq$
- (11)  $IRL(G) = 6.4864pq$
- (12)  $IRLU(G) = 8pq$
- (13)  $IRLF(G) = 6.5312pq$
- (14)  $IRDI(G) = 11.0896pq$
- (15)  $IRLA(G) = 6.4pq$
- (16)  $IRGA(G) = 0.3264pq$

TABLE 1: Definitions of irregularity indices.

Irregularity index	Mathematical form
VAR	$\sum_{uv \in V} (d_u - 2m/n)^2 = M_1(G)/n - (2m/n)^2$
AL	$\sum_{uv \in E(G)}  d_u - d_v $
IR1	$\sum_{uv \in V} (d_u)^3 - 2m/n \sum_{uv \in V} (d_u)^2 = F(G) - 2m/n M_1(G)$
IR2	$\sqrt{\sum_{uv \in E(G)} d_u d_v / m - 2m/n} = \sqrt{M_2(G)/m - 2m/n}$
IRF	$\sum_{uv \in E(G)} (d_u - d_v)^2 = F(G) - 2M_2(G)$
IRFW	$IRF(G)/M_2(G)$
IRA	$\sum_{uv \in E(G)} (d_u^{-1/2} - d_v^{-1/2})^2 = n - 2R(G)$
IRB	$\sum_{uv \in E(G)} (d_u^{1/2} - d_v^{1/2})^2 = M_1(G) - 2RR(G)$
IRC	$\sum_{uv \in E(G)} \sqrt{d_u d_v} / m - 2m/n = RR(G)/m - 2m/n$
IRDIF	$\sum_{uv \in E(G)}  d_u/d_v - d_v/d_u  = \sum_{i < j} m_{i,j} (j/i - i/j)$
IRL	$\sum_{uv \in E(G)}  \ln d_u - \ln d_v  = \sum_{i < j} m_{i,j} \ln(j/i)$
IRLU	$\sum_{uv \in E(G)}  d_u - d_v  / \min(d_u, d_v) = \sum_{i < j} m_{i,j} \ln(j - i/i)$
IRLF	$\sum_{uv \in E(G)}  d_u - d_v  / \sqrt{(d_u d_v)} = \sum_{i < j} m_{i,j} (j - i / \sqrt{ij})$
IRLA	$2 \sum_{uv \in E(G)}  d_u - d_v  / (d_u + d_v) = 2 \sum_{i < j} m_{i,j} (j - i / i + j)$
IRDI	$\sum_{uv \in E(G)} \ln 1 +  d_u - d_v  = \sum_{i < j} m_{i,j} \ln(i + j - 1)$
IRGA	$\sum_{uv \in E(G)} \ln(d_u + d_v / 2 \sqrt{d_u d_v}) = \sum_{i < j} m_{i,j} (i + j / 2 \sqrt{ij})$

*Proof*

$$\begin{aligned} VAR(G) &= \sum_{uv \in V} \left( d_u - \frac{2m}{n} \right)^2 = \frac{M_1(G)}{n} - \left( \frac{2m}{n} \right)^2 \\ &= \left( \frac{176pq - 20p - 20q}{24pq} \right) - \left( \frac{2(32pq - 2p - 2q)}{24pq} \right)^2 \\ &= \frac{8p^2q^2 + 2p^2q + 2pq^2 - p^2 - 2pq - q^2}{(6pq)^2}, \end{aligned} \quad (2)$$

$$\begin{aligned} AL(G) &= \sum_{uv \in E(G)} |d_u - d_v| \\ &= |2 - 2|(4p + 4q) + |2 - 3|(16pq) \\ &\quad + |3 - 3|(16pq - 6p - 6q) = 16pq, \end{aligned} \quad (3)$$

$$\begin{aligned} IR1(G) &= \sum_{uv \in V} d_u^3 - \frac{2m}{n} \sum_{uv \in V} d_u^2 = F(G) - \left( \frac{2m}{n} \right) M_1(G) \\ &= (496pq - 76p - 76q) - \frac{2(32pq - 2p - 2q)}{24pq} \\ &\quad (176pq - 20p - 20q) \\ &= \frac{10(8p^2q^2 + 2p^2q + 2pq^2 - p^2 - 2pq - q^2)}{3pq}, \end{aligned} \quad (4)$$

$$\begin{aligned} IR2(G) &= \sqrt{\frac{\sum_{uv \in E(G)} d_u d_v}{m} - \frac{2m}{n}} = \sqrt{\frac{M_2(G)}{m} - \frac{2m}{n}} \\ &= \sqrt{\frac{240pq - 38p - 38q}{32pq - 2p - 2q} - \frac{2(32pq - 2p - 2q)}{24pq}} \\ &= \sqrt{\frac{240pq - 38p - 38q}{32pq - 2p - 2q} - \frac{32pq - 2p - 2q}{12pq}}, \end{aligned} \quad (5)$$

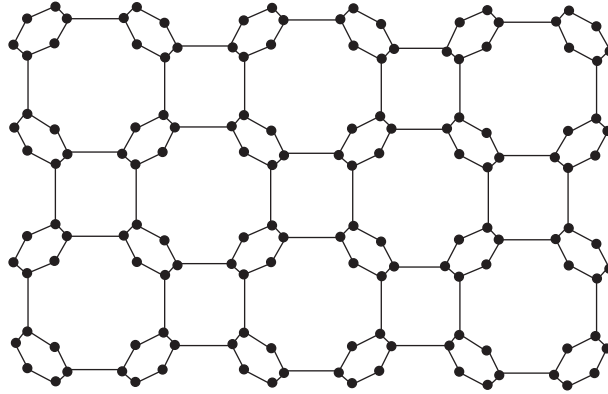


FIGURE 1:  $BR_p$ .

TABLE 2:  $E(BR_p)$ .

$(d_u, d_v)$	Frequency
(2,2)	$4p + 4q$
(2,3)	$16pq$
(3,3)	$16pq - 6p - 6q$

$$\begin{aligned}
 IRC(G) &= \frac{\sum_{uv \in E(G)} \sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{RR(G)}{m} - \frac{2m}{n} \\
 &= \frac{16(\sqrt{6} + 3)pq}{32pq - 2p - 2q} - \frac{2(32pq - 2p - 2q)}{24pq} \\
 &= \frac{48\sqrt{6}p^2q^2 - 112p^2q^2 + 2p^2q + 2pq^2 - p^2 - 2pq - q^2}{6pq(16pq - p - q)},
 \end{aligned} \tag{10}$$

$$\begin{aligned}
 IRF(G) &= \sum_{uv \in E(G)} (d_u - d_v)^2 \\
 &= (2 - 2)^2(4p + 4q) + (2 - 3)^2(16pq) \\
 &\quad + (3 - 3)^2(16pq - 6p - 6q) \\
 &= 16pq,
 \end{aligned} \tag{6}$$

$$\begin{aligned}
 IRFW(G) &= \frac{IRF(G)}{M_2(G)} \\
 &= \frac{16pq}{240pq - 38p - 38q},
 \end{aligned} \tag{7}$$

$$\begin{aligned}
 IRA(G) &= \sum_{uv \in E(G)} (d_u^{-1/2} - d_v^{-1/2})^2 \\
 &= n - 2R(G) = (24pq) - \frac{16}{3}(\sqrt{6} + 2)pq \\
 &= \frac{8}{3}(5 - 2\sqrt{6})pq,
 \end{aligned} \tag{8}$$

$$\begin{aligned}
 IRB(G) &= \sum_{uv \in E(G)} (d_u^{1/2} - d_v^{1/2})^2 = M_1(G) - 2RR(G) \\
 &= (176pq - 20p - 20q) - 32(\sqrt{6} + 3)pq \\
 &= 16(5 - 2\sqrt{6})pq,
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 IRDIF(G) &= \sum_{uv \in E(G)} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| \\
 &= \left| \frac{2}{2} - \frac{2}{2} \right| (4p + 4q) + \left| \frac{2}{3} - \frac{3}{2} \right| (24pq) \\
 &\quad + \left| \frac{3}{3} - \frac{3}{3} \right| (16pq - 6p - 6q) \\
 &= 13.3328pq,
 \end{aligned} \tag{11}$$

$$\begin{aligned}
 IRL(G) &= \sum_{uv \in E(G)} |\ln d_u - \ln d_v| \\
 &= |\ln 2 - \ln 2|(4p + 4q) + |\ln 2 - \ln 3|(24pq) \\
 &\quad + |\ln 3 - \ln 3|(16pq - 6p - 6q) \\
 &= 6.4864pq,
 \end{aligned} \tag{12}$$

$$\begin{aligned}
 IRLU(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\min(d_u, d_v)} \\
 &= \frac{|2 - 2|}{2} (4p + 4q) + \frac{|2 - 3|}{2} (24pq) \\
 &\quad + \frac{|3 - 3|}{3} (16pq - 6p - 6q) \\
 &= 8pq,
 \end{aligned} \tag{13}$$

$$\begin{aligned}
 \text{IRLF}(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\sqrt{d_u \cdot d_v}} \\
 &= \frac{|2-2|}{\sqrt{4}}(4p+4q) + \frac{|2-3|}{\sqrt{6}}(24pq) \\
 &\quad + \frac{|3-3|}{\sqrt{9}}(16pq-6p-6q) \\
 &= 6.5312pq,
 \end{aligned} \tag{14}$$

$$\begin{aligned}
 \text{IRLA}(G) &= \sum_{uv \in E(G)} 2 \frac{|d_u - d_v|}{(d_u + d_v)} \\
 &= 2 \frac{|2-2|}{4}(4p+4q) + 2 \frac{|2-3|}{5}(24pq) \\
 &\quad + 2 \frac{|3-3|}{6}(16pq-6p-6q) \\
 &= 6.4pq,
 \end{aligned} \tag{15}$$

$$\begin{aligned}
 \text{IRD1}(G) &= \sum_{uv \in E(G)} \ln\{1 + |d_u - d_v|\} \\
 &= \ln\{1 + |2-2|\}(4p+4q) + \ln\{1 + |2-3|\}(24pq) \\
 &\quad + \ln\{1 + |3-3|\}(16pq-6p-6q) \\
 &= 11.0896pq,
 \end{aligned} \tag{16}$$

$$\begin{aligned}
 \text{IRGA}(G) &= \sum_{uv \in E(G)} \ln\left(\frac{d_u + d_v}{2\sqrt{d_u d_v}}\right) \\
 &= \ln\left(\frac{2+2}{2\sqrt{2 \times 2}}\right)(4p+4q) + \ln\left(\frac{2+3}{2\sqrt{2 \times 3}}\right)(24pq) \\
 &\quad + \ln\left(\frac{3+3}{2\sqrt{3 \times 3}}\right)(16pq-6p-6q) \\
 &= 0.3264pq.
 \end{aligned} \tag{17}$$

□

### 3. Irregularity Indices for $SBR_p$

Here, we will discuss the irregularity indices for  $SBR_p$ . The molecular graph of  $SBR_p$  is given in Figures 2 and 3. The cardinalities of different types of edges are given in Table 3. The cardinality of the edge set of  $SBR_p$  is  $32pq - 2p - 2q$  (Figure 4).

The edge partition of  $SBR_p$  is given in Table 3.

**Theorem 2.** Let  $G$  be  $SBR_p$ . The irregularity indices are

- (1)  $\text{VAR}(G) = 4(576p^2q^2 - 88p^2q - 88pq^2 - 78p^2 - 572pq - 78q^2 + 209p + 209q - 86)/(8pq - 2p - 2q + 1)^2$
- (2)  $\text{AL}(G) = 64pq - 28p - 28q + 8$
- (3)  $\text{IR1}(G) = 4(7792p^2q^2 - 2660p^2q - 2660pq^2 - 690p^2 - 540pq - 690q^2 + 97p + 97q + 227)/8pq - 2p - 2q + 1$

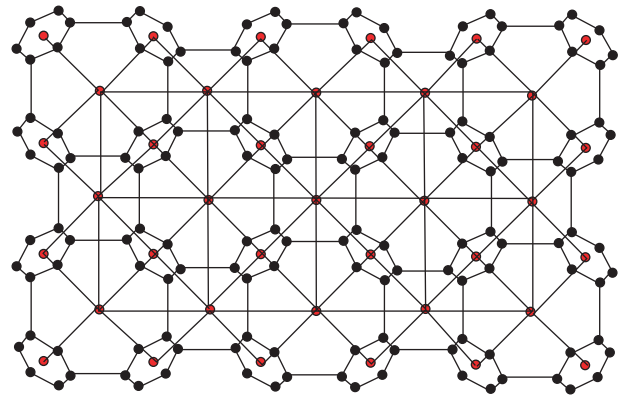


FIGURE 2: Construction of  $SBR_p$ .

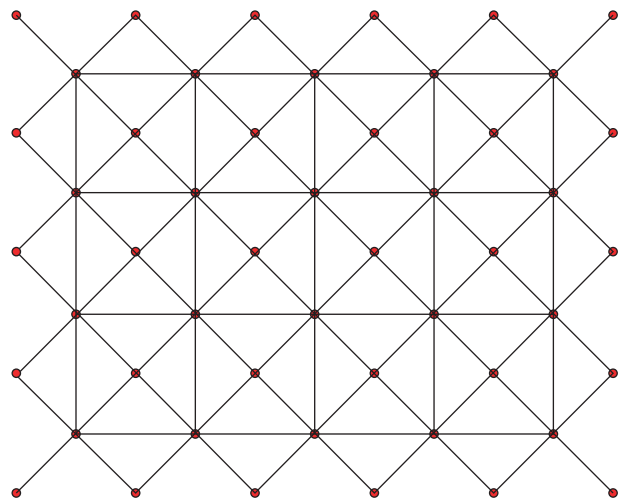


FIGURE 3:  $SBR_p$ .

TABLE 3:  $E(SBR_p)$ .

$(d_u, d_v)$	Frequency
(1,6)	4
(2,6)	8
(2,7)	$8p + 8q - 24$
(4,6)	4
(4,7)	$8p + 8q - 24$
(4,8)	$16pq - 24p - 24q + 36$
(6,7)	8
(7,7)	$4p + 4q - 16$
(7,8)	$4p + 4q - 12$
(8,8)	$24pq - 14p - 14q + 8$

- (4)  $\text{IR2}(G) = \sqrt{(2048pq - 908p - 908q - 248)/24pq - 14p - 14q + 8} - (2(24pq - 14p - 14q + 8)/8pq - 2p - 2q + 1)$
- (5)  $\text{IRF}(G) = 256pq - 108p - 108q - 4$
- (6)  $\text{IRFW}(G) = 256pq - 108p - 108q - 4/2048pq - 908p - 908q - 248$

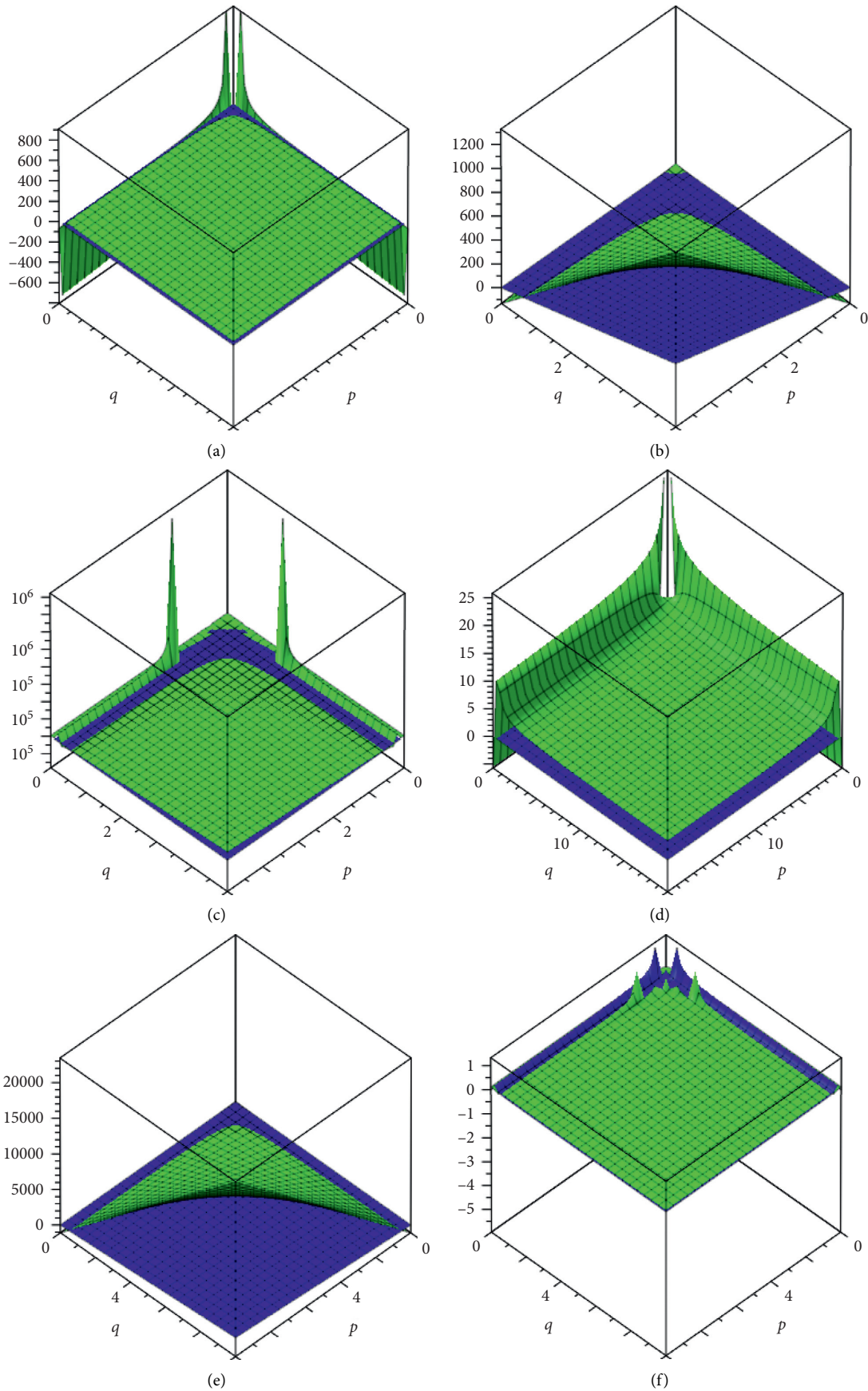


FIGURE 4: Continued.

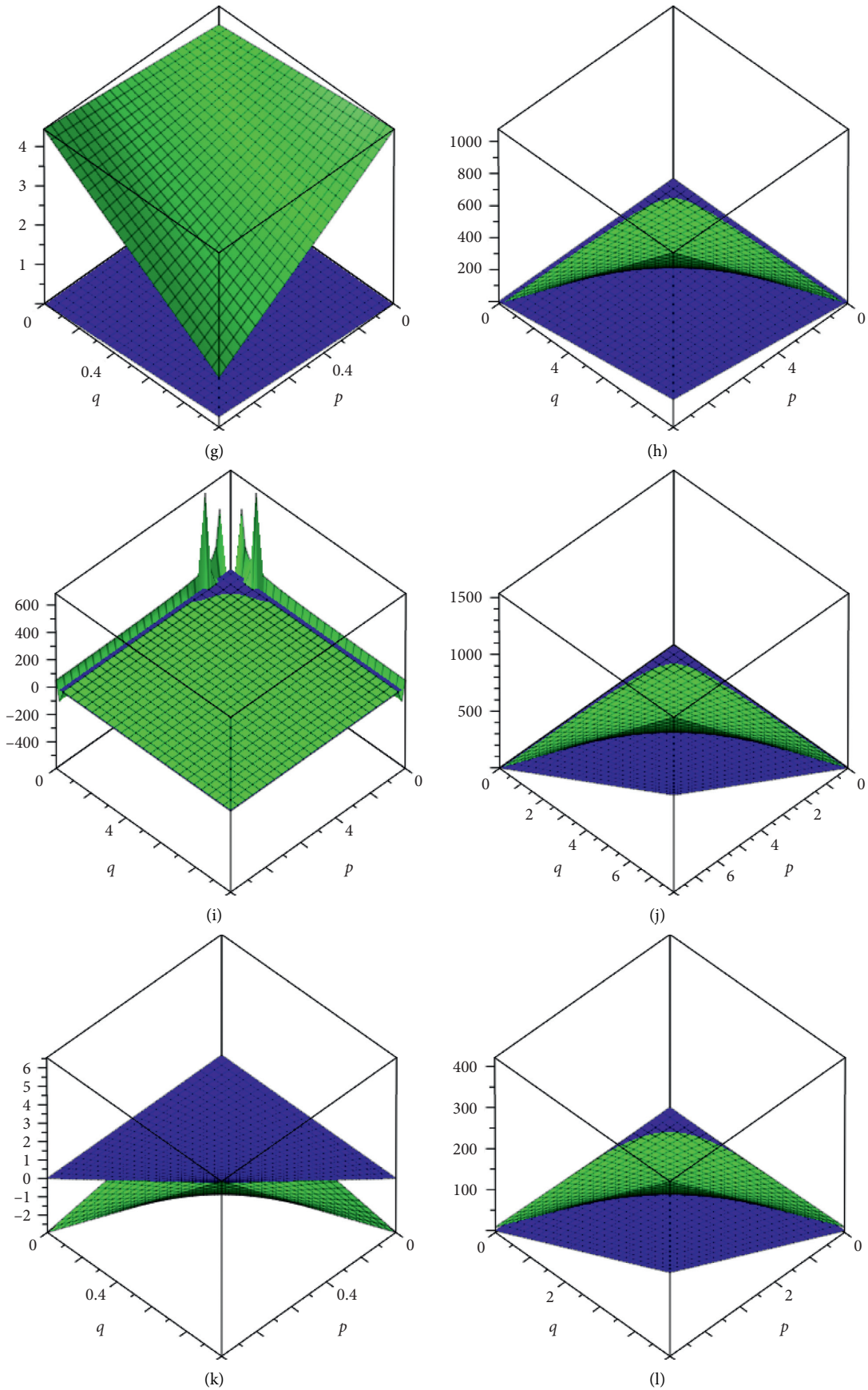


FIGURE 4: Continued.



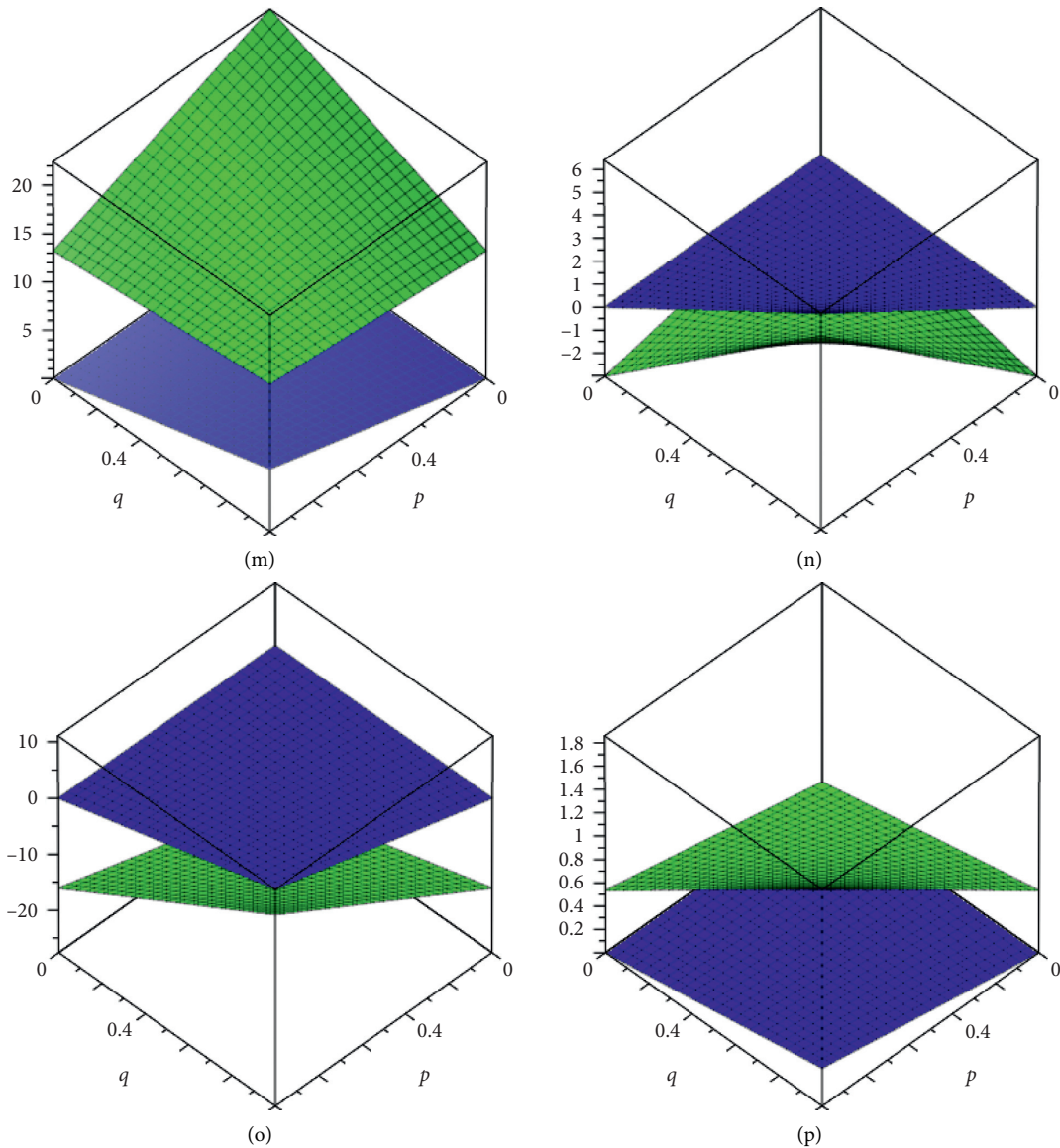


FIGURE 4: Comparison of results. (a) VAR. (b) AL. (c) IR1. (d) IR2. (e) IRF. (f) IRFW. (g) IRA. (h) IRB. (i) IRC. (j) IRDIF. (k) IRL. (l) IRLU. (m) IRLF. (n) IRLA. (o) IRD1. (p) IRGA.

$$(6) \text{ IRF}(G) = 256pq - 108p - 108q - 4$$

$$(7) \text{ IRA}(G) = (2 - 4\sqrt{2})pq + 1/14(5 + 84\sqrt{2} - 16\sqrt{7} - 20\sqrt{14})p + 1/14(5 + 84\sqrt{2} - 16\sqrt{7} - 20\sqrt{14})q + 1/21(75 - 189\sqrt{2} + 56\sqrt{3} - 42\sqrt{6} + 72\sqrt{7} + 30\sqrt{14} - 8\sqrt{42})$$

$$(8) \text{ IRB}(G) = (192 - 128)pq + (192\sqrt{2} - 32\sqrt{7} - 32\sqrt{14} - 68)p + (192\sqrt{2} - 32\sqrt{7} - 32\sqrt{14} - 68)q - (288\sqrt{2} + 32\sqrt{3} + 24\sqrt{6} - 96\sqrt{7} - 96\sqrt{14} + \sqrt{42})$$

$$(9) \text{ IRC}(G) = (1/(24pq - 2p - 2q + 1))(8pq - 2p - 2q + 1)(4(-16 + 16\sqrt{42}pq - 8\sqrt{3}p - 8\sqrt{7}q^2 + 48\sqrt{2}q^2 - 8\sqrt{3}q + 28\sqrt{7}p + 28\sqrt{7}q - 96\sqrt{2}p - 96\sqrt{2}q + 3\sqrt{6} - 12\sqrt{14} + 48\sqrt{2}p^2 - 8\sqrt{7}p^2 + 4\sqrt{3} + 36\sqrt{2} - 12\sqrt{7} - 8\sqrt{14}p^2 + 32\sqrt{7}p^2q + 128\sqrt{2}p^2q^2 - 8\sqrt{14}q^2 + 32\sqrt{7}pq^2 - 6\sqrt{6}q + 28\sqrt{14}q + 2\sqrt{42} - 6\sqrt{6}p - 224\sqrt{2}p^2q + 400\sqrt{2}pq -$$

$$224\sqrt{2}pq^2 + 32\sqrt{3}pq - 112\sqrt{7}pq + 28\sqrt{14}p + 18p + 96p^2q^2 - 72p^2q - 72pq^2 + 32\sqrt{14}p^2q + 32\sqrt{14}pq^2 - 4\sqrt{42}p - 4\sqrt{42}q + 24\sqrt{6}pq - 112\sqrt{14}pq + 28p^2 + 28q^2 + 18q - 100pq)$$

$$(10) \text{ IRDIF}(G) = 24pq + 0.2128p + 0.2128q - 4.1624$$

$$(11) \text{ IRL}(G) = 11.0896pq - 1.6020p - 1.6020q - 1.3356$$

$$(12) \text{ IRLU}(G) = 16pq + 2.56p + 2.56q - 4.40$$

$$(13) \text{ IRLF}(G) = 11.3136pq - 9.2104p - 9.2104q + 22.4452$$

$$(14) \text{ IRLA}(G) = 10.6656pq - 2.21308p - 2.21308q - 8.1396$$

$$(15) \text{ IRD1}(G) = 11.1008pq + 11.5444p + 11.5444q - 27.5856$$

$$(16) \text{ IRGA}(G) = 0.9408pq + 0.3824p + 0.3824q + 0.1532$$



Proof

$$\begin{aligned}
 \text{VAR}(G) &= \sum_{u \in V} \left( d_u - \frac{2m}{n} \right)^2 \\
 &= \frac{M_1(G)}{n} - \left( \frac{2m}{n} \right)^2 = \left( \frac{576pq - 236p - 236q - 88}{8pq - 2p - 2q + 1} \right) \\
 &\quad - \left( \frac{2(24pq - 14p - 14q + 8)}{8pq - 2p - 2q + 1} \right)^2 \\
 &= \frac{4(576p^2q^2 - 88p^2q - 88pq^2 - 78p^2 - 572pq - 78q^2 + 209p + 209q - 86)}{(8pq - 2p - 2q + 1)^2},
 \end{aligned} \tag{18}$$

$$\begin{aligned}
 \text{AL}(G) &= \sum_{uv \in E(G)} |d_u - d_v| \\
 &= |1 - 6|(4) + |2 - 6|(8) + |2 - 7|(8p + 8q - 24) + |4 - 6|(4) + |4 - 7|(8p + 8q - 24) + |4 - 8|(16pq - 24p - 24q + 36) \\
 &\quad + |6 - 7|(8) + |7 - 7|(4p + 4q - 16) \\
 &\quad + |7 - 8|(4p + 4q - 12) \\
 &\quad + |8 - 8|(24pq - 14p - 14q + 8) \\
 &= 64pq - 28p - 28q + 8,
 \end{aligned} \tag{19}$$

$$\begin{aligned}
 \text{IR1}(G) &= \sum_{u \in V} d_u^3 - \frac{2m}{n} \sum_{u \in V} d_u^2 = F(G) - \left( \frac{2m}{n} \right) M_1(G) \\
 &= (4352pq - 1924p - 1924q - 500) \\
 &\quad - \left( \frac{2(24pq - 14p - 14q + 8)}{8pq - 2p - 2q + 1} \right) (576pq - 236p - 236q - 88) \\
 &= \frac{4(7792p^2q^2 - 2660p^2q - 2660pq^2 - 690p^2 - 540pq - 690q^2 + 97p + 97q + 227)}{8pq - 2p - 2q + 1},
 \end{aligned} \tag{20}$$

$$\begin{aligned}
 \text{IR2}(G) &= \sqrt{\frac{\sum_{uv \in E(G)} d_u d_v}{m}} - \frac{2m}{n} = \sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n} \\
 &= \sqrt{\frac{2048pq - 908p - 908q - 248}{8pq - 2p - 2q + 1}} - \left( \frac{2(24pq - 14p - 14q + 8)}{8pq - 2p - 2q + 1} \right) \\
 &= \sqrt{\frac{2048pq - 908p - 908q - 248}{24pq - 14p - 14q + 8}} - \frac{2(24pq - 14p - 14q + 8)}{8pq - 2p - 2q + 1},
 \end{aligned} \tag{21}$$

$$\begin{aligned}
 \text{IRF}(G) &= \sum_{uv \in E(G)} (d_u - d_v)^2 \\
 &= (1 - 6)^2(4) + (2 - 6)^2(8) + (2 - 7)^2(8p + 8q - 24) + (4 - 6)^2(4) \\
 &\quad + (4 - 7)^2(8p + 8q - 24) + (4 - 8)^2(16pq - 24p - 24q + 36) \\
 &\quad + (6 - 7)^2(8) + (7 - 7)^2(4p + 4q - 16) + (7 - 8)^2(4p + 4q - 12) \\
 &\quad + (8 - 8)^2(24pq - 14p - 14q + 8) \\
 &= 256pq - 108p - 108q - 4,
 \end{aligned} \tag{22}$$

$$\begin{aligned} \text{IRFW}(G) &= \frac{\text{IRF}(G)}{M_2(G)} \\ &= \frac{256pq - 108p - 108q - 4}{2048pq - 908p - 908q - 248} \end{aligned} \quad (23)$$

$$\begin{aligned} \text{IRA}(G) &= \sum_{uv \in E(G)} (d_u^{-1/2} - d_v^{-1/2})^2 = n - 2R(G) \\ &= (8pq - 2p - 2q + 1) - 2 \left( \begin{array}{c} \sqrt{6} + \frac{4}{3}\sqrt{3} + \frac{1}{14}\sqrt{14}(8p + 8q - 24) \\ + \frac{1}{14}\sqrt{7}(8p + 8q - 24) + \frac{1}{8}\sqrt{2}(16pq - 24p - 24q + 36) \\ + \frac{4}{21}\sqrt{42} - \frac{33}{28}p - \frac{33}{28}q - \frac{9}{7} + \frac{1}{28}\sqrt{14}(4p + 4q - 12) + 3pq \end{array} \right) \\ &= (2 - 4\sqrt{2})pq + \frac{1}{14}(5 + 84\sqrt{2} - 16\sqrt{7} - 20\sqrt{14})p \\ &\quad + \frac{1}{14}(5 + 84\sqrt{2} - 16\sqrt{7} - 20\sqrt{14})q \\ &\quad + \frac{1}{21}(75 - 189\sqrt{2} + 56\sqrt{3} - 42\sqrt{6} + 72\sqrt{7} + 30\sqrt{14} - 8\sqrt{42}), \end{aligned} \quad (24)$$

$$\begin{aligned} \text{IRB}(G) &= \sum_{uv \in E(G)} (d_u^{1/2} - d_v^{1/2})^2 = M_1(G) - 2RR(G) \\ &= (2048pq - 908p - 908q - 248) - 2 \left( \begin{array}{c} 12\sqrt{6} + 16\sqrt{3} + \sqrt{14}(8p + 8q - 24) \\ + 2\sqrt{7}(8p + 8q - 24) + 4\sqrt{2}(16pq - 24p - 24q + 36) \\ + 8\sqrt{42} - 84p - 84q - 48 + 2\sqrt{14}(4p + 4q - 12) + 192pq \end{array} \right) \\ &= (192 - 128)pq + (192\sqrt{2} - 32\sqrt{7} - 32\sqrt{14} - 68)p \\ &\quad + (192\sqrt{2} - 32\sqrt{7} - 32\sqrt{14} - 68)q \\ &\quad - \left( \begin{array}{c} 288\sqrt{2} + 32\sqrt{3} \\ + 24\sqrt{6} - 96\sqrt{7} - 96\sqrt{14} + \sqrt{42} \end{array} \right), \end{aligned} \quad (25)$$

$$\begin{aligned} \text{IRC}(G) &= \frac{\sum_{uv \in E(G)} \sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{RR(G)}{m} - \frac{2m}{n} \\ &= \frac{1}{24pq - 14p - 14q + 8} (12\sqrt{6} + 16\sqrt{3} + \sqrt{14}(8p + 8q - 24) + 2\sqrt{7}(8p + 8q - 24) + 4\sqrt{2}(16pq - 24p - 24q + 36)) \\ &\quad + 8\sqrt{42} - 84p - 84q - 48 + 2\sqrt{14}(4p + 4q - 12) + 192pq - \left( \frac{2(24pq - 14p - 14q + 8)}{8pq - 2p - 2q + 1} \right) \\ &= \frac{1}{(24pq - 2p - 2q + 1)(8pq - 2p - 2q + 1)} \left( 4(-16 + 16\sqrt{42}pq - 8\sqrt{3}p - 8\sqrt{7}q^2 + 48\sqrt{2}q^2 - 8\sqrt{3}q + 28\sqrt{7}p \right. \\ &\quad \left. + 28\sqrt{7}q - 96\sqrt{2}p) \right. \\ &\quad \left. - 96\sqrt{2}q + 3\sqrt{6} - 12\sqrt{14} + 48\sqrt{2}p^2 - 8\sqrt{7}p^2 + 4\sqrt{3} + 36\sqrt{2} - 12\sqrt{7} - 8\sqrt{14}p^2 \right. \\ &\quad \left. + 32\sqrt{7}p^2q + 128\sqrt{2}p^2q^2 - 8\sqrt{14}q^2 + 32\sqrt{7}pq^2 - 6\sqrt{6}q + 28\sqrt{14}q + 2\sqrt{42} - 6\sqrt{6}p - 224\sqrt{2}p^2q \right. \\ &\quad \left. + 400\sqrt{2}pq - 224\sqrt{2}p^2q^2 + 32\sqrt{3}pq - 112\sqrt{7}pq + 28\sqrt{14}p + 18p + 96p^2q^2 - 72p^2q - 72pq^2 \right. \\ &\quad \left. + 32\sqrt{14}p^2q + 32\sqrt{14}pq^2 - 4\sqrt{42}p - 4\sqrt{42}q + 24\sqrt{6}pq - 112\sqrt{14}pq + 28p^2 + 28q^2 + 18q - 100pq, \right. \end{aligned} \quad (26)$$

$$\begin{aligned}
\text{IRDIF}(G) &= \sum_{uv \in E(G)} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| \\
&= \left| \frac{1}{6} - \frac{6}{1} \right| (4) + \left| \frac{2}{6} - \frac{6}{2} \right| (8) + \left| \frac{2}{7} - \frac{7}{2} \right| (8p + 8q - 24) + \left| \frac{4}{6} - \frac{6}{4} \right| (4) \\
&\quad + \left| \frac{4}{7} - \frac{7}{4} \right| (8p + 8q - 24) + \left| \frac{4}{8} - \frac{8}{4} \right| (16pq - 24p - 24q + 36) \\
&\quad + \left| \frac{6}{7} - \frac{7}{6} \right| (8) + \left| \frac{7}{7} - \frac{7}{7} \right| (4p + 4q - 16) + \left| \frac{7}{8} - \frac{8}{7} \right| (4p + 4q - 12) \\
&\quad + \left| \frac{8}{8} - \frac{8}{8} \right| (24pq - 14p - 14q + 8) \\
&= 24pq + 0.2128p + 0.2128q - 4.1624,
\end{aligned} \tag{27}$$

$$\begin{aligned}
\text{IRL}(G) &= \sum_{uv \in E(G)} |\ln d_u - \ln d_v| \\
&= |\ln 2 - \ln 2| (4) + |\ln 2 - \ln 6| (8) + |\ln 2 - \ln 7| (8p + 8q - 24) + |\ln 4 - \ln 6| (4) \\
&\quad + |\ln 4 - \ln 7| (8p + 8q - 24) + |\ln 4 - \ln 8| (16pq - 24p - 24q + 36) \\
&\quad + |\ln 6 - \ln 7| (8) + |\ln 7 - \ln 7| (4p + 4q - 16) + |\ln 7 - \ln 8| (4p + 4q - 12) \\
&\quad + |\ln 8 - \ln 8| (24pq - 14p - 14q + 8) = 11.0896pq - 1.6020p - 1.6020q - 1.3356,
\end{aligned} \tag{28}$$

$$\begin{aligned}
\text{IRLU}(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\min(d_u, d_v)} \\
&= \frac{|1-2|}{1} (4) + \frac{|2-6|}{2} (8) + \frac{|2-7|}{2} (8p + 8q - 24) + \frac{|4-6|}{4} (4) \\
&\quad + \frac{|4-7|}{4} (8p + 8q - 24) + \frac{|4-8|}{4} (16pq - 24p - 24q + 36) \\
&\quad + \frac{|6-7|}{6} (8) + \frac{|7-7|}{7} (4p + 4q - 16) + \frac{|7-8|}{7} (4p + 4q - 12) \\
&\quad + \frac{|8-8|}{8} (24pq - 14p - 14q + 8) \\
&= 16pq + 2.56p + 2.56q - 4.40,
\end{aligned} \tag{29}$$

$$\begin{aligned}
\text{IRLF}(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\sqrt{d_u \cdot d_v}} \\
&= \frac{|1-6|}{\sqrt{6}} (4) + \frac{|2-6|}{\sqrt{12}} (8) + \frac{|2-7|}{\sqrt{14}} (8p + 8q - 24) + \frac{|4-6|}{\sqrt{24}} (4) \\
&\quad + \frac{|4-7|}{\sqrt{28}} (8p + 8q - 24) + \frac{|4-8|}{\sqrt{32}} (16pq - 24p - 24q + 36) \\
&\quad + \frac{|6-7|}{\sqrt{42}} (8) + \frac{|7-7|}{\sqrt{49}} (4p + 4q - 16) + \frac{|7-8|}{\sqrt{56}} (4p + 4q - 12) \\
&\quad + \frac{|8-8|}{\sqrt{64}} (24pq - 14p - 14q + 8) \\
&= 11.3136pq - 9.2104p - 9.2104q + 22.4452,
\end{aligned} \tag{30}$$

$$\begin{aligned}
 \text{IRLA}(G) &= \sum_{uv \in E(G)} 2 \frac{|d_u - d_v|}{(d_u + d_v)} \\
 &= 2 \frac{|1 - 6|}{7} (4) + 2 \frac{|2 - 6|}{8} (8) + 2 \frac{|2 - 7|}{9} (8p + 8q - 24) + 2 \frac{|4 - 6|}{10} (4) \\
 &\quad + 2 \frac{|4 - 7|}{11} (8p + 8q - 24) + 2 \frac{|4 - 8|}{12} (16pq - 24p - 24q + 36) \\
 &\quad + 2 \frac{|6 - 7|}{13} (8) + 2 \frac{|7 - 7|}{14} (4p + 4q - 16) \\
 &\quad + 2 \frac{|7 - 8|}{15} (4p + 4q - 12) \\
 &\quad + 2 \frac{|8 - 8|}{16} (24pq - 14p - 14q + 8) \\
 &= 10.6656pq - 2.21308p - 2.21308q - .81396,
 \end{aligned} \tag{31}$$

$$\begin{aligned}
 \text{IRD1}(G) &= \sum_{uv \in E(G)} \ln\{1 + |d_u - d_v|\} \\
 &= \ln\{1 + |1 - 6|\} (4) + \ln\{1 + |2 - 6|\} (8) + \ln\{1 + |2 - 7|\} (8p + 8q - 24) \\
 &\quad + \ln\{1 + |4 - 6|\} (4) + \ln\{1 + |4 - 7|\} (8p + 8q - 24) \\
 &\quad + \ln\{1 + |4 - 8|\} (16pq - 24p - 24q + 36) \\
 &\quad + \ln\{1 + |6 - 7|\} (8) + \ln\{1 + |7 - 7|\} (4p + 4q - 16) \\
 &\quad + \ln\{1 + |7 - 8|\} (4p + 4q - 12) \\
 &\quad + \ln\{1 + |8 - 8|\} (24pq - 14p - 14q + 8) \\
 &= 11.1008pq + 11.5444p + 11.5444q - 27.5856,
 \end{aligned} \tag{32}$$

$$\begin{aligned}
 \text{IRGA}(G) &= \sum_{uv \in E(G)} \ln\left(\frac{d_u + d_v}{2\sqrt{d_u d_v}}\right) \\
 &= \ln\left(\frac{1 + 6}{2\sqrt{1 \times 6}}\right) (4) + \ln\left(\frac{2 + 6}{2\sqrt{2 \times 6}}\right) (8) \\
 &\quad + \ln\left(\frac{2 + 7}{2\sqrt{2 \times 7}}\right) (8p + 8q - 24) + \ln\left(\frac{4 + 6}{2\sqrt{4 \times 6}}\right) (4) + \ln\left(\frac{4 + 7}{2\sqrt{4 \times 7}}\right) (8p + 8q - 24) \\
 &\quad + \ln\left(\frac{4 + 8}{2\sqrt{4 \times 8}}\right) (16pq - 24p - 24q + 36) + \ln\left(\frac{6 + 7}{2\sqrt{6 \times 7}}\right) (8) \\
 &\quad + \ln\left(\frac{7 + 7}{2\sqrt{7 \times 7}}\right) (4p + 4q - 16) + \ln\left(\frac{7 + 8}{2\sqrt{7 \times 8}}\right) (4p + 4q - 12) \\
 &\quad + \ln\left(\frac{8 + 8}{2\sqrt{8 \times 8}}\right) (24pq - 14p - 14q + 8) \\
 &= 0.9408pq + 0.3824p + 0.3824q + 0.1532.
 \end{aligned} \tag{33}$$

#### 4. Graphical Representation

In this section, we will give the comparison of sixteen irregularity indices of  $BR_p$  and  $BR_p$ . The colour blue is fixed for  $BR_p$ , and the green colour is fixed for  $SBR_p$ . From the plots, one can observe the behaviour of computed results with respect to involved parameters.

#### 5. Conclusions

The irregularity of a graph can be defined by different so-called graph topological indices [34, 39]. It is known that the irregularity measures are not always compatible. We foresee that our results could play an important role in determining properties of the understudy material such as enthalpy, □

toxicity, resistance, and entropy. It is hoped that this article will aid the reader in understanding the rationale and utility of a simple quantitative tool which could be used in malocclusion assessment.

## Data Availability

All data are including in this paper.

## Conflicts of Interest

The authors do not have conflicts of interest.

## Authors' Contributions

All authors contributed equally in this paper.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant nos. 11971142, 11871202, 61673169, 11701176, 11626101, and 11601485) and also by the research foundation of Huainan Normal University (Grant no. 2019XJZD05).

## References

- [1] M. ÓKeefe, G. B. Adams, and O. F. Sankey, "Predicted new low energy forms of carbon," *Physical Review Letters*, vol. 68, no. 15, pp. 2325–2328, 1992.
- [2] M. V. Diudea, *Nanostructures: Novel Architecture*, Nova Publishers, Hauppauge, NY, USA, 2005.
- [3] M. V. Diudea and C. L. Nagy, *Periodic nanostructures*, vol. 7, Springer Science & Business Media, Berlin, Germany, 2007.
- [4] K. Y. Amsharov and M. Jansen, "A C78Fullerene precursor: toward the direct synthesis of higher fullerenes," *The Journal of Organic Chemistry*, vol. 73, no. 7, pp. 2931–2934, 2008.
- [5] Y. Ma, S. Cao, Y. Shi, I. Gutman et al., "From the connectivity index to various Randic-type descriptors," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 80, no. 1, pp. 85–106, 2018.
- [6] F. M. Brückler, T. Dolic, A. Graovac, and I. Gutman, "On a class of distance-based molecular structure descriptors," *Chemical Physics Letters*, vol. 503, no. 4–6, pp. 336–338, 2011.
- [7] H. Gonzalez-Diaz, S. Vilar, L. Santana, and E. Uriarte, "Medicinal chemistry and bioinformatics - current trends in drugs discovery with networks topological indices," *Current Topics in Medicinal Chemistry*, vol. 7, no. 10, pp. 1015–1029, 2007.
- [8] H. Hosoya, K. Hosoi, and I. Gutman, "A topological index for the totalp-electron energy," *Theoretica Chimica Acta*, vol. 38, no. 1, pp. 37–47, 1975.
- [9] T. Réti, R. Sharafadini, A. Dregelyi-Kiss, and H. Haghbin, "Graph irregularity indices used as molecular descriptors in QSPR studies," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 79, pp. 509–524, 2018.
- [10] E. Deutsch and S. Klavzar, "S. M-Polynomial and degree-based topological indices," *Iranian Journal of Mathematical Chemistry*, vol. 6, pp. 93–102, 2015.
- [11] M. Ajmal, W. Nazeer, M. Munir, S. M. Kang, and C. Y. Jung, "The M-polynomials and topological indices of generalized prism network," *International Journal of Mathematical Analysis*, vol. 11, no. 6, pp. 293–303, 2017.
- [12] M. Munir, W. Nazeer, Z. Shahzadi, and S. Kang, "Some invariants of circulant graphs," *Symmetry*, vol. 8, no. 11, p. 134, 2016.
- [13] A. A. Dobrynin, R. Entringer, and I. Gutman, "Wiener index of trees: theory and applications," *Acta Applicandae Mathematicae*, vol. 66, no. 3, pp. 211–249, 2001.
- [14] I. Gutman and O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer Science & Business Media, Berlin, Germany, 2012.
- [15] Y. Hu, X. Li, Y. Shi, T. Xu, and I. Gutman, "On molecular graphs with smallest and greatest zeroth-order general Randić index," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 54, no. 2, pp. 425–434, 2005.
- [16] G. Caporossi, I. Gutman, P. Hansen, and L. Pavlović, "Graphs with maximum connectivity index," *Computational Biology and Chemistry*, vol. 27, no. 1, pp. 85–90, 2003.
- [17] X. Li and I. Gutman, *Mathematical Chemistry Monographs No. 1*, University of Kragujevac, Kragujevac, Macedonia, 2006.
- [18] X. Li, I. Gutman, and M. Randić, *Mathematical Aspects of Randić-type Molecular Structure Descriptors*, University, Faculty of Science, Kragujevac, Macedonia, 2006.
- [19] I. Gutman, B. Furtula, and C. Elphick, "Three new/old vertex-degree-based topological indices," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 72, no. 3, pp. 617–632, 2014.
- [20] X. Li and Y. Shi, "A survey on the Randić index," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 59, no. 1, pp. 127–156, 2008.
- [21] S. Nikolić, G. Kovacević, A. Milicević, and N. Trinajstić, "The Zagreb indices 30 years after," *Croatica Chimica Acta*, vol. 76, no. 2, pp. 113–124, 2003.
- [22] I. Gutman and K. C. Das, "The first Zagreb index 30 years after," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 50, no. 1, pp. 83–92, 2004.
- [23] J.-B. Liu, J. Zhao, J. Min, and J. Cao, "The hosoya index of graphs formed by a fractal graph," *Fractals*, vol. 27, no. 8, Article ID 1950135, 2019.
- [24] D. Vukicević and A. Graovac, "Valence connectivity versus Randić, Zagreb and modified Zagreb index: A linear algorithm to check discriminative properties of indices in acyclic molecular graphs," *Croatica Chimica Acta*, vol. 77, no. 3, pp. 501–508, 2004.
- [25] A. Milicević, S. Nikolić, and N. Trinajstić, "On reformulated Zagreb indices," *Molecular Diversity*, vol. 8, no. 4, pp. 393–399, 2004.
- [26] C. K. Gupta, V. Lokesh, S. B. Shwetha, and P. S. Ranjini, "On the symmetric division deg index of graph," *Southeast Asian Bulletin of Mathematics*, vol. 40, no. 1, 2016.
- [27] O. Favaron, M. Mahéo, and J. F. Saclé, "Some eigenvalue properties in graphs (conjectures of Graffiti II)," *Discrete Mathematics*, vol. 111, no. 1–3, pp. 197–220, 1993.
- [28] A. T. Balaban, "Highly discriminating distance-based topological index," *Chemical Physics Letters*, vol. 89, no. 5, pp. 399–404, 1982.
- [29] S. M. Hosamani, V. Lokesh, I. N. Cangul, and K. M. Devendraiah, "On certain topological indices of the derived graphs of subdivision graphs," *TWMS Journal of Applied and Engineering Mathematics*, vol. 6, no. 2, p. 324, 2016.
- [30] J.-B. Liu, Z.-Y. Shi, Y.-H. Pan, J. Cao, M. Abdel-Aty, and A.-J. Udai, "Computing the Laplacian spectrum of linear octagonal-quadrilateral networks and its applications," *Polycyclic Aromatic Compounds*, pp. 1–12, 2020.
- [31] A. Aslam, M. K. Jamil, W. Gao, and W. Nazeer, "Topological aspects of some dendrimer structures," *Nanotechnology Reviews*, vol. 7, no. 2, pp. 123–129, 2018.

- [32] M. Randic, "Characterization of molecular branching," *Journal of the American Chemical Society*, vol. 97, no. 23, pp. 6609–6615, 1975.
- [33] D. Wang, H. Ahmad, and W. Nazeer, "Hosoya and Harary polynomials of TUC 4 nanotube," *Mathematical Methods in the Applied Sciences*, 2020.
- [34] A. R. Virk, M. A. Rehman, and W. Nazeer, "New definition of atomic bond connectivity index to overcome deficiency of structure sensitivity and abruptness in existing definition," *Scientific Inquiry and Review*, vol. 3, no. 4, pp. 1–20, 2019.
- [35] W. Gao, M. Younas, A. Farooq, A. Mahboob, and W. Nazeer, "M-polynomials and degree-based topological indices of the crystallographic structure of molecules," *Biomolecules*, vol. 8, no. 4, p. 107, 2018.
- [36] M. Munir, W. Nazeer, S. Rafique, and S. Kang, "M-polynomial and related topological indices of nanostar dendrimers," *Symmetry*, vol. 8, no. 9, p. 97, 2016.
- [37] M. Munir, W. Nazeer, S. Rafique, and S. Kang, "M-polynomial and degree-based topological indices of polyhex nanotubes," *Symmetry*, vol. 8, no. 12, p. 149, 2016.
- [38] Z. Shao, A. R. Virk, M. S. Javed, M. A. Rehman, and M. R. Farahani, "Degree based graph invariants for the molecular graph of Bismuth Tri-iodide," *Engineering and Applied Science Letters*, vol. 2, no. 1, pp. 1–11, 2019.
- [39] A. u. R. Virk, M. N. Jhangeer, M. N. Jhangeer, and M. A. Rehman, "Reverse Zagreb and reverse hyper-zagreb indices for silicon carbide  $Si_2C_3 - I[r, s]$  and  $Si_2C_3 - II[r, s]$ ," *Engineering and Applied Science Letters*, vol. 1, no. 2, pp. 37–50, 2018.



## Research Article

# Computing Vertex-Based Eccentric Topological Descriptors of Zero-Divisor Graph Associated with Commutative Rings

Abdullah Ali H. Ahmadini,<sup>1</sup> Ali N. A. Koam ,<sup>1</sup> Ali Ahmad ,<sup>2</sup> Martin Bača ,<sup>3</sup>  
and Andrea Semaničová–Feňovčíková <sup>3</sup>

<sup>1</sup>Department of Mathematics, College of Science, Jazan University, New Campus, Jazan 2097, Saudi Arabia

<sup>2</sup>College of Computer Science and Information Technology, Jazan University, Jazan, Saudi Arabia

<sup>3</sup>Department of Applied Mathematics and Informatics, Technical University, Košice 04200, Slovakia

Correspondence should be addressed to Ali Ahmad; ahmadsms@gmail.com

Received 15 June 2020; Accepted 27 July 2020; Published 24 August 2020

Guest Editor: Muhammad Javaid

Copyright © 2020 Abdullah Ali H. Ahmadini et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The applications of finite commutative ring are useful substances in robotics and programmed geometric, communication theory, and cryptography. In this paper, we study the vertex-based eccentric topological indices of a zero-divisor graphs of commutative ring  $\mathbb{Z}_{p^2} \times \mathbb{Z}_q$ , where  $p$  and  $q$  are primes.

## 1. Introduction

One of the most significant issues in science is to change over chemical structure into numerical molecular descriptors that are pertinent to the physical, chemical, or organic properties. Atomic structure is one of the essential ideas of science since properties and chemical and organic practices of atoms are controlled by it.

Molecular descriptors called topological indices are graph invariants that play a significant job in science, and engineering, since they can be connected with huge physico-chemical properties of particles. We utilize topological descriptors during the time spent associating the chemical structures with different attributes, for example, boiling points and molar heats of formation. The computation of these topological descriptors for various chemical graphs is a very attractive direction for researchers. The chemical structure of a molecule is represented by molecular descriptors.

Atoms and atomic structures are frequently displayed by a molecular graph. An atomic structure is a graph in which vertices are atoms and edges are its atomic bonds. In this manner, a topological descriptor is a numeric amount related with a graph which portrays the topology of graph and

its invariants. There are some significant classes of topological descriptors and related polynomials which can be seen [1–9].

## 2. Definitions and Notations

Let  $V(G)$  and  $E(G)$  be the set of vertices and edges of connected graph  $G$ , respectively. The basic notations and definitions are taken from the book [10]. Let  $d_\mu$  be the degree of vertex  $\mu$  and  $d(\mu, \nu)$  be the distance between two vertices  $\mu$  and  $\nu$ . In mathematics, eccentricity is defined as

$$\varepsilon(\mu) = \max\{d(\mu, \nu) : \nu \in V(G)\}. \quad (1)$$

In [11], Sharma et al. introduced “eccentric connectivity index,” and the general formula of eccentric connectivity index is defined as

$$\xi(G) = \sum_{\nu \in V(G)} d_\nu \varepsilon(\nu). \quad (2)$$

Detail of applications and results of eccentric connectivity index can be seen in [12–15]. Farooq and Malik [16] introduced the “total eccentricity index” and defined as

$$\zeta(G) = \sum_{v \in V(G)} \varepsilon(v). \quad (3)$$

The “first Zagreb index” (and its new version by Ghorbani and Hosseinzadeh [17]) of a graph  $G$  was studied in [18] and defined as follows:

$$M_1^*(G) = \sum_{v \in V(G)} \varepsilon(v)^2. \quad (4)$$

The “eccentric connectivity polynomial” [19, 20], “augmented eccentric connectivity index” [21–23], “connective eccentric index” [23], “Ediz eccentric connectivity index,” and “reverse eccentric connectivity index” [24, 25] are defined in equations (5)–(8) and (9), respectively.

$$\text{ECP}(G, x) = \sum_{v \in V(G)} d(v)x^{\varepsilon(v)}, \quad (5)$$

$$\xi^{ac}(G) = \sum_{v \in V(G)} \frac{M(v)}{\varepsilon(v)}, \quad (6)$$

$$\xi^C(G) = \sum_{v \in V(G)} \frac{d(v)}{\varepsilon(v)}, \quad (7)$$

$$E\zeta(G) = \sum_{v \in V(G)} \frac{S(v)}{\varepsilon(v)}, \quad (8)$$

$$\text{Re}\zeta(G) = \sum_{v \in V(G)} \frac{\varepsilon(v)}{S(v)}, \quad (9)$$

where  $M(v) = \prod_{\mu \in N(v)} d_\mu$  and  $S(v) = \sum_{\mu \in N(v)} d_\mu$ .  $N(v)$  denotes the set of all vertices adjacent  $v$ .

### 3. Results and Discussion

Beck [26] defined the zero-divisor graph as “for a commutative ring with identity  $\mathbb{R}$  and set of its all zero divisors  $\mathbb{Z}(\mathbb{R})$ ,” its zero-divisor graph  $G(\mathbb{R})$  is constructed as  $\alpha, \beta \in V(G(\mathbb{R})) = G(\mathbb{R})$  and  $(\alpha, \beta) \in E(G(\mathbb{R}))$ , if  $\alpha\beta = 0$ . For further study of zero divisor, see [27–32].

Let  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  denotes the zero-divisor graph of the commutative ring  $\mathbb{Z}_{p^2} \times \mathbb{Z}_q$ . It is defined as follows: for  $a \in \mathbb{Z}_{p^2}$  and  $b \in \mathbb{Z}_q$ ,  $(a, b) \notin V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))$  if and only if  $a \neq kp$ ,  $1 \leq k \leq p-1$ ,  $b \neq 0$ . Let  $J = \{(a, b) \notin V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) : a \neq kp, 1 \leq k \leq p-1, b \neq 0\}$ , then  $|J| = (p^2 - p)(q - 1)$ . The vertices of the set  $J$  are the nonzero divisors of the commutative ring  $\mathbb{Z}_{p^2} \times \mathbb{Z}_q$ . Also,  $(0, 0) \in \mathbb{Z}_{p^2} \times \mathbb{Z}_q$  is a nonzero divisor. Therefore, the total number of nonzero divisors is  $|J| + 1 = (p^2 - p)(q - 1) + 1 = p^2q - p^2 - pq + p + 1$ . There are  $p^2q$  total vertices of the commutative ring  $\mathbb{Z}_{p^2} \times \mathbb{Z}_q$ . Hence, there are  $p^2q - (p^2q - p^2 - pq + p + 1) = p^2 + pq - p - 1$  total number of zero divisors. This implies that the order of the zero-divisor graph  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is  $p^2 + pq - p - 1$ , i.e.,  $|V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))| = p^2 + pq - p - 1$ .

From the definition, we partitioned the vertex set of the graph  $G$  into the following four partitions corresponding to their degrees:

$$\begin{aligned} P_1 &= \{(a, b) : b \in \{1, 2, 3, \dots, q-1\}\}, \\ P_2 &= \{(a, b) : a = kp, 1 \leq k \leq p-1, b \in \{1, 2, 3, \dots, q-1\}\}, \\ P_3 &= \{(a, 0) : a \in \mathbb{Z}_{p^2} \setminus \{0, p, 2p, \dots, (p-1)p\}\}, \\ P_4 &= \{(a, 0) : a = kp, 1 \leq k \leq p-1\}. \end{aligned} \quad (10)$$

This implies that  $V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) = P_1 \cup P_2 \cup P_3 \cup P_4$ . Let  $d_A(a)$  denotes the degree of a vertex  $a$  in  $A$  and  $d(A, B)$  denotes the distance between the vertices of two sets  $A$  and  $B$ . It is easy to see that  $|P_1| = q - 1$ ,  $|P_2| = (p - 1)(q - 1)$ ,  $|P_3| = p^2 - p$ ,  $|P_4| = p - 1$ ,  $d_{P_1}(a) = p^2 - 1$ ,  $d_{P_2}(a) = p - 1$ ,  $d_{P_3}(a) = q - 1$ , and  $d_{P_4}(a) = pq - 2$ . In the next theorem, we determined  $\varepsilon(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))$ .

**Lemma 1.** *The eccentricity of the vertices of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is 2 or 3.*

*Proof.* From the definition of zero-divisor graph, the vertices of the partition set  $P_1$  are adjacent with the vertices of partitions  $P_3$  and  $P_4$ , i.e.,  $d(P_1, P_3) = d(P_1, P_4) = 1$ . Similarly, it is observed that  $d(P_2, P_4) = 1$  and  $d(P_4, P_4) = 1$ , for any  $a \neq b \in P_4$ . Now,  $d(P_1, P_2) = d(P_1, P_4) + d(P_4, P_2) = 2$  and  $d(P_3, P_4) = d(P_3, P_1) + d(P_1, P_4) = 2$ . Also, the distance between any two different vertices of the same set of partition  $P_1, P_2$ , and  $P_3$  is 2. This shows that the maximum distance of the vertices of the partitions  $P_1$  and  $P_4$  is 2; therefore, their eccentricity is 2, i.e.,  $\varepsilon(P_1) = \varepsilon(P_4) = 2$ .

Since partitions  $P_2$  and  $P_3$  are at distance 1 with partitions  $P_4$  and  $P_1$ , respectively, and  $d(P_1, P_4) = 1$ , it follows that  $d(P_2, P_3) = d(P_2, P_4) + d(P_4, P_1) + d(P_1, P_3) = 3$  or  $d(P_2, P_3) = d(P_2, P_1) + d(P_1, P_3) = 2 + 1 = 3$ . This implies that  $\varepsilon(P_2) = \varepsilon(P_3) = 3$ . This shows that the eccentricity of the vertices of  $G$  is 2 or 3. This completes the proof.  $\square$

We summarize the above discussion in Table 1.

We determined the eccentric connectivity index of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  in the following theorem.

**Theorem 1.** *The eccentric connectivity index of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is*

$$\xi(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) = 10p^2q - 8p^2 - 11pq + 5p + q + 3. \quad (11)$$

*Proof.* Using the values from Table 1, formula (2) implies that

$$\begin{aligned} \xi(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \sum_{v \in V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))} d_v \varepsilon(v) \\ &= 3(p^2 - p)(q - 1) + 2(q - 1)(p^2 - 1) \\ &\quad + 2(p - 1)(pq - 2) \\ &\quad + 3(p - 1)^2(q - 1) = 10p^2q - 8p^2 - 11pq + 5p + q + 3. \end{aligned} \quad (12)$$

We arrive at the desired result.

TABLE 1: The summary of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$ .

Partition sets of vertices	Degree	Eccentricity	Frequency
$P_1$	$p^2 - 1$	2	$q - 1$
$P_2$	$p - 1$	3	$(p - 1)(q - 1)$
$P_3$	$q - 1$	3	$p(p - 1)$
$P_4$	$pq - 2$	2	$p - 1$

By using Lemma 1 and Table 1 in equations (3) and (4), we obtain the total eccentricity index and the first eccentricity Zagreb index for  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  in the following corollaries.  $\square$

**Corollary 1.** *The total eccentricity index of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is given by*

$$\zeta(\mathbb{Z}_{p^2} \times \mathbb{Z}_q) = 3p^2 + 3pq - 4p - q - 1. \quad (13)$$

**Corollary 2.** *The first eccentricity Zagreb index of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is given by*

$$M_1^*(\mathbb{Z}_{p^2} \times \mathbb{Z}_q) = 9p^2 + 9pq - 14p - 5q + 1. \quad (14)$$

**Theorem 2.** *The eccentric connectivity polynomial of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is*

$$\begin{aligned} \text{ECP}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q), x) &= (2p^2q - 2p^2 - 3pq + 3p + q - 1)x^3 \\ &\quad + (2p^2q - p^2 - pq - 2p - q + 3)x^2. \end{aligned} \quad (15)$$

*Proof.* By using the degree and its corresponding eccentricity for each partition set in Table 1, equation (5) gives

$$\begin{aligned} \text{ECP}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)x) &= \sum_{v \in V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)x)} d_v x^{\varepsilon(v)} \\ &= (p^2 - p)(q - 1)x^3 + (p^2 - 1)(q - 1)x^2 + (pq - 2)(p - 1)x^2 + (q - 1)(p - 1)^2x^3 \\ &= (2p^2q - 2p^2 - 3pq + 3p + q - 1)x^3 + (2p^2q - p^2 - pq - 2p - q + 3)x^2. \end{aligned} \quad (16)$$

**Theorem 3.** *The augmented eccentric connectivity index of  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is*

$$\begin{aligned} \xi^{ac}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \left( \frac{p-1}{3} + \frac{(q-1)^{p^2-p}}{2} \right) (pq-2)^{p-1} (q-1) \\ &\quad + \left( \frac{p}{3} + \frac{p^{(q-1)(q-1)}}{2} \right) (p^2-1)^{q-1} (p-1). \end{aligned} \quad (17)$$

*Proof.* Inserting values from the proof of Lemma 1 and Table 1 to equation (6), we obtain

$$\begin{aligned} \xi^{ac}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \sum_{v \in V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))} \frac{M(v)}{\varepsilon(v)} = (p-1)(q-1) \frac{(pq-2)^{p-1}}{3} + (q-1) \frac{(pq-2)^{p-1} (q-1)^{p^2-p}}{2} \\ &\quad + (p^2-p) \frac{(p^2-1)^{q-1}}{3} + (p-1) \frac{(p-1)^{(p-1)(q-1)} (p^2-1)^{q-1} (pq-2)^{p-2}}{2}. \end{aligned} \quad (18)$$

After simplification, we get

$$\begin{aligned} \xi^{ac}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \left( \frac{p-1}{3} + \frac{(q-1)^{p^2-p}}{2} \right) (pq-2)^{p-1} (q-1) \\ &\quad + \left( \frac{p}{3} + \frac{p^{(p-1)(q-1)} (pq-2)^{p-2}}{2} \right) (p^2-1)^{q-1} (p-1). \end{aligned} \quad (19)$$

$\square$

**Theorem 4.** The connective eccentric index of the graph  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is

$$\xi^C(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) = \frac{(p-1)(10pq - 7p + q - 7)}{6}. \quad (20)$$

*Proof.* We apply the values of degrees and their eccentricity from Table 1. Then, formula (7) gives

$$\begin{aligned} \xi^C(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \sum_{v \in V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))} \frac{d_v}{\varepsilon(v)} \\ &= \frac{(p^2 - p)(q - 1)}{3} + \frac{(p^2 - 1)(q - 1)}{2} + \frac{(p - 1)(pq - 2)}{2} \\ &\quad + \frac{(p - 1)^2(q - 1)}{3} = \frac{(p - 1)(10pq - 7p + q - 7)}{6}. \end{aligned} \quad (21)$$

□

**Theorem 5.** The Ediz eccentric connectivity index of the graph  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is

$$E\check{\zeta}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) = \frac{(p^2 + pq + p - 2)(p - 1)^2(q - 1)}{3} + \frac{(p - 1)(3p^2q + 2pq^2 - 2p^2 - 7pq + p - 2q + 6)}{2}. \quad (22)$$

*Proof.* By using Table 1 and equation (8), we get

$$\begin{aligned} E\check{\zeta}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \sum_{v \in V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))} \frac{S(v)}{\varepsilon(v)} \\ &= \frac{(p - 1)^2(q - 1)(pq - 2)}{3} + \frac{(p - 1)^3(q - 1) + (p - 1)(p^2 - 1)(q - 1) + (p - 2)(pq - 2)}{2} \\ &\quad + \frac{(p - 1)(pq - 2)(q - 1) + (q - 1)^2(p^2 - p)}{2} + \frac{(p^2 - p)(q - 1)(p^2 - 1)}{3}. \end{aligned} \quad (23)$$

After simplification, we get

$$E\check{\zeta}(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) = \frac{(p^2 + pq + p - 2)(p - 1)^2(q - 1)}{3} + \frac{(p - 1)(3p^2q + 2pq^2 - 2p^2 - 7pq + p - 2q + 6)}{2}. \quad (24)$$

□

The following theorem determines the reverse eccentric connectivity index of the graph  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$ .

**Theorem 6.** The reverse eccentric connectivity index of the graph  $\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)$  is

$$\begin{aligned} \operatorname{Re} \zeta(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z})) &= \frac{3(p^2 + pq + p - 2)}{p(p+1)(p-1)^2(q-1)(pq-2)} \\ &+ \frac{2}{p(p-1)^2(q-1) + (p-2)(pq-2)} + \frac{2}{(p-1)(q-1)(2pq-p-2)}. \end{aligned} \tag{25}$$

*Proof.* For the values given in Table 1, formula (9) implies that

$$\begin{aligned} \operatorname{Re} \zeta(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) &= \sum_{v \in V(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q))} \frac{\varepsilon(v)}{S(v)} \\ &= \frac{3}{(p-1)^2(q-1)(pq-2)} + \frac{2}{(p-1)^3(q-1) + (p-1)(p^2-1)(q-1) + (p-2)(pq-2)} \\ &+ \frac{2}{(p-1)(pq-2)(q-1) + (q-1)^2(p^2-p)} + \frac{3}{(p^3-p)(q-1)(p^2-1)}. \end{aligned} \tag{26}$$

After simplification, we get

$$\operatorname{Re} \zeta(\Gamma(\mathbb{Z}_{p^2} \times \mathbb{Z}_q)) = \frac{3(p^2 + pq + p - 2)}{p(p+1)(p-1)^2(q-1)(pq-2)} + \frac{2}{p(p-1)^2(q-1) + (p-2)(pq-2)} + \frac{2}{(p-1)(q-1)(2pq-p-2)}, \tag{27}$$

and we are done. □

### Data Availability

No data were used to support this study.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Acknowledgments

This work was supported by the Slovak Research and Development Agency under contract no. APVV-19-0153 and by VEGA 1/0233/18.

### References

- [1] A. Ahmad, "On the degree based topological indices of benzene ring embedded in p-type-surface in 2D network," *Hacettepe Journal of Mathematics and Statistics*, vol. 47, no. 1, pp. 9–18, 2018.
- [2] M. Bača, J. Horváthová, M. Mokrišová, and A. Suhányiová, "On topological indices of fullerenes," *Applied Mathematics and Computation*, vol. 251, pp. 154–161, 2015.
- [3] M. Javaid, J.-B. Liu, M. A. Rehman, and S. Wang, "On the certain topological indices of titania nanotube  $\text{TiO}_2 [m, n]$ ," *Zeitschrift für Naturforschung A*, vol. 72, no. 7, pp. 647–654, 2017.
- [4] A. Gravovac, M. Ghorbani, and M. A. Hosseinzadeh, "Computing fifth geometric-arithmetic index  $ABC_4$  of nanostar dendrimers, Optoelectron," *Advanced Materials-Rapid Communications*, vol. 4, no. 9, pp. 1419–1422, 2010.
- [5] S. Hayat and M. Imran, "Computation of topological indices of certain networks," *Applied Mathematics and Computation*, vol. 240, pp. 213–228, 2014.
- [6] J.-B. Liu, S. Javed, M. Javaid, and K. Shabbir, "Computing first general Zagreb index of operations on graphs," *IEEE Access*, vol. 7, pp. 47494–47502, 2019.
- [7] M. F. Nadeem, S. Zafar, and Z. Zahid, "On certain topological indices of the line graph of subdivision graphs," *Applied Mathematics and Computation*, vol. 271, pp. 790–794, 2015.
- [8] M. F. Nadeem, S. Zafar, and Z. Zahid, "On topological properties of the line graphs of subdivision graphs of certain nanostructures," *Applied Mathematics and Computation*, vol. 273, pp. 125–130, 2016.
- [9] J.-B. Liu, M. Javaid, and H. M. Awais, "Computing Zagreb indices of the subdivision-related generalized operations of graphs," *IEEE Access*, vol. 7, pp. 105479–105488, 2019.
- [10] K. H. Rosen, *Discrete Mathematics and Its Applications*, McGraw-Hill Education, New York City, NY, USA, Seventh edition, 2012.
- [11] V. Sharma, R. Goswami, and A. K. Madan, "Eccentric connectivity index: a novel highly discriminating topological descriptor for structure-property and structure-activity studies," *Journal of Chemical Information and Computer Sciences*, vol. 37, no. 2, pp. 273–282, 1997.
- [12] H. Dureja and A. K. Madan, "Topochemical models for prediction of cyclin-dependent kinase 2 inhibitory activity of indole-2-ones," *Journal of Molecular Modeling*, vol. 11, no. 6, pp. 525–531, 2005.
- [13] A. Ilic and I. Gutman, "Eccentric-connectivity index of chemical trees," *MATCH Communications in Mathematical and in Computer*, vol. 65, pp. 731–744, 2011.

- [14] V. Kumar and A. K. Madan, "Application of graph theory: prediction of cytosolic phospholipase A2 inhibitory activity of propan-2-ones," *Journal of Mathematical Chemistry*, vol. 39, no. 3-4, pp. 511-521, 2006.
- [15] B. Zhou, "On eccentric-connectivity index," *MATCH Communications in Mathematical and in Computer*, vol. 63, pp. 181-198, 2010.
- [16] R. Farooq and M. A. Malik, "On some eccentricity based topological indices of nanostar dendrimers," *Optoelectronics and Advanced Materials-Rapid Communications*, vol. 9, no. 5-6, pp. 842-849, 2015.
- [17] M. Ghorbani and M. Hosseinzadeh, "A new version of Zagreb indices," *Filomat*, vol. 26, no. 1, pp. 93-100, 2012.
- [18] X. Li and H. Zhao, "Trees with the first three smallest and largest generalized topological indices," *MATCH Communications in Mathematical and in Computer*, vol. 50, pp. 57-62, 2004.
- [19] M. Alaeiyan, J. Asadpour, and R. Mojarad, "A numerical method for MEC polynomial and MEC index of one-pentagonal carbon nanocones," *Fullerenes, Nanotubes and Carbon Nanostructures*, vol. 21, no. 10, pp. 825-835, 2013.
- [20] A. R. Ashrafi, M. Ghorbani, and M. A. Hossein-Zadeh, "The eccentric-connectivity polynomial of some graph operations," *Serdica Journal of Computing*, vol. 5, pp. 101-116, 2011.
- [21] A. Ali, Z. Raza, and A. A. Bhatti, "On the augmented Zagreb index," *Kuwait Journal of Science*, vol. 43, no. 2, pp. 123-138, 2016.
- [22] A. Ali, A. A. Bhatti, and Z. Raza, "The augmented Zagreb index, vertex connectivity, and matching number of graphs," *Bulletin of the Iranian Mathematical Society*, vol. 42, no. 2, pp. 417-425, 2016.
- [23] S. Gupta, M. Singh, and A. K. Madan, "Connective eccentricity index: a novel topological descriptor for predicting biological activity," *Journal of Molecular Graphics and Modelling*, vol. 18, no. 1, pp. 18-25, 2000.
- [24] S. Ediz, "On the ediz eccentric connectivity index of a graph," *Optoelectronics and Advanced Materials-Rapid Communications*, vol. 5, no. 11, pp. 1263-1264, 2011.
- [25] S. Ediz, "Reverse eccentric connectivity index," *Optoelectronics and Advanced Materials-Rapid Communications*, vol. 6, no. 5-6, pp. 664-667, 2012.
- [26] I. Beck, "Coloring of commutative rings," *Journal of Algebra*, vol. 116, no. 1, pp. 208-226, 1988.
- [27] D. F. Anderson and A. Badawi, "On the zero-divisor graph of a ring," *Communications in Algebra*, vol. 36, no. 8, pp. 3073-3092, 2008.
- [28] S. Akbari and A. Mohammadian, "On the zero-divisor graph of a commutative ring," *Journal of Algebra*, vol. 274, no. 2, pp. 847-855, 2004.
- [29] D. F. Anderson and S. B. Mulay, "On the diameter and girth of a zero-divisor graph," *Journal of Pure and Applied Algebra*, vol. 210, no. 2, pp. 543-550, 2008.
- [30] T. Asir and T. T. Chelvam, "On the total graph and its complement of a commutative ring," *Communications in Algebra*, vol. 41, no. 10, pp. 3820-3835, 2013.
- [31] K. Samei, "The zero-divisor graph of a reduced ring," *Journal of Pure and Applied Algebra*, vol. 209, no. 3, pp. 813-821, 2007.
- [32] T. T. Chelvam and T. Asir, "A note on total graph of  $\mathbb{Z}_n$ ," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 14, no. 1, pp. 1-7, 2011.



## Research Article

# Model Test and Numerical Simulation of the Mudcake Thickness Effect on the Bearing Capacity of Vertically Loaded Single Piles

Ziguang Zhang <sup>1</sup>, Jian Cui,<sup>2</sup> Maoyan Ma <sup>1</sup> and Peisheng Xi<sup>1</sup>

<sup>1</sup>Anhui Province Key Laboratory of Building Structure and Underground Engineering, Anhui Jianzhu University, Hefei, China

<sup>2</sup>Anhui Road and Bridge Engineering Group Co., Ltd, Hefei, China

Correspondence should be addressed to Maoyan Ma; maoyanma2019@163.com

Received 4 July 2020; Accepted 29 July 2020; Published 20 August 2020

Guest Editor: Shaohui Wang

Copyright © 2020 Ziguang Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mudcake is an important factor affecting the bearing capacity of vertically loaded single piles. The model test and numerical calculation were used to analyse the stress and displacement characteristics of vertically loaded single piles under the conditions of different mudcake thickness and loading capacity. The effect of mudcake thickness on bearing capacity has been analyzed systematically, and its mechanism and changing laws were revealed. Eventually, the concept of the mudcake thickness effect was proposed, which must be considered fully during the construction of mud protection bored piles. The research results have important scientific value and guiding significance for understanding of the influence of pile side mudcake on the bearing capacity of vertically loaded single piles.

## 1. Introduction

Mudcake is an important factor that affects the normal performance of vertically loaded single pile [1–3]. It is generally accepted that the existence of mudcake weakens pile side resistance and leads to the decrease in the bearing capacity of a single pile. It is of great scientific value and practical significance to correctly understand the influence of mudcake thickness on the bearing capacity of a single pile.

The effect of mudcake on the bearing capacity of a vertically loaded single pile has received wide attention. Hosoi [4] found by experiment that the friction resistance between concrete and mudcake soil was significantly lower than that between concrete and the same type of mudcake-free soil and proposed that if the mudcake on the side of a mud protection bored pile could not be completely removed, the influence of mudcake on the friction resistance of the pile should be fully considered. Majana et al. [5] carried out a laboratory model test to study the influence of mudcake thickness on the load transfer law of cast-in-place pile in saturated sandy soil and believed that attenuation of pile side resistance was mainly due to excessive mudcake thickness. Based on a bridge reconstruction project, Wu [6] expounded the mechanism of

the influence of mudcake thickness on the bearing capacity of a single pile by an on-site test. Hu et al. [7] studied the mudcake thickness influence on the bearing capacity of a single pile by the method of numerical calculation and model test on the special condition of red bed mudstone, put forward the concept of single pile bearing capacity correction coefficient, and modified the single pile bearing capacity calculation formula in the Technical Specification for the Building Pile Foundation (JGJ 94-2008); however, the pile length and pile diameter which also are influence factors have not been considered adequately in their research. Chen et al. [8] discussed the influence of mudcake on pile-soil rough contact surface mechanical characteristics including the relationship of shear stress and displacement and peak shear strength and peak friction angle by the model test. It was exposed that the shear stress and shear displacement relationship curves showed characteristics of softening under the condition of no mudcake and hardening under the condition of mudcake. The thickness of mudcake was an important factor for the attenuation of shear strength of the contact surface. Chen et al. [9] discussed the changing laws of displacement of the pile top, lateral resistance, and tip resistance of vertically loaded single piles with mudcake by the model test, and pointed out

that the cement as mud additive could effectively improve the bearing capacity of the single pile and reduced the mudcake influence on pile side friction.

In conclusion, many scholars and engineers have carried out multiple studies on pile side mudcake. However, the comprehensive influence of mudcake thickness on the bearing capacity of a vertically loaded single pile is still insufficient. In view of this, the model test and numerical calculation were used in this paper to analyze stress and displacement characteristics of vertically loaded single piles under the conditions of different mudcake thickness and load. The effect of mudcake thickness on the bearing capacity of a vertically loaded single pile has been analyzed systematically, and its mechanism and changing laws were revealed. Eventually, the concept of the mudcake thickness effect was proposed, which must be considered fully during the construction of mud protection bored piles.

## 2. Model Test

### 2.1. Test Overview

**2.1.1. Test Device.** The model test of the bearing performance of a vertically loaded single pile was conducted in the Anhui Province Key Laboratory of Building Structure and Underground Engineering. The model test was carried out in a self-made test box which was composed of two cylindrical steel drums welded by semicylindrical high-strength steel plates. The steel drums are 1.8 m in height and 1.2 m in diameter, which not only eliminate the influence of test boundary conditions but also ensure the test stiffness conditions. The test device is mainly composed of a loading device (jack), reaction device (reaction frame), and measuring device (dial indicator and strain gauge) (Figure 1).

### 2.1.2. Test Materials

**(1) Test Sand.** The sand used for the test was taken from the project site of an industrial park in Hefei, which is located in the middle of the Jianghuai Hills. The selected sand is relatively pure medium sand, and its particle grading analysis curve is as shown in Figure 2.

In the model test, sand was filled in layers and compacted manually. In order to evenly fill the foundation soil in each layer, the thickness of compaction was controlled. Each layer was filled to 10 cm and compacted to 8 cm. Samples for geotechnical tests were taken by cutting ring from the top surface and the levels of -30 cm, -60 cm, -90 cm, and -120 cm of the sand for testing. Physical and mechanical parameters of sand are shown in Table 1.

**(2) Model Pile.** Based on the reduced-scale of the model test and the similar percentage of the material stiffness of the pile material, the hollow Perspex tube with a diameter of 50 mm and an inner wall thickness of 5 mm was selected for the model pile in this test. The length of the model pile is 1.2 m, and the elastic modulus is 3.48 GPa. In order to facilitate loading, the actual length of the model pile used in the test is 1.4 m. During the test, the model pile exceeded the upper

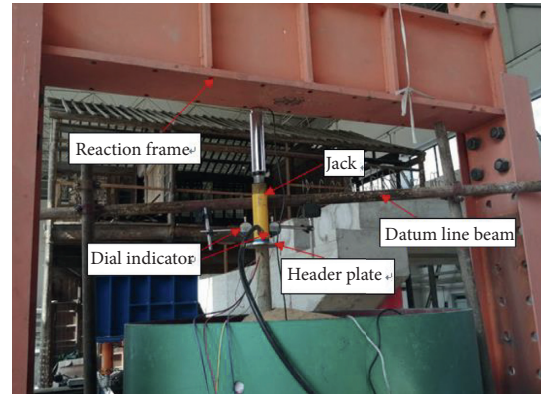


FIGURE 1: Model test apparatus.

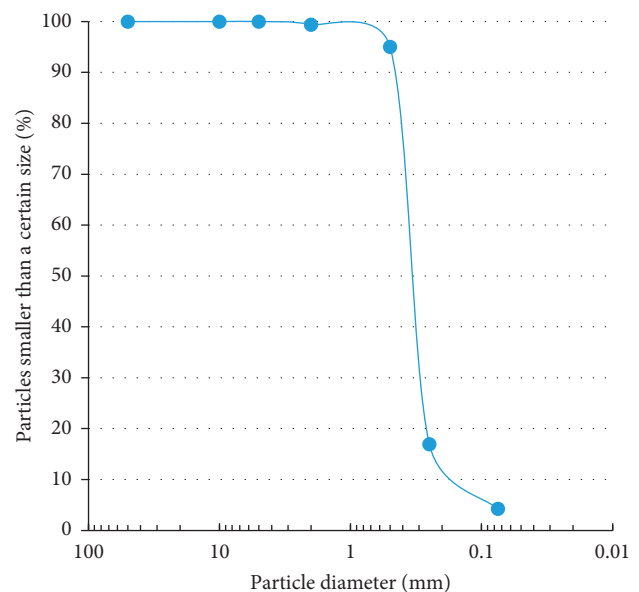


FIGURE 2: Grading curve of sand particles for test.

surface of the test soil by 0.2 m. Strain gauges were pasted inside both the pile top and pile bottom, and axial forces of the pile top and pile bottom were measured during the test. Model piles are shown in Figure 3.

**(3) Mudcake Thickness Control.** The mudcake used in this test was the mud collected at a project site in Hefei, and the mudcake density is  $1.5 \text{ g/cm}^3$ . The method of quality control was used to control the thickness of mudcake, and the manufacturing process of a single pile with mudcake was as follows: (1) the quality and external area of the model pile were weighed and calculated without mudcake; (2) the predetermined quality of piles was calculated with mudcake; (3) the surface of test piles was covered evenly with mud. After several hours each time, when the mud was solidified and could be touched gently, the test pile was covered with another layer of mud evenly. The total mass of the pile with mudcake was weighed after each layer of mud was brushed until the predetermined mass was reached; and (4) it was covered with a layer of plastic wrap when the mudcake was dry and placed in the shade for curing (Figure 4).

TABLE 1: Physical and mechanical parameters of sand used in the test.

Unit weight ( $\text{kN/m}^3$ )	Compression modulus (MPa)	Internal friction angle ( $^\circ$ )	Cohesion (kPa)
18.2	40	33	0



FIGURE 3: Model piles.



FIGURE 4: Mudcake thickness control.

**2.1.3. Test Process.** In the model test of the bearing capacity of a vertically loaded single pile, the mudcake thickness of the pile was selected as 0 mm, 0.5 mm, 1.0 mm, 1.5 mm, and 2.0 mm, respectively. The load was applied by means of a hand-operated separate hydraulic jack. The output oil pressure was measured by using an oil pressure meter installed on the oil road and controlled according to the load-oil pressure curve calibrated on a pressure machine. During the test, the internal force of the pile top and bottom was measured in real time by using the strain gauge affixed to the pile top and bottom (Figure 5).

The load started from 0.5 kN and increased by 0.5 kN step by step. Loading was stopped when the pile displacement increased sharply. The vertical displacement of the pile top was measured at 5 min, 10 min, 15 min, 30 min, 45 min, and 60 min, respectively, after each load level, and when the

settlement of each level was less than 0.1 mm, within 30 min, the next load level began [10–12].

## 2.2. Test Results

**2.2.1. Stress Characteristics of the Vertically Loaded Single Pile.** Based on the measured data of strain gauges at the top and bottom of the pile in the model test, curves of load and pile tip resistance changing with time can be obtained under the conditions of different mudcake thickness (Figure 6).

It can be seen that the maximum stable bearing capacity of the pile without mudcake is about 5800 N, and the pile tip resistance is within the range of 7%–10% of the load, showing significant bearing characteristics of the end bearing friction pile. The existence of pile side mudcake reduces the bearing capacity of the single pile significantly. When the thickness of mudcake is 0.5 mm, 1.0 mm, 1.5 mm, and 2.0 mm, the



maximum stable bearing capacity of the single pile is 4000 N, 5400 N, 4200 N, and 3800 N, respectively.

The vertical load of the single pile is borne by pile side resistance and pile tip resistance. Pile side resistance can be regarded as the difference between load and pile tip resistance. In this paper, the difference between the arithmetic mean value of the two measured values of load in the model test and that of the two measured values of the tip resistance was the value of single pile side resistance, and Figure 7 shows the measured values of the load and the side resistance change with time under the conditions of different mudcake thickness.

It can be seen that the existence of mudcake on the side of the pile reduces the overall side resistance of the vertically loaded single pile significantly. When the mudcake thickness is 0.0 mm, 0.5 mm, 1.0 mm, 1.5 mm, and 2.0 mm, the percentage of the side resistance is 90%–93%, 60%–80%, 75%–82%, 45%–55%, and 78%–82% of the load, respectively. The existence of mudcake on the pile side changes the direct combination between the pile and the soil mass, affects the interaction between the pile and the soil mass on the pile side, and then, causes the reduction of the pile side friction.

*2.2.2. Displacement Characteristics of the Vertically Loaded Single Pile.* Based on the model test, Q-S curves of pile top load and pile top vertical displacement with different mudcake thickness ( $k$ ) can be obtained (Figure 8). It can be seen that the Q-S curves of the single pile drop slowly. Q-S curves of the pile with mudcake deviate from the curve of pile without mudcake at the beginning of loading. With the increase in mudcake thickness, the Q-S curve of the mudskin pile becomes steeper and steeper.

The bearing capacity of a vertically loaded single pile is closely related to the vertical displacement of the pile top, which is an important index of the bearing capacity of a single pile. With the increasing requirements of engineering structures on foundation deformation and differential deformation, the method of determining the bearing capacity of the vertically loaded single pile based on the vertical displacement control standard of the pile top is gradually favored by the majority of experts, scholars, and engineers [13–15]. The vertical displacement control standard of the pile top is 2.5 mm ( $0.05d$ ,  $d$  is the pile diameter, the same as below), 3.0 mm ( $0.06d$ ), and 3.5 mm ( $0.07d$ ), respectively. Under the abovementioned standard, the bearing capacity of a single pile with different mudcake thickness and the percentage of bearing capacity of a single pile with mudcake accounting for that without mudcake were obtained (Table 2 and Figure 9).

With the increase of mudcake thickness, the bearing capacity of the vertically loaded single pile decreases slowly at first, then decreases suddenly, and finally, tends to be stable. The larger the displacement control standard is, the more remarkable the change gets. When the control standard is 2.5 mm, 3.0 mm, and 3.5 mm, respectively, the bearing capacity of the single pile with mudcake of 0.5 mm accounts for about 96%, 98%, and 98%, respectively, of that of single pile without mudcake. When the thickness of mudcake is 1.0 mm, the bearing capacity of the single pile



FIGURE 5: Stress acquisition instrument.

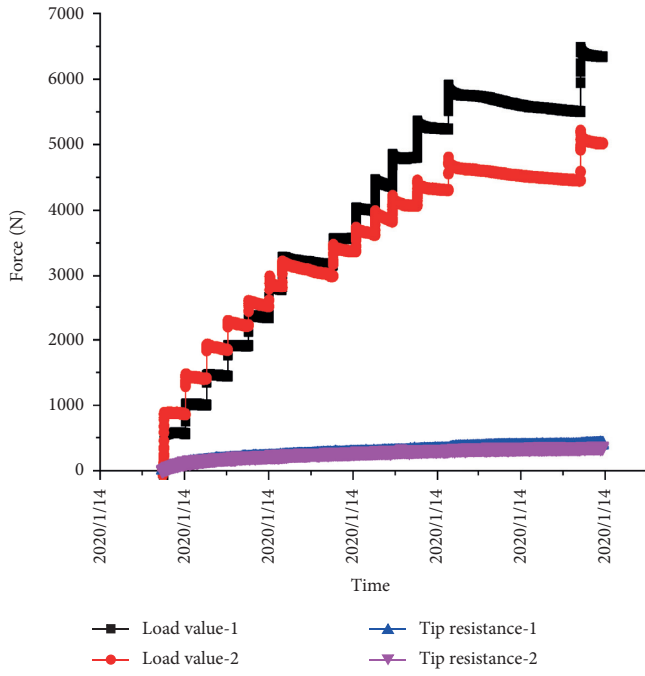
with mudcake accounts for about 76%, 81%, and 85%, respectively. When the thickness of mudcake reaches or exceeds 1.5 mm, the bearing capacity of the single pile with mudcake is within the range of 62%–65%, 67%–71%, and 72%–75%, respectively.

For a single pile with different thickness of mudcake, when the mudcake thickness is thin, the interaction between piles and soil is controlled mainly by the interface between the pile and the soil surrounding the pile, and the mudcake only acts as a lubricant. With the increase in the thickness of mudcake, single pile bearing capacity shows a trend of slow decline. With continuous increase of the thickness of mudcake, the binding effect of the soil around the mudcake on mudcake is becoming weaker and weaker and it is easier for the slip to happen between the pile and mudcake. The interaction between piles and soil is controlled by the interface between the pile and the mudcake, and at this stage, single pile bearing capacity rapidly decreases with the increase of the thickness of mudcake. When the mudcake thickness exceeds a certain value, the interaction between the pile and the soil mainly depends on the strength of the mudcake itself. At this stage, the bearing capacity of a single pile tends to be stable with the increase of the mudcake thickness.

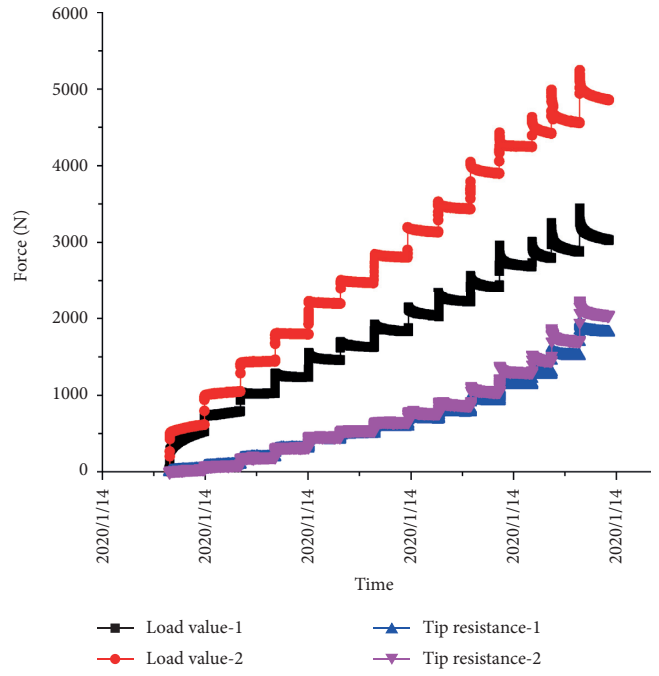
### 3. Numerical Calculation

*3.1. Calculation Model and Parameters.* Midas/GTS software is widely used in the pile performance calculation because it is suitable for large deformation calculation and contact processing and provides complete three-dimensional dynamic analysis and a variety of constitutive models [16–19].

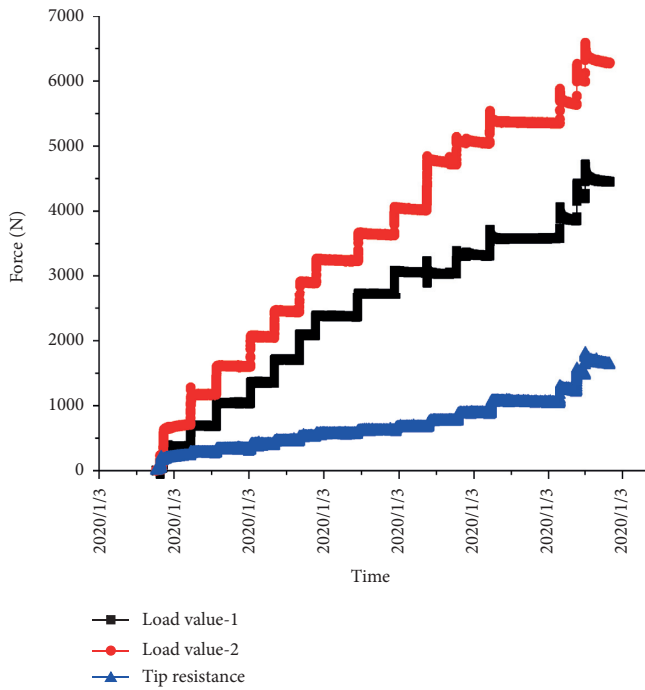
The pile side soil and mudcake were considered as ideal elastic-plastic material, and the pile was considered as linear elastic material in the calculation. The Mohr–Coulomb failure



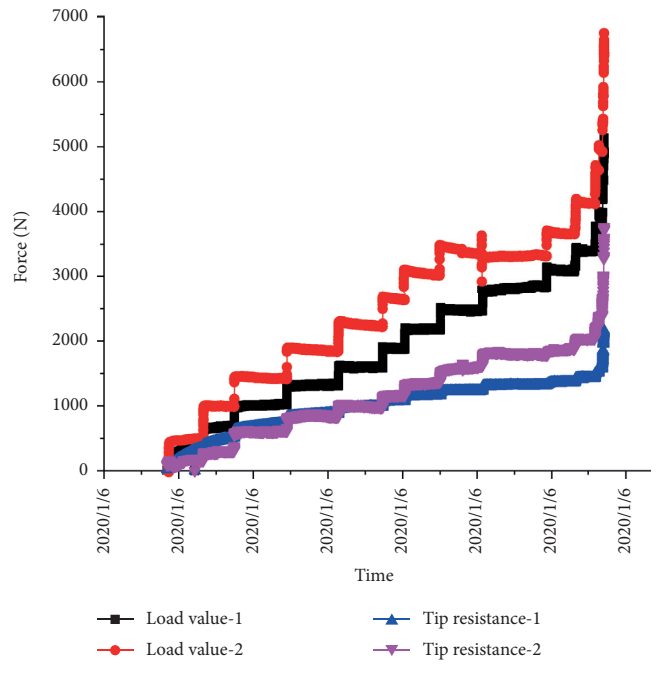
(a)



(b)



(c)



(d)

FIGURE 6: Continued.

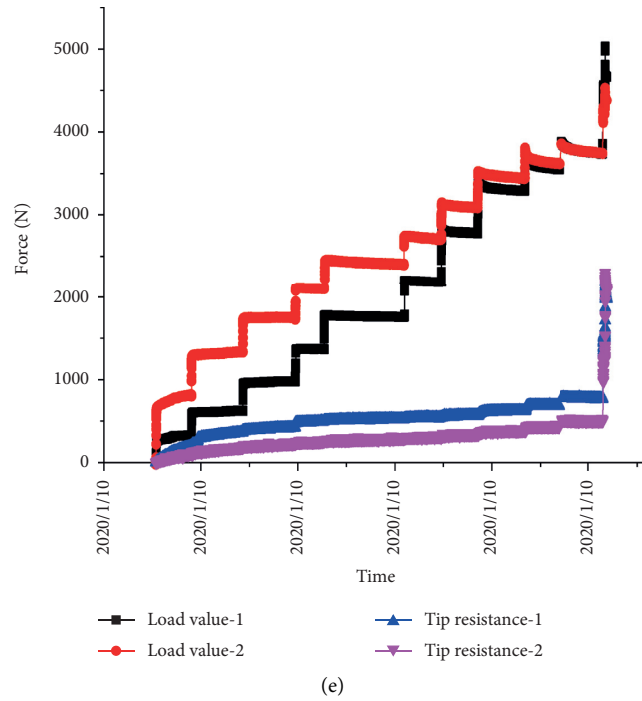


FIGURE 6: Change of load and tip resistance of the vertically loaded single pile with time. (a) 0 mm. (b) 0.5 mm. (c) 1.0 mm. (d) 1.5 mm. (e) 2.0 mm.

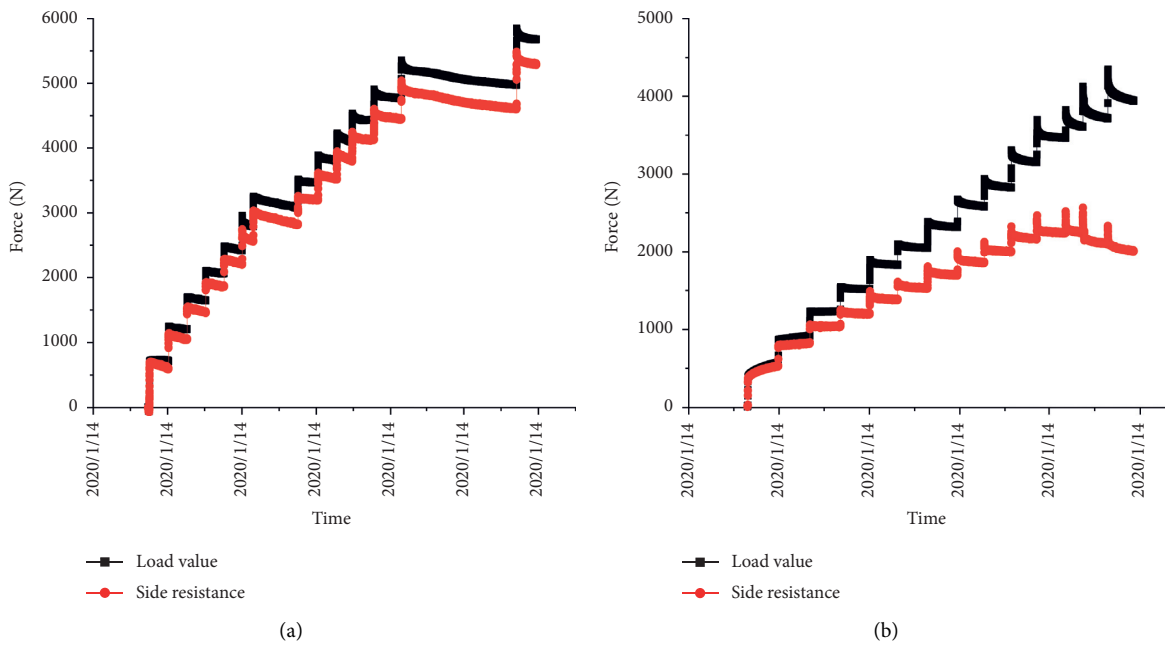


FIGURE 7: Continued.



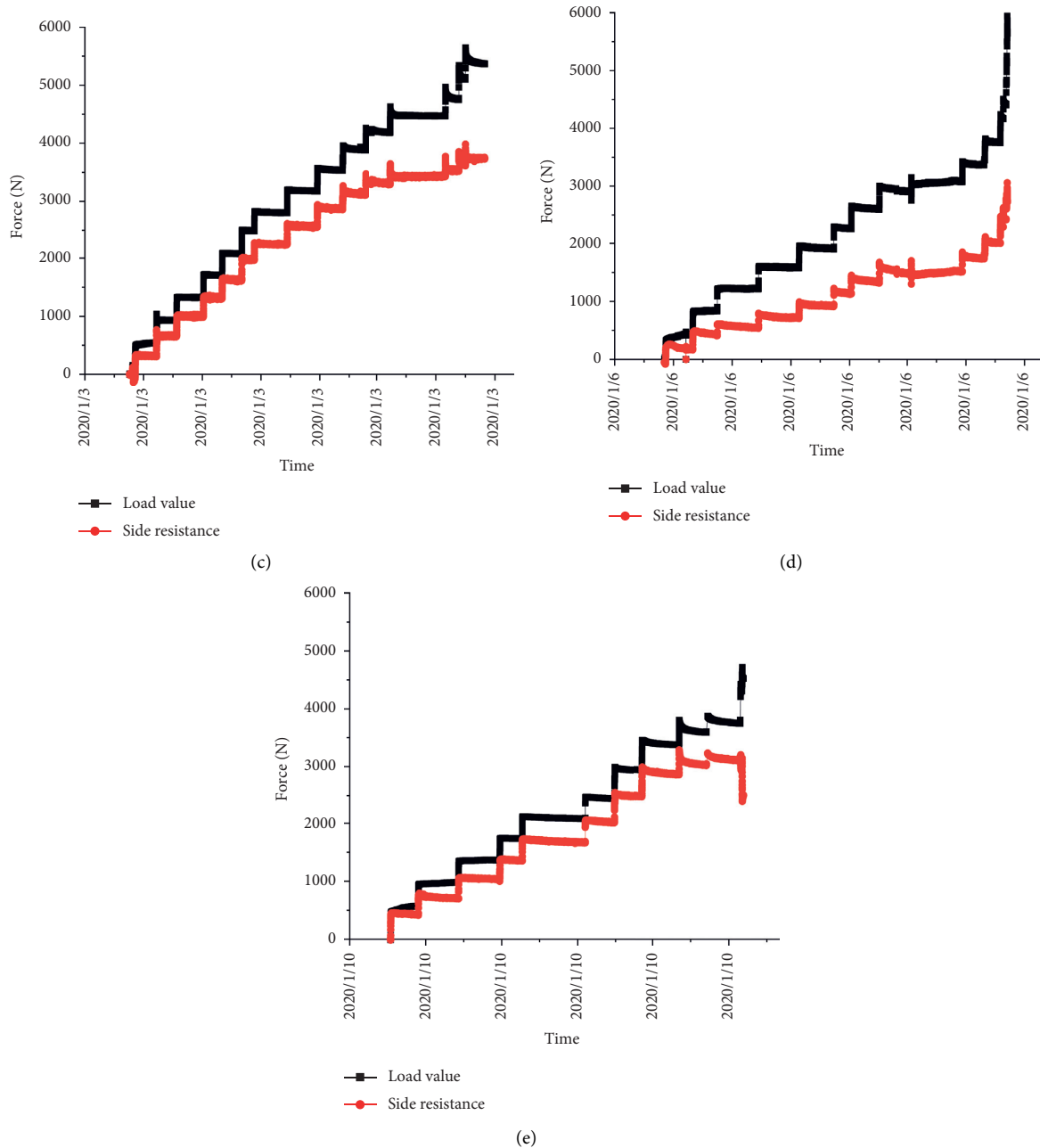


FIGURE 7: Change of load and side resistance of the vertically loaded single pile with time. (a) 0 mm. (b) 0.5 mm. (c) 1.0 mm. (d) 1.5 mm. (e) 2.0 mm.

criterion was used. The contact between the pile and soil and the changes of the material parameters of the mudcake with different thickness were not considered. The diameter and length of the pile were 2.5 m and 60m, respectively, and the calculation model was 50 m × 50 m × 90 m (Figure 10). According to the actual data of the project site, the numerical calculation parameters are as shown in Table 3.

3.2. Calculation Process and Analysis of Results

3.2.1. Calculation Process. The mudcake thickness of 0 mm, 10 mm, 20 mm, 30 mm, and 50 mm were calculated to study

the effect of mudcake on the bearing capacity of the vertically loaded single pile. Uniformly distributed load in vertical direction was applied on the pile top. The load was applied step by step with an increase of 3000 kN starting from 5000 kN. When the vertical load on the pile top reached 32000 kN finally, the tip resistance and displacement of the pile shaft were calculated (Figure 11).

3.2.2. Bearing Capacity of the Single Pile. After the pile tip resistance was calculated (Figure 11(a)), side resistance can be obtained by taking the difference between the load of the single pile and the pile tip resistance. Table 4 and Figure 12

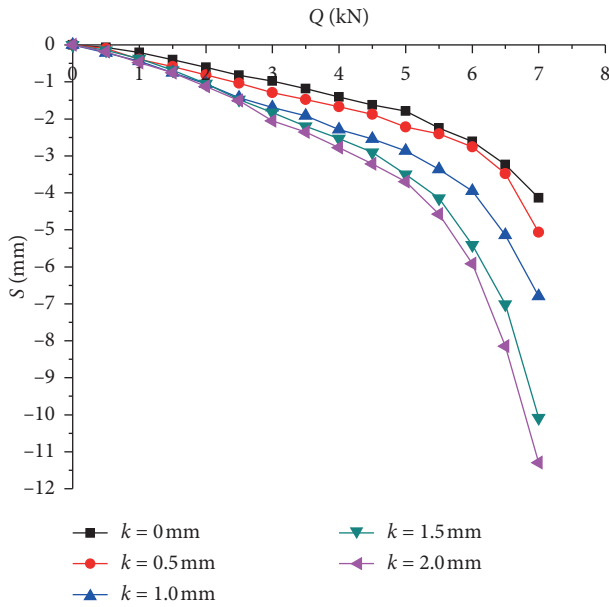


FIGURE 8: Load-settlement curves of pile with different mudcake thickness.

TABLE 2: Bearing capacity of the single pile with mudcake and its percentage accounting for that without mudcake.

Mudcake thickness (mm)		0	0.5	1.0	1.5	2.0
2.5 mm	Bearing capacity (kN)	5.85	5.63	4.42	3.95	3.63
	Percentage (%)	100	96	76	67	62
3.0 mm	Bearing capacity (kN)	6.31	6.17	5.13	4.48	4.25
	Percentage (%)	100	98	81	71	67
3.5 mm	Bearing capacity (kN)	6.64	6.51	5.62	5.02	4.79
	Percentage (%)	100	98	85	75	72

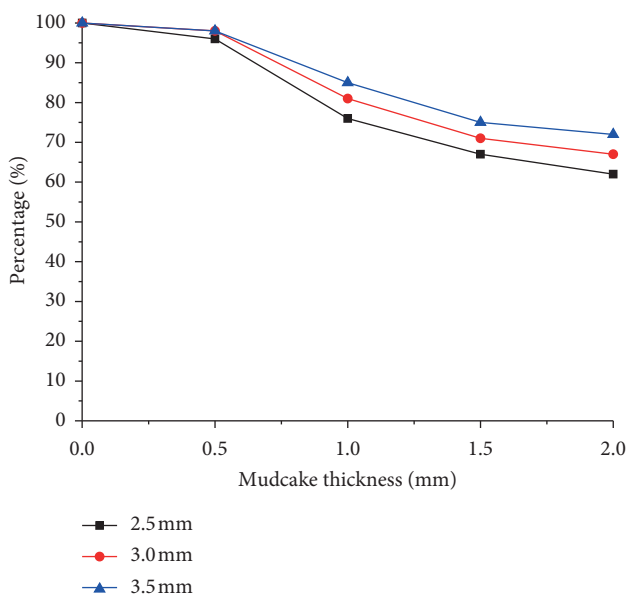


FIGURE 9: The percentages of single pile bearing capacity with mudcake accounting for that without mudcake

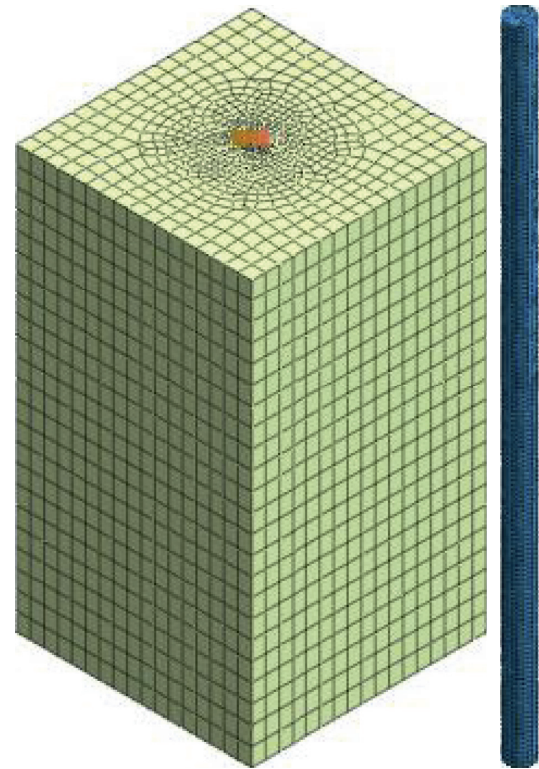


FIGURE 10: Calculation model and meshing.

show the side resistance and its percentage accounting for loading capacity under vertical loads.

It can be seen that, with the increase of mudcake thickness, the pile side resistance decreases first and, then, tends to be stable. When vertical loads of 2000 kN and 32000 kN were applied to the pile top, pile side resistance accounts for 92%–94% of the load, indicating that the bearing capacity of the single pile is mainly borne by pile side resistance. Under the pile top load of 50000 kN and 62000 kN, when the mudcake thickness is 10 mm, pile side resistance accounts for about 90% and 88% of the load, respectively. After the thickness of mudcake reaches 20 mm, the percentage of pile side resistance is within the range of 86%–88%. It shows that the existence of mudcake affects the normal development of side resistance of the single pile.

*3.2.3. Displacement of the Vertically Loaded Single Pile.* According to the results of displacement(Figure 11(b)), the numerical calculation results of the Q-S curve of the vertical load of the single pile and the vertical displacement of the pile top are as shown in Figure 13.

The bearing capacity of the single pile was determined according to the vertical displacement control standard of the pile top. When the control standard is 1%d, 2%d, and 3% d, respectively, the bearing capacity of the single pile with different mudcake thickness and its percentage accounting for that without mudcake are as shown in Table 5 and Figure 14.

It can be seen that, with the increase of mudcake thickness, the bearing capacity of the vertically loaded single

TABLE 3: Numerical calculation parameters.

Material	Unit weight (kN/m <sup>3</sup> )	Elasticity modulus (kPa)	Poisson ratio	Internal friction angle (°)	Cohesion (kPa)
Pile shaft	27.0	$2.9 \times 10^7$	0.25	-	-
Pile side soil	20.5	$3.7 \times 10^4$	0.30	30	0
Mudcake	15.0	$5.0 \times 10^3$	0.45	10	10

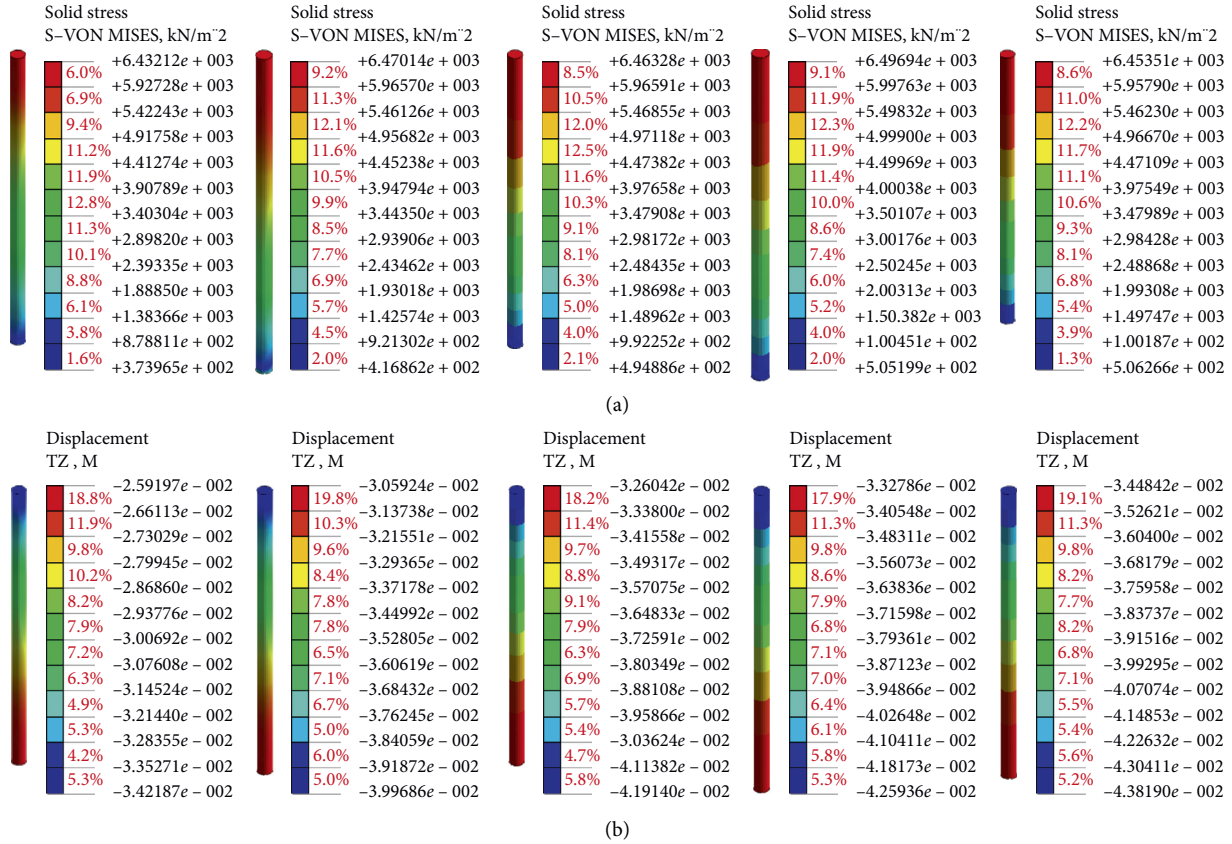


FIGURE 11: Stress and displacement with different mudcake thickness. (a) Stress. (b) Displacement (right-to-left: 0 mm, 10 mm, 20 mm, 30 mm, and 50 mm).

TABLE 4: Pile side resistance and its percentage accounting for loading capacity with different mudcake thickness.

	Loading (kN)	Thickness (mm)				
		0	10	20	30	50
20000	Side resistance (kN)	18785	18587	18411	18407	18424
	Percentage (%)	93.93	92.93	92.05	92.03	92.12
32000	Side resistance (kN)	30168	29957	29575	29525	29519
	Percentage (%)	94.27	93.62	92.42	92.26	92.25
50000	Side resistance (kN)	46309	45128	44056	43977	44171
	Percentage (%)	94.62	90.26	88.11	87.95	88.34
62000	Side resistance (kN)	58754	54730	53311	53267	53592
	Percentage (%)	94.76	88.27	85.99	85.91	86.44

pile decreases rapidly at first and, then, tends to be stable. The larger the vertical displacement control standard of the pile top is, the smaller the bearing capacity is when it tends to be stable. When the control standard is 1% d, 2% d, and 3% d, respectively, the bearing capacity of the single pile with 10 mm mudcake thickness accounts for about 96%, 85%, and

78% of the corresponding bearing capacity of the pile without mudcake. When the thickness of mudcake exceeds 20 mm, the bearing capacity of the single pile with mudcake is within the range of 88.5%–91.5%, 71.1%–72.7%, and 66.3%–66.2%, respectively.

#### 4. Mudcake Thickness Effect on the Bearing Capacity of the Vertically Loaded Single Pile

Mudcake is a layer of sticky clay between the pile and the soil on the side of the pile. The pile side mudcake is the product of a series of complex physical and chemical interactions of mud for wall protection, pile shaft, and pile side soil, and its engineering properties are often inferior to the pile side soil [1]. The existence of pile side mudcake prevents the direct combination between the pile and the pile side soil, affects the interaction between the pile and pile side soil, causes the increase of the vertical displacement of the pile, reduces the side friction resistance of the pile, and then, affects the normal performance of the bearing capacity of vertically

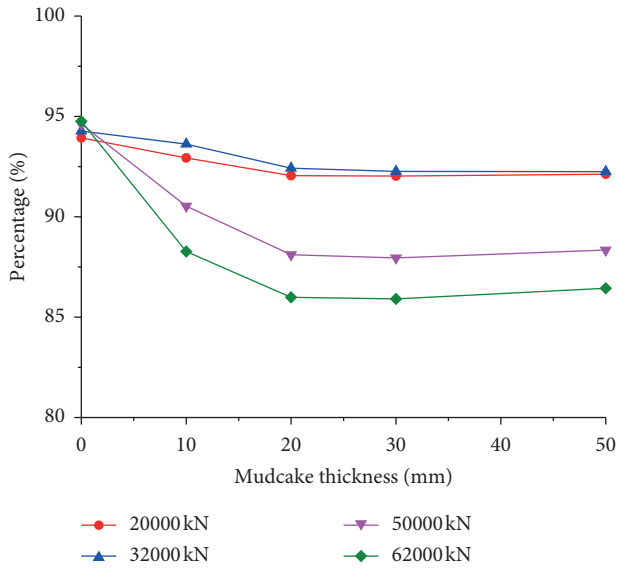


FIGURE 12: Variation of percentages of side resistance accounting for load with mudcake thickness.

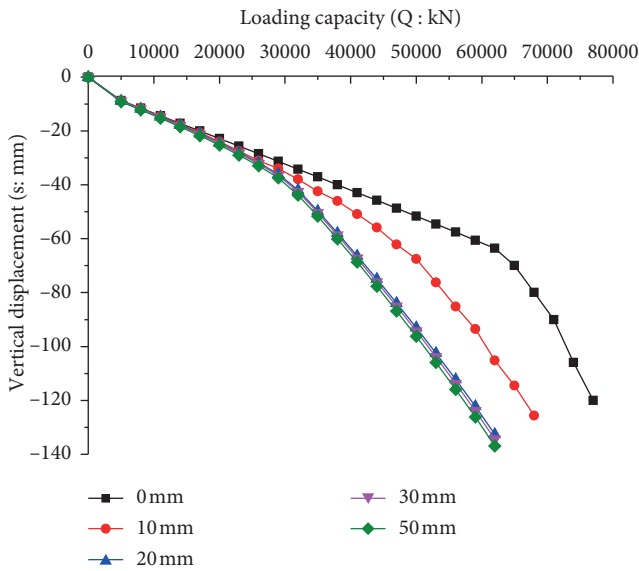


FIGURE 13: Q-S curve of the single vertically loaded pile.

loaded single piles. Effect refers to a specific scientific phenomenon under specific environmental conditions, which is produced by some specific motivation or reason. In this paper, the specific scientific phenomenon that the vertical load bearing performance of the single pile is affected by the existence of mudcake of different thickness between the pile and the pile side soil is defined as the mudcake thickness effect on the bearing performance of the vertically loaded single pile.

The mudcake thickness effect on the bearing capacity of the vertically loaded single pile often causes the reduction of the actual bearing capacity of the single pile, which poses a hazard to engineering construction. During the construction of mud protection bored piles, the mudcake thickness effect

TABLE 5: Bearing capacity of the single pile with different mudcake thickness and its percentage accounting for that without mudcake

Mudcake thickness (mm)	0	10	20	30	50	
1% d	Bearing capacity (kN)	22324	21486	20432	20226	19758
	Percentage (%)	100.00%	96.25%	91.52%	90.60%	88.51%
2% d	Bearing capacity (kN)	48298	40921	35110	34839	34365
	Percentage (%)	100.00%	84.73%	72.69%	72.13%	71.15%
3% d	Bearing capacity (kN)	66500	52018	44000	43644	42753
	Percentage (%)	100.00%	78.22%	66.17%	65.63%	64.29%

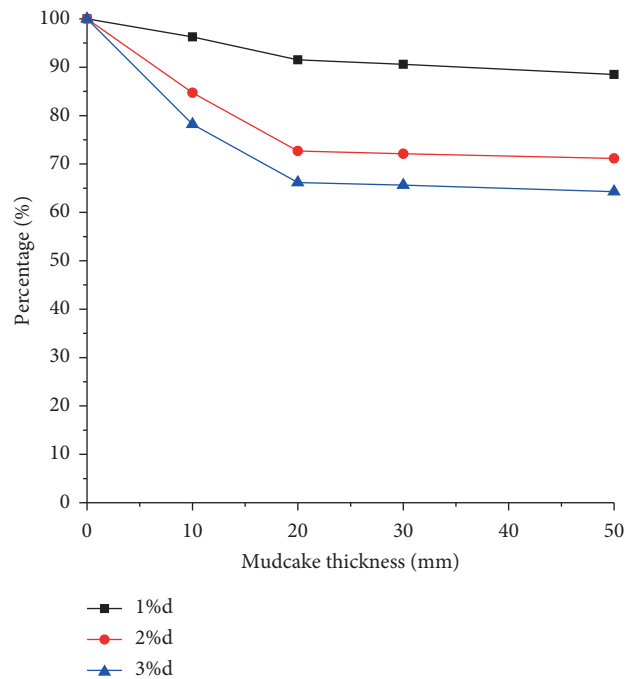


FIGURE 14: The percentage of single pile bearing capacity with mudcake accounting for that without mudcake

on the bearing capacity of the vertically loaded single pile should be fully considered. If necessary, technical measures such as improving mud properties and high-pressure grouting can be adopted to reduce the pile side mudcake thickness.

### 5. Conclusions

- (1) The change laws of the mudcake thickness effect on the bearing capacity of the vertically loaded single pile were obtained. With the increase of the mudcake thickness, the side resistance, vertical displacement, and bearing capacity of the vertically loaded single pile decrease gradually, firstly, and then, increase suddenly and become stable eventually.

- (2) The concept of the mudcake thickness effect on the bearing capacity of the vertically loaded single pile was put forward, and it was pointed out that the thickness effect of the mudcake on the bearing capacity of the vertically loaded single pile should be fully considered during the construction of mud protection bored piles.
- (3) The mechanism of the mudcake thickness effect on the bearing capacity of the vertically loaded single pile was revealed. Because pile side mudcake blocks the direct connection between the pile and the pile side soil, the interaction between pile and pile side soil is affected. As a result, the vertical displacement of the pile increases, pile side friction decreases, and the normal performance of vertically loaded bearing capacity of the single pile is affected greatly.

### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was funded by the Scientific Research Project of Anhui Jianzhu University (No. 2019QDZ24) and the Natural Science Foundation of Anhui Province (No. 2008085ME165).

### References

- [1] Z.-M. Zhang, J. Yu, G.-X. Zhang, and X.-M. Zhou, "Test study on the characteristics of mudcakes and in situ soils around bored piles," *Canadian Geotechnical Journal*, vol. 46, no. 3, pp. 241–255, 2009.
- [2] J.-J. Zhou, X.-N. Gong, K.-H. Wang, and R.-H. Zhang, "Shaft capacity of the pre-bored grouted planted pile in dense sand," *Acta Geotechnica*, vol. 13, no. 5, pp. 1227–1239, 2018.
- [3] Z. Ai-rong, Z.-J. Chen, and A.-J. Zhuge, "Pile load test of post-grouting bored pile at Beijing capital international airport," *IOP Conference Series: Earth and Environmental Science*, vol. 267, no. 5, 2019.
- [4] Hosoit, "Consideration to the skin friction of diaphragm wall foundation," in *Proceedings of the 3rd International Conference on Deep Foundation Practice Incorporating on Pile Talk*, pp. 213–221, Singapore, 1994.
- [5] R. E. Majano, M. W. O'Neill, and K. M. Hassan, "Perimeter load transfer in model drilled shafts formed under slurry," *Journal of Geotechnical Engineering*, vol. 120, no. 12, pp. 2136–2154, 1994.
- [6] Yu-jun Wu, "Effect of the mud thickness around cast-in-place piles upon the bearing capacity of single-pile," *Journal of Jilin Institute of Architecture & Civil Engineering*, vol. 27, no. 6, pp. 16–18, 2010.
- [7] Q. Hu, J. Mei, L.-P. He, Y.-H. Xu, and Y.-H. Zhang, "Correction of bearing capacity formula of pile foundation in red mudstone considering mudcake thickness," *Journal of Highway and Transportation Research and Development*, vol. 34, no. 7, pp. 60–68, 2017.
- [8] C. Chen, W.-M. Leng, Q. Yang, Z.-H. Jin, R.-S. Nie, and J. Qiu, "Experimental study of mechanical properties of concrete pile-slurry-sand interface," *Rock and Soil Mechanics*, vol. 39, no. 7, pp. 2461–2472, 2018.
- [9] Y.-Q. Chen, J.-S. Lei, L. Xu, G. Wang, and Z. Dai, "Model test study on influence of mud cake on friction performance of pouring pile," *Journal of Railway Science and Engineering*, vol. 16, no. 7, pp. 1660–1665, 2019.
- [10] J.-J. Zhou, K.-H. Wang, X.-N. Gong, and R.-H. Zhang, "Bearing capacity and load transfer mechanism of a static drill rooted nodular pile in soft soil areas," *Journal of Zhejiang University-Science A (Applied Physics & Engineering)*, vol. 14, pp. p705–719, 2013.
- [11] J.-J. Zhou, X.-N. Gong, K.-H. Wang, R.-H. Zhang, and T.-L. Yan, "A model test on the behavior of a static drill rooted nodular pile under compression," *Marine Georesources & Geotechnology*, vol. 34, no. 3, pp. p293–301, 2016.
- [12] M. Xub, P. Nia, X. Dingc, and G. Meib, "Physical and numerical modelling of axially loaded bored piles with debris at the pile tip," *Computers and Geotechnics*, vol. 114, Article ID 103146, 13 pages, 2019.
- [13] E. G. Balakrishnan, A. S. Balasubramaniam, and N. Phien-Wej, "Load deformation analysis of bored piles in residual weathered formation," *Journal of Geotechnical and Environmental Engineering*, vol. 125, no. 2, pp. 122–131, 1999.
- [14] M. Yang, "Study of reducing settlement pile foundation based on controlling settlement principle," *Chinese Journal of Geotechnical Engineering*, vol. 22, no. 4, pp. 481–486, 2000.
- [15] Z. Muszynski, J. Rybak, and P. Kaczor, "Accuracy assessment of semi-automatic measuring techniques applied to displacement control in self-balanced pile capacity testing appliance," *Sensors*, vol. 8, no. 4067, pp. p1–23, 2018.
- [16] P. Deb and S. K. Pal, "Numerical analysis of piled raft foundation under combined vertical and lateral loading," *Ocean Engineering*, vol. 190, no. 10643, pp. p1–13, 2019.
- [17] K. Chatterjee and D. Choudhury, "Analytical and numerical approaches to compute the influence of vertical load on lateral response of single pile," in *Proceedings of the 15th Asian Regional Conference on Soil Mechanics and Geotechnical Engineering: New Innovations and Sustainability*, vol. 1032, pp. p1319–1322, Fukuoka, Japan, November 2015.
- [18] Z. Liu, A. Zhou, Y. Hu, X. Sui, and Q. Liu, "Numerical simulation analysis of supporting stability of cast-in-place piles in pipe jacking," *Earth and Environmental Science*, vol. 267, no. 4, pp. p1–7, 2019.
- [19] X. Wang and Y. Ruan, "The comparative analysis research of the finite element between sand lining teathed pile and smooth pile," *Advanced Materials Research*, vol. 1032, pp. p905–907, 2014.

## Research Article

# Computation of Irregularity Indices of Certain Computer Networks

Jiangnan Liu,<sup>1</sup> Lulu Cai ,<sup>2</sup> Abaid ur Rehman Virk ,<sup>3</sup> Waheed Akhtar,<sup>4</sup> Shahzad Ahmed Maitla,<sup>4</sup> and Yang Wei<sup>5</sup>

<sup>1</sup>College of Electrical and Information Engineering, Quzhou University, Quzhou 324000 China and Southeast Digital Economic Development Institute, Quzhou, 324000, China

<sup>2</sup>College of Electrical and Information Engineering, Quzhou University, Quzhou 324000, China

<sup>3</sup>Department of Mathematics, University of Management and Technology, Lahore, Pakistan

<sup>4</sup>Department of Mathematics, University of Management and Technology Lahore, Sialkot Campus, Pakistan

<sup>5</sup>State GRID Quzhou Power Supply Company, Quzhou 324000, China

Correspondence should be addressed to Lulu Cai; [ysu-fbg@163.com](mailto:ysu-fbg@163.com) and Abaid ur Rehman Virk; [abaid.math@gmail.com](mailto:abaid.math@gmail.com)

Received 24 April 2020; Accepted 4 July 2020; Published 17 August 2020

Guest Editor: Muhammad Javaid

Copyright © 2020 Jiangnan Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A graph is said to be a regular graph if all its vertices have the same degree; otherwise, it is irregular. In general, irregularity indices are used for computational analysis of nonregular graph topological composition. The creation of irregular indices is based on the conversion of a structural graph into a total count describing the irregularity of the molecular design on the map. It is important to be notified how unusual a molecular structure is in various situations and problems in structural science and chemistry. In this paper, we will compute irregularity indices of certain networks.

## 1. Introduction

In mathematics, graph theory can be used to describe different types of graphs that are computational structures. It is also used to model sensible item connections [1, 2]. An irregularity index is a statistical value connected with a graph that defines a graph's irregularity. The theory of networks is a part of computer science and network engineering graph theory.

A topological invariant  $TOP(G)$  is referred to as a graph  $G$  irregularity index if the topological invariant  $TOP(G) \geq 0$  and  $TOP(G) = 0$  if  $G$  is a regular graph. However, if all its vertices have the same degree, it is said that the graph is regular. Topological invariant  $TOP(G)$  is a numeric value of a molecular structure of a chemical compound. Nonetheless, the creation of irregular indices is

based on the conversion of a structural graph into a total count describing the irregularity of the molecular design on the map [3, 4]. Many networks including silicate, chain silicate, oxide, hexagonal, and honeycomb networks are identical to networks of atomic or chemical structure. There are very important unusual characteristics in such networks.

In this paper, we are concerned with simple connected graphs symbolized by  $G(V, E)$ , where  $V(G)$  and  $E(G)$  represent the set of vertices and edges of  $G$ , respectively. The degree of a vertex  $u$  of a graph  $G$  is the count of first neighbors of  $u$ . And  $uv$  represents an edge for  $G$ , connecting vertices  $u$  and  $v$  [5, 6].

Graph theory was established in 1736 when Leonhard Euler presented "Solutio problematica as situspertinntis geometriis" (the solution of a problem related to place theory).



Wiener is the pioneer of topological indices; he discovered the first topological index and found out the boiling point of a compound (paraffin, a member of the alkane family) in 1947. It was named as path number, but latterly, it was renamed as the Wiener index [7].

$$W(G) = \frac{1}{2} \sum_{u,v \in V(G)} (d_u, d_v). \quad (1)$$

Wiener also invented the Wiener polarity index. Milan Randić invented the first and oldest degree-based index named as the Randić index in 1975 [8]:

$$R_{-1/2}(G) = \sum_{uv \in E(G)} \frac{1}{\sqrt{d_u d_v}}. \quad (2)$$

The graph invariant denoted by  $M_1(G)$  is called the first Zagreb index, which is equal to the sum of square of the degrees of the vertices of a graph; it was introduced by Trinjastic and Gutman in 1972 [9].  $M_1(G)$  is linked with sum of quantities in the field of chemical graph theory.  $M_1(G)$  is known as the Gutman index;  $M_1(G)$  is bounded and attains lower and upper bound [10]:

$$M_1(G) = \sum_{uv \in E(G)} (d_u + d_v). \quad (3)$$

The second Zagreb index is a graph invariant denoted by  $M_2(G)$  which is defined as the aggregate of the product of degrees of connected pairs of vertices of the molecular compound, and it was introduced by Trinjastic and Gutman in 1972.

$$M_2(G) = \sum_{uv \in E(G)} (d_u \cdot d_v). \quad (4)$$

Bo-Zhou and Ivan Gutman presented the upper bound for these Zagreb indices w.r.t the min-max degree [11, 12].

Nevertheless, the emergence of new topological descriptors with extremely detecting capacity retains a concern without deformation for the scientific world [13, 14]. Therefore, there is a great willingness to change novel graph invariants with enormous detecting power combined with insignificant degeneration. In this paper, we compute irregularity indices for certain networks.

## 2. Irregularity Indices

All these selected irregularity indices belong to the family of degree-based topological indices. Tamas Reti et al. selected these irregularity indices as a molecular descriptor in the QSPR study to predict physicochemical properties of octane isomers [15]. The selected irregularity indices for certain networks are represented by

$$VAR(G) = \frac{M_1(G)}{n} - \left(\frac{2m}{n}\right)^2,$$

$$AL(G) = \sum_{uv \in E} |d_u - d_v|,$$

$$IR1(G) = F(G) - \frac{2m}{n} M_1(G),$$

$$IR2(G) = \sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n},$$

$$IRF(G) = F(G) - 2M_2(G),$$

$$IRFW(G) = \frac{IRF(G)}{M_2(G)},$$

$$IRA(G) = \sum_{uv \in E} (d_u^{-1/2} - d_v^{-1/2})^2,$$

$$IRB(G) = \sum_{uv \in E} (d_u^{1/2} - d_v^{1/2})^2,$$

$$IRC(G) = \sum_{uv \in E} \frac{\sqrt{d_u d_v}}{m} - \frac{2m}{n} \quad (5)$$

$$IRDIF(G) = \sum_{uv \in E} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right|,$$

$$IRL(G) = \sum_{uv \in E} |\ln d_u - \ln d_v|,$$

$$IRLU(G) = \sum_{uv \in E} \frac{|d_u - d_v|}{\min(d_u, d_v)},$$

$$IRLF(G) = \sum_{uv \in E} \frac{|d_u - d_v|}{\sqrt{d_u d_v}},$$

$$IRLA(G) = 2 \sum_{uv \in E} \frac{|d_u - d_v|}{d_u + d_v},$$

$$IRD1(G) = \sum_{uv \in E} \ln\{1 + |d_u - d_v|\},$$

$$IRGA(G) = \sum_{uv \in E} \ln \frac{d_u + d_v}{(2)\sqrt{d_u d_v}}$$

## 3. Main Results and Discussion

In this section of the paper, we will discuss about the silicate network, chain silicate network, oxide network, hexagonal network, and honeycomb network briefly and will compute the irregularity indices for these networks.

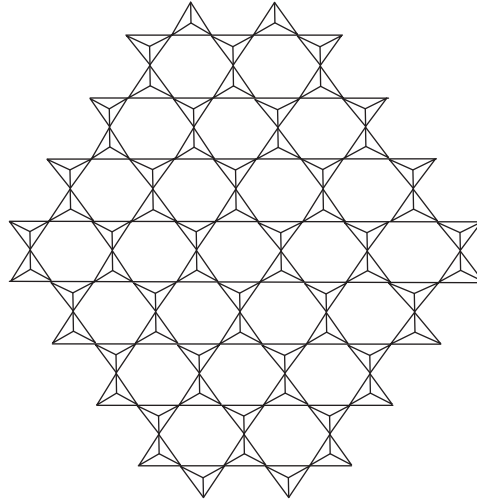


FIGURE 1: Structure of the silicate network.

TABLE 1: Partition of the edge set of  $SL_n$  which is based on degrees of end vertices.

$(d_u, d_v)$	(3, 3)	(3, 6)	(6, 6)
Number of edges	$6n$	$18n^2 + 6n$	$18n^2 - 12n$

3.1. *Silicate Network.* Silicates are the most popular, significant, and most complex mineral class. Tetrahedron ( $SiO_4$ ) is the basic chemical unit of silicates. By fusing metal oxides or metal carbonates with sand, silicates are produced. The fact is that all silicates contain tetrahedron of  $SiO_4$ . For chemistry, the corner vertices of tetrahedron  $SiO_4$  reflect oxygen ions, and the silicon ion is represented by the middle vertex. We name the corner vertices as oxygen nodes in graph theory and the middle vertex as the silicon node. Although the tetrahedron is arranged linearly, chain silicates are produced (refer Figure 1).

The total number of vertices and edges in  $SL_n$  is  $15n^2 + 3n$  and  $36n^2$ , respectively.

**Theorem 1.** *The irregularity indices for the silicate network ( $SL_n$ ) for  $n > 1$  are*

$$VAR(SL_n) = \frac{54n^2 + 16n - 22}{(5n + 1)^2},$$

$$AL(SL_n) = 3n(36n + 6),$$

$$IR1(SL_n) = \frac{9072n^3 - 3690n^2 + 486n}{5n + 1},$$

$$IR2(SL_n) = \sqrt{\frac{(324n - 90)}{12n} - \frac{24n}{5n + 1}},$$

$$IRF(SL_n) = 3n(54n + 18),$$

$$IRFW(SL_n) = \frac{54n + 18}{324n - 90},$$

$$IRA(SL_n) = 0.1716n(3n + 1),$$

$$IRB(SL_n) = 3.0883n(3n + 1),$$

$$IRC(SL_n) = \frac{(30.7279n - 3.7279)}{6n} - \frac{24n}{5n + 1},$$

$$IRDIF(SL_n) = 27n^2 + 9n,$$

$$IRL(SL_n) = 0.6931(18n^2 + 6n),$$

$$IRLU(SL_n) = 6n(3n + 1),$$

$$IRLF(SL_n) = 4.2426n(3n + 1),$$

$$IRLA(SL_n) = 4n(3n + 1),$$

$$IRD1(SL_n) = 8.3178n(3n + 1),$$

$$IRGA(SL_n) = 0.3533n(3n + 1).$$

(6)

*Proof.* By using the edge partition based on degrees of end vertices of each edge of the silicate network ( $SL_n$ ) given in Table 1, we compute the irregularity indices of the silicate network ( $SL_n$ ), and the computations are given as follows:

$$VAR(SL_n) = \frac{M_1(SL_n)}{n} - \left(\frac{2m}{n}\right)^2 = \frac{3n(126n - 22)}{3n(5n + 1)} - \left(\frac{2(36n^2)}{3n(5n + 1)}\right)^2 = \frac{54n^2 + 16n - 22}{(5n + 1)^2},$$

$$AL(SL_n) = \sum_{uv \in E} |d_u - d_v| = |3 - 3|(6n) + |3 - 6|(18^2 + 6n) + |6 - 6|(18n^2 - 12n) = 3n(36n + 6),$$

$$IR1(SL_n) = F(SL_n) - \frac{2m}{n}M_1(SL_n) = (2106n^2 - 486n) - \frac{(2)(36n^2)(378n^2 - 66n)}{3n(5n + 1)} = \frac{9072n^3 - 3690n^2 + 486n}{5n + 1},$$

$$IR2(SL_n) = \sqrt{\frac{M_2(SL_n)}{m}} - \frac{2m}{n} = \sqrt{\frac{3n(324n - 90)}{36n^2}} - \frac{2(36n^2)}{3n(5n + 1)} = \sqrt{\frac{(324n - 90)}{12n}} - \frac{24n}{5n + 1},$$

$$IRF(SL_n) = F(SL_n) - 2M_2(SL_n) = (2106n^2 - 486n) - 2(972n^2 - 270n) = 3n(54n + 18),$$

$$IRFW(SL_n) = \frac{IRF(SL_n)}{M_2(SL_n)} = \frac{3n(54n + 18)}{3n(324n - 90)} = \frac{54n + 18}{324n - 90},$$

$$\begin{aligned} IRA(SL_n) &= \sum_{uv \in E} (d_u^{-1/2} - d_v^{-1/2})^2 = \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}}\right)^2 (6n) + \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{6}}\right)^2 (18n^2 + 6n) \\ &\quad + \left(\frac{1}{\sqrt{6}} - \frac{1}{\sqrt{6}}\right)^2 (18^2 - 12n) = 0.1716n(3n + 1), \end{aligned} \tag{7}$$

$$\begin{aligned} IRB(SL_n) &= \sum_{uv \in E} (d_u^{1/2} - d_v^{1/2})^2 = (\sqrt{3} - \sqrt{3})^2 (6n) + (\sqrt{3} - \sqrt{6})^2 (18n^2 + 6n) \\ &\quad + (\sqrt{6} - \sqrt{6})^2 (18n^2 - 6n) = 3.0883n(3n + 1), \end{aligned}$$

$$\begin{aligned} IRC(SL_n) &= \sum_{uv \in E} \frac{\sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{\sqrt{9}(6n) + (18n^2 + 6n)\sqrt{18} + (18n^2 - 12n)\sqrt{36}}{36n^2} \\ &\quad - \frac{72n^2}{3n(5n + 1)} = \frac{(30.7279n - 3.7279)}{6n} - \frac{24n}{5n + 1}, \end{aligned}$$

$$IRDIF(SL_n) = \sum_{uv \in E} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| = (6n) \left| \frac{3}{3} - \frac{3}{3} \right| + (18n^2 + 6n) \left| \frac{3}{6} - \frac{6}{3} \right| + (18n^2 - 12n) \left| \frac{6}{6} - \frac{6}{6} \right| = 27n^2 + 9n,$$

$$\begin{aligned} IRL(SL_n) &= \sum_{uv \in E} |\ln d_u - \ln d_v| = (6n)|\ln(3) - \ln(3)| + (18n^2 + 6n)|\ln(3) - \ln(6)| \\ &\quad + (18n^2 - 12n)|\ln(6) - \ln(6)| = 0.6931(18n^2 + 6n), \end{aligned}$$

$$IRLU(SL_n) = \sum_{uv \in E} \frac{|d_u - d_v|}{\min(d_u, d_v)} = (6n) \frac{|3 - 3|}{3} + (6n(3n + 1)) \frac{|3 - 6|}{3} + (18n^2 - 12n) \frac{|6 - 6|}{6} = 6n(3n + 1).$$

Specific values of irregularity indices of  $SL_n$  for different values of parameters are given in Table 2.

$$\begin{aligned}
 IRLF(SL_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\sqrt{d_u d_v}} = (6n) \frac{|3-3|}{\sqrt{9}} + (18n^2 + 6n) \frac{|6-3|}{\sqrt{18}} + (18n^2 - 12n) \frac{|6-6|}{\sqrt{36}} = 4.2426n(3n+1), \\
 IRLA(SL_n) &= 2 \sum_{uv \in E} \frac{|d_u - d_v|}{d_u + d_v} = 2(6n) \frac{|3-3|}{6} + 2(18n^2 + 6n) \frac{|6-3|}{9} + 2(18n^2 - 12n) \frac{|6-6|}{12} = 4n(3n+1), \\
 IRD1(SL_n) &= \sum_{uv \in E} \ln\{1 + |d_u - d_v|\} = \ln\{1 + |3-3|\} (6n) + \ln\{1 + |3-6|\} (18n^2 + 6n) \\
 &\quad + \ln\{1 + |6-6|\} (18n^2 - 12n) = 8.3178n(3n+1), \\
 IRGA(SL_n) &= \sum_{uv \in E} \ln \frac{d_u + d_v}{(2)\sqrt{d_u d_v}} = (6n) \ln \frac{3+3}{(2)\sqrt{9}} + (18n^2 + 6n) \ln \frac{3+6}{(2)\sqrt{18}} + (18n^2 - 12n) \ln \frac{6+6}{(2)\sqrt{36}} = 0.3533n(3n+1).
 \end{aligned} \tag{8}$$

3.2. Chain Silicate Network. Chain silicate network is obtained when tetrahedra are organized in a sequence. An  $n$ -dimensional chain silicate network is represented by  $CS_n$ , and it is generated by sequentially organizing  $n$  tetrahedra. An  $n$ -dimensional chain silicate network is shown in Figure 2.

The edge partition of the chain silicate network is given in Table 3.

The total number of vertices and edges in  $CS_n$  is  $3n + 1$  and  $6n$ , respectively.

**Theorem 2.** The irregularity indices for the chain silicate network ( $CS_n$ ) for  $n > 1$  are

$$\begin{aligned}
 VAR(CS_n) &= \frac{18n^2 - 18}{(3n+1)^2}, \\
 AL(CS_n) &= 12n - 6, \\
 IR1(CS_n) &= \frac{162(n^2 - 1)}{3n+1}, \\
 IR2(CS_n) &= \sqrt{\frac{9(13n-8)}{6n}} - \frac{12n}{3n+1}, \\
 IRF(CS_n) &= 36n - 18, \\
 IRFW(CS_n) &= \frac{4n-2}{13n-8}, \\
 IRA(CS_n) &= 0.02857(4n-2),
 \end{aligned}$$

$$\begin{aligned}
 IRB(CS_n) &= 0.5147(4n-2), \\
 IRC(CS_n) &= \frac{3(n+4) + 3\sqrt{2}(4n-2) + 6(n-2)}{6n} \\
 &\quad - \frac{12n}{3n+1}, \\
 IRDIF(CS_n) &= 6n - 3, \\
 IRL(CS_n) &= 0.6931(4n-2), \\
 IRLU(CS_n) &= 4n - 2, \\
 IRLF(CS_n) &= \sqrt{2}(2n-1), \\
 IRLA(CS_n) &= \frac{4}{3}(2n-1), \\
 IRD1(CS_n) &= (4n-2)\ln 4, \\
 IRGA(CS_n) &= (2n-1)\ln \frac{3}{\sqrt{2}}.
 \end{aligned} \tag{9}$$

*Proof.* By using the edge partition based on degrees of end vertices of each edge of the chain silicate network ( $CS_n$ ) given in Table 3, the computations for irregularities indices are given as follows:

$$VAR(CS_n) = \frac{M_1(CS_n)}{n} - \left(\frac{2m}{n}\right)^2 = \frac{18(3n-1)}{3n+1} - \left(\frac{2(6n)}{3n+1}\right)^2 = \frac{18n^2 - 18}{(3n+1)^2},$$

$$AL(CS_n) = \sum_{uv \in E} |d_u - d_v| = |3-3|(n+4) + |3-6|(4n-2) + |6-6|(n-2) = 12n - 6,$$

$$IR1(CS_n) = F(CS_n) - \frac{2m}{n}M_1(CS_n) = (270n - 162) - \frac{(6n)(2)(54n-18)}{3n+1} = \frac{162n^2 - 162}{3n+1},$$

$$IR2(CS_n) = \sqrt{\frac{M_2(CS_n)}{m}} - \frac{2m}{n} = \sqrt{\frac{9(13n-8)}{6n}} - \frac{2(6n)}{3n+1} = \sqrt{\frac{9(13n-8)}{6n}} - \frac{12n}{3n+1},$$

$$IRF(CS_n) = F(CS_n) - 2M_2(CS_n) = (270n - 162) - 18(13n - 8) = 36n - 18,$$

$$IRFW(CS_n) = \frac{IRF(CS_n)}{M_2(CS_n)} = \frac{18(2n-1)}{9(13n-8)} = \frac{4n-2}{13n-8},$$

$$IRA(CS_n) = \sum_{uv \in E} (d_u^{-1/2} - d_v^{-1/2})^2 = \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}}\right)^2 (n+4) + \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{6}}\right)^2 (4n-1) \\ + \left(\frac{1}{\sqrt{6}} - \frac{1}{\sqrt{6}}\right)^2 (n-2) = 0.02857(4n-2),$$

$$IRB(CS_n) = \sum_{uv \in E} (d_u^{1/2} - d_v^{1/2})^2 = (\sqrt{3} - \sqrt{3})^2 (n+4) + (\sqrt{3} - \sqrt{6})^2 (4n-2) + (\sqrt{6} - \sqrt{6})^2 (n-2) = 0.5147(4n-2),$$

$$IRC(CS_n) = \sum_{uv \in E} \frac{\sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{\sqrt{9}(n+4) + (4n-2)\sqrt{18} + (n-2)\sqrt{36}}{6n} - \frac{12n}{3n+1} \\ = \frac{3(n+4) + 3\sqrt{2}(4n-2) + 6(n-2)}{6n} - \frac{12n}{3n+1},$$

$$IRDI F(CS_n) = \sum_{uv \in E} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| = (n+4) \left| \frac{3}{3} - \frac{3}{3} \right| + (4n-2) \left| \frac{3}{6} - \frac{6}{3} \right| + (n-2) \left| \frac{6}{6} - \frac{6}{6} \right| = 6n - 3,$$

$$IRL(CS_n) = \sum_{uv \in E} |\ln d_u - \ln d_v| = (n+4)|\ln(3) - \ln(3)| + (4n-2)|\ln(3) - \ln(6)| + (n-2)|\ln(6) - \ln(6)| = 0.6931(4n-2),$$

$$IRLU(CS_n) = \sum_{uv \in E} \frac{|d_u - d_v|}{\min(d_u, d_v)} = (n+4) \frac{|3-3|}{3} + (4n-2) \frac{|3-6|}{3} + (n-2) \frac{|6-6|}{6} = 4n-2,$$

$$IRLF(CS_n) = \sum_{uv \in E} \frac{|d_u - d_v|}{\sqrt{d_u d_v}} = (n+4) \frac{|3-3|}{\sqrt{9}} + (4n-2) \frac{|6-3|}{\sqrt{18}} + (n-2) \frac{|6-6|}{\sqrt{36}} = \sqrt{2}(2n-1),$$

$$IRLA(CS_n) = 2 \sum_{uv \in E} \frac{|d_u - d_v|}{d_u + d_v} = 2(n+4) \frac{|3-3|}{6} + 2(4n-2) \frac{|6-3|}{9} + 2(n-2) \frac{|6-6|}{12} = (2n-1) \frac{4}{3},$$

$$IRD1(CS_n) = \sum_{uv \in E} \ln\{1 + |d_u - d_v|\} = \ln\{1 + |3-3|\}(n+4) + \ln\{1 + |3-6|\}(4n-2) + \ln\{1 + |6-6|\}(n-2) = (4n-2)\ln 4,$$

$$IRGA(CS_n) = \sum_{uv \in E} \ln \frac{d_u + d_v}{(2)\sqrt{d_u d_v}} = (n+4)\ln \frac{3+3}{(2)\sqrt{9}} + (4n-2)\ln \frac{3+6}{(2)\sqrt{18}} + (n-2)\ln \frac{6+6}{(2)\sqrt{36}} = (2n-1)\ln \frac{3}{\sqrt{2}}$$

TABLE 2: Values of irregularity indices for  $SL_n$ .

Irregularity indices	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
$VAR(SL_n)$	1.3333	1.8678	2.0000	2.0544	2.0828
$AL(SL_n)$	126	468	1026	1800	2790
$IR1(SL_n)$	978	5344.3636	13324.5000	24929.1429	40160.7692
$IR2(SL_n)$	0.4159	0.4582	0.4497	0.4411	0.4344
$IRF(SL_n)$	216	756	1620	2808	4320
$IRFW(SL_n)$	0.3077	0.2258	0.2041	0.1940	0.1882
$IRA(SL_n)$	0.6864	2.4024	5.1480	8.9232	13.7280
$IRB(SL_n)$	12.3532	43.2362	92.6490	160.5916	247.6640
$IRC(SL_n)$	0.5000	0.4470	0.4142	0.3945	0.3816
$IRDIF(SL_n)$	36	126	270	468	720
$IRL(SL_n)$	16.6344	58.2204	124.7580	216.2472	332.6880
$IRLU(SL_n)$	24	84	180	312	480
$IRLF(SL_n)$	16.9704	59.3964	127.2780	220.6152	339.4080
$IRLA(SL_n)$	16	56	120	208	320
$IRD1(SL_n)$	33.2712	116.4492	249.5340	432.5256	665.4240
$IRGA(SL_n)$	1.4132	4.9462	10.5990	18.3716	28.2640

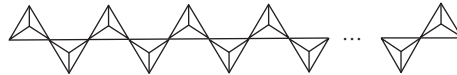


FIGURE 2: Structure of the chain silicate network.

TABLE 3: Edge partition of  $CS_n$ .

$(d_u, d_v)$ , where $uv \in E(G)$	(3, 3)	(3, 6)	(6, 6)
Number of edges	$n + 4$	$2(2n - 1)$	$n - 2$

Specific values of irregularity indices for different values of involved parameters for the chain silicon network are given in Table 4.

3.3. Oxide Network. Oxide networks play a crucial role in the analysis of silicate networks. If we detach silicone vertices out of a silicate network, we get an oxide network  $OX_n$  that is referred to as an  $n$ -dimensional oxide network as shown in Figure 3.

There are two types of edge partition in the oxide network ( $OX_n$ ) centered on degrees of end vertices. Table 5 shows the edge partition for  $OX_n$ .

The total number of vertices and edges in  $OX_n$  is  $9n^2 + 3n$  and  $18n^2$ , respectively.

**Theorem 3.** The irregularity indices of oxide networks ( $OX_n$ ) for  $n > 1$  are

$$VAR(OX_n) = \frac{72n^2(3n-1)}{(9n^2+3n)^2},$$

$$AL(OX_n) = 72n,$$

$$IR1(OX_n) = \frac{144n(3n-1)}{3n+1},$$

$$IR2(OX_n) = \sqrt{\frac{16(3n-1)}{3n}} - \frac{12n}{3n+1},$$

$$IRF(OX_n) = 48n,$$

$$IRFW(OX_n) = \frac{7}{24(3n-1)},$$

$$IRA(OX_n) = 0.5136n,$$

$$IRB(OX_n) = 4.1177n,$$

$$IRC(OX_n) = \frac{2(2\sqrt{2}+6n-4)}{3n} - \frac{12n}{3n+1},$$

$$IRDIF(OX_n) = 18n, \tag{11}$$

$$IRL(OX_n) = 8.3177n,$$

$$IRLU(OX_n) = 12n,$$

$$IRLF(OX_n) = \frac{12n}{\sqrt{2}},$$

$$IRLA(OX_n) = 8n,$$

$$IRD1(OX_n) = 13.1833n,$$

$$IRGA(OX_n) = 0.7068n.$$

*Proof.* By using the edge partition based on degrees of end vertices of each edge of the oxide network ( $OX_n$ ) given in Table 5, we have the following computations for irregularities of the oxide network ( $OX_n$ ):



$$\begin{aligned}
VAR(OX_n) &= \frac{M_1(OX_n)}{n} - \left(\frac{2m}{n}\right)^2 = \frac{24n(6n-1)}{3n(3n+1)} - \left(\frac{2 \cdot 18n^2}{3n(3n+1)}\right)^2 = \frac{72n^2(3n-1)}{(9n^2+3n)^2}, \\
AL(OX_n) &= \sum_{uv \in E} |d_u - d_v| = |2-4|12n + |4-4|(18n^2-12n) = 72n, \\
IR1(OX_n) &= F(OX_n) - \frac{2m}{n}M_1(OX_n) = (576n^2 - 144n) - \frac{(18n^2)(2)(24n(6n-1))}{9n^2+3n} = \frac{144n(3n-1)}{3n+1}, \\
IR2(OX_n) &= \sqrt{\frac{M_2(OX_n)}{m}} - \frac{2m}{n} = \sqrt{\frac{96n(3n-1)}{18n^2}} - \frac{2(18n^2)}{3n(3n+1)} = \sqrt{\frac{16(3n-1)}{3n}} - \frac{12n}{3n+1}, \\
IRF(OX_n) &= F(OX_n) - 2M_2(OX_n) = (576n^2 - 144n) - 2(288n^2 - 96n) = 48n, \\
IRFW(OX_n) &= \frac{IRF(OX_n)}{M_2(OX_n)} = \frac{28n}{96n(3n-1)} = \frac{7}{24(3n-1)}, \\
IRA(OX_n) &= \sum_{uv \in E} (d_u^{-1/2} - d_v^{-1/2})^2 = \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{4}}\right)^2 (12n) + \left(\frac{1}{\sqrt{4}} - \frac{1}{\sqrt{4}}\right)^2 (18n^2 - 12n) = 0.5136n, \\
IRB(OX_n) &= \sum_{uv \in E} (d_u^{1/2} - d_v^{1/2})^2 = (\sqrt{2} - \sqrt{4})^2 (12n) + (\sqrt{4} - \sqrt{4})^2 (18n^2 - 12n) = 4.1177n, \\
IRC(OX_n) &= \sum_{uv \in E} \frac{\sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{2(2\sqrt{2} + 6n - 4)}{3n} - \frac{12n}{3n+1}, \\
IRDIF(OX_n) &= \sum_{uv \in E} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| = 12n \left| \frac{2}{4} - \frac{4}{2} \right| + 18n^2 \left| \frac{4}{4} - \frac{4}{4} \right| = 18n.
\end{aligned} \tag{12}$$

Specific values of irregularity indices for different values of involved parameters for the oxide network are given in Table 6.

$$\begin{aligned}
IRL(OX_n) &= \sum_{uv \in E} |\ln d_u - \ln d_v| = 12n |\ln 2 - \ln 4| + 18n^2 |\ln 4 - \ln 4| = 8.3177n, \\
IRLU(OX_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\min(d_u, d_v)} = 12n \frac{|2-4|}{2} + 18n^2 \frac{|4-4|}{4} = 12n, \\
IRLF(OX_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\sqrt{d_u d_v}} = 12n \frac{|4-2|}{\sqrt{8}} + 18n^2 \frac{|4-4|}{\sqrt{16}} = \frac{12n}{\sqrt{2}}, \\
IRLA(OX_n) &= 2 \sum_{uv \in E} \frac{|d_u - d_v|}{d_u + d_v} = 2(12n) \frac{|4-2|}{6} + 2(18n^2) \frac{|4-4|}{8} = 8n, \\
IRD1(OX_n) &= \sum_{uv \in E} \ln\{1 + |d_u - d_v|\} = \ln\{1 + |4-2|\}(12n) + \ln\{1 + |4-4|\}(18n^2) = 13.1833n, \\
IRGA(OX_n) &= \sum_{uv \in E} \ln \frac{d_u + d_v}{2\sqrt{d_u d_v}} = (12n) \ln \frac{4+2}{2\sqrt{8}} + (18n^2) \ln \frac{4+4}{2\sqrt{16}} = 0.7068n.
\end{aligned} \tag{13}$$

TABLE 4: Values of irregularity indices for  $CS_n$ .

Irregularity indices	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
$VAR(CS_n)$	0.0000	1.1020	1.4400	1.5976	1.6875
$AL(CS_n)$	6	18	30	42	54
$IR1(CS_n)$	0.0000	69.4286	129.6000	186.9231	243.0000
$IR2(CS_n)$	-0.2614	0.2457	0.3337	0.3669	0.3852
$IRF(CS_n)$	18	54	90	126	162
$IRFW(CS_n)$	0.4000	0.3333	0.3226	0.3182	0.3158
$IRA(CS_n)$	0.0572	0.1716	0.2860	0.4003	0.5147
$IRB(CS_n)$	1.0294	3.0882	5.1470	7.2058	9.2646
$IRC(CS_n)$	-0.0858	0.1927	0.2570	0.2826	0.2956
$IRDIF(CS_n)$	3	9	15	21	27
$IRL(CS_n)$	1.3862	4.1586	6.9310	9.7034	12.4758
$IRLU(CS_n)$	2	6	10	14	18
$IRLF(CS_n)$	1.4142	4.2426	7.0711	9.8995	12.7279
$IRLA(CS_n)$	1.3333	4.0000	6.6667	9.3333	12.0000
$IRD1(CS_n)$	2.7726	8.3178	13.8629	19.4081	24.9533
$IRGA(CS_n)$	0.7520	2.2561	3.7602	5.2643	6.7683

TABLE 6: Values of irregularity indices for  $OX_n$ .

Irregularity indices	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
$VAR(OX_n)$	1.0000	0.8163	0.6400	0.5207	0.4375
$AL(OX_n)$	72	144	216	288	360
$IR1(OX_n)$	72	205.7143	345.6000	487.3846	630.0000
$IR2(OX_n)$	0.2660	3.8744	7.7137	11.6265	15.5718
$IRF(OX_n)$	48	96	144	192	240
$IRFW(OX_n)$	0.1458	0.0583	0.0365	0.0265	0.0208
$IRA(OX_n)$	0.5136	1.0272	1.5408	2.0544	2.5680
$IRB(OX_n)$	4.1177	8.2354	12.3531	16.4708	20.5885
$IRC(OX_n)$	0.2189	0.1809	0.1396	0.1124	0.0938
$IRDIF(OX_n)$	18	36	54	72	90
$IRL(OX_n)$	8.3177	16.6354	24.9531	33.2708	41.5885
$IRLU(OX_n)$	12	24	36	48	60
$IRLF(OX_n)$	8.4853	16.9706	25.4558	33.9411	42.4264
$IRLA(OX_n)$	8	16	24	32	40
$IRD1(OX_n)$	13.1833	26.3666	39.5499	52.7332	65.9165
$IRGA(OX_n)$	0.7068	1.4136	2.1204	2.8272	3.5340

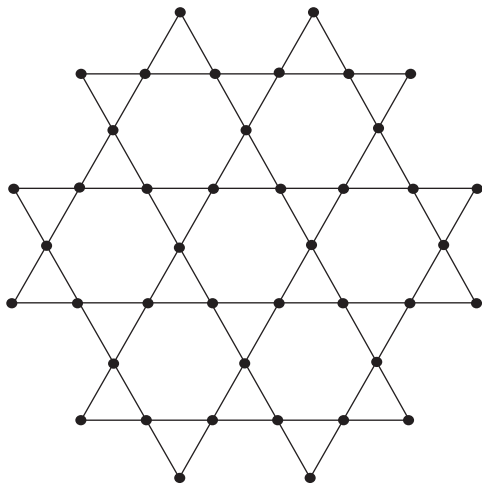


FIGURE 3: Structure of the oxide network.

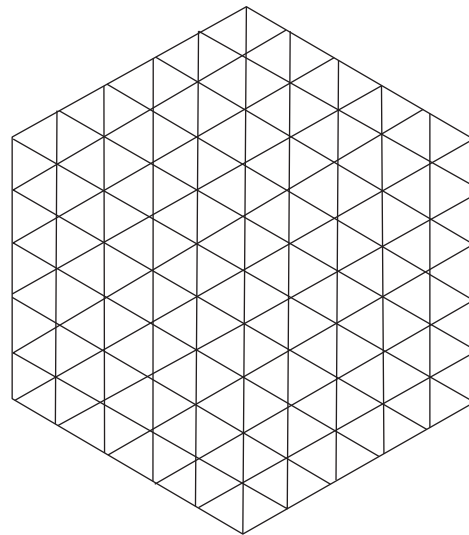


FIGURE 4: Structure of the hexagonal network.

TABLE 5: Edge partition for the oxide network.

$(d_u, d_v)$ , where $uv \in (G)$	(2, 4)	(4, 4)
Number of edges	$12n$	$18n^2 - 12n$

3.4. Hexagonal Network. It is well known that there are three normal plane tilings with the same regular polygon type as triangle, hexagon-shaped, and square structures. Triangular tiling is used in the design of hexagonal networks. An  $n$ -dimensional hexagonal network is commonly referred to as  $HX_n$ , where  $n$  is each side's number of hexagonal vertices. A hexagonal network  $HX_n$  is shown in Figure 4.

TABLE 7: Edge partition for the hexagonal network.

$(d_u, d_v)$ , where $uv \in (G)$	(3, 4)	(3, 6)	(4, 4)	(4, 6)	(6, 6)
Number of edges	12	6	$6(n - 3)$	$12(n - 2)$	$9^2 - 33n + 30$

Table 7 shows the edge partition of the hexagonal network ( $HX_n$ ). The total number of vertices and edges in hexagonal networks ( $HX_n$ ) is  $3n^2 - 3n + 1$  and  $9n^2 - 15n + 6$ , respectively.

**Theorem 4.** *The irregularity indices of the hexagonal network  $(HX_n)$  for  $n > 1$  are*

$$VAR(HX_n) = \frac{72n^3 + 243n^2 + 15n - 30}{(3n^2 - 3n + 1)^2},$$

$$AL(HX_n) = 6(4n - 3),$$

$$IR1(HX_n) = \frac{720n^3 - 2142n^2 + 1878n - 462}{3n^2 - 3n + 1},$$

$$IR2(HX_n) = \sqrt{\frac{108n^2 - 268n + 156}{3n^2 - 5n + 2}} - \frac{18n^2 - 30n + 12}{3n^2 - 3n + 1},$$

$$IRF(HX_n) = 6(8n - 5),$$

$$IRFW(HX_n) = \frac{(2)(8n - 5)}{(108n^2 - 268n + 156)},$$

$$IRA(HX_n) = 0.1008n - 0.0418,$$

$$IRB(HX_n) = 2.4240n - 0.8981, \tag{14}$$

$$IRC(HX_n) = \frac{4\sqrt{12} + \sqrt{18}(2) + (8n - 24) + (4n - 8)\sqrt{24} + (2)(n^2 - 33n + 30)}{3n^2 - 5n + 2} - \frac{18n^2 - 30n + 12}{3n^2 - 3n + 1},$$

$$IRDIF(HX_n) = 10n - 4,$$

$$IRL(HX_n) = 4.8656n + 2.1201,$$

$$IRLU(HX_n) = 6n - 2,$$

$$IRLF(HX_n) = 4.8990n - 20.0913,$$

$$IRLA(HX_n) = \frac{168n - 76}{35},$$

$$IRDI(HX_n) = 13.1833n - 9.7311,$$

$$IRGA(HX_n) = 0.2449n - 0.0126.$$

*Proof.* By using the edge partition based on degrees of end vertices of each edge of the hexagonal network ( $HX_n$ ) given in Table 7, we have the following computations for irregularities of the hexagonal network ( $HX_n$ ):

$$\begin{aligned}
 VAR(HX_n) &= \frac{M_1(HX_n)}{n} - \left(\frac{2m}{n}\right)^2 = \frac{108n^2 - 228n + 114}{3n^2 - 3n + 1} - \left(\frac{2(9n^2 - 15n + 6)}{3n^2 - 3n + 1}\right)^2 = \frac{72n^3 + 243n^2 + 15n - 30}{(3n^2 - 3n + 1)^2}, \\
 AL(HX_n) &= \sum_{uv \in E} |d_u - d_v| = |3 - 4|(12) + |3 - 6|(6) + |4 - 4|(6n - 18) + |4 - 6|(12n - 24) \\
 &\quad + |6 - 6|(9n^2 - 33n + 30) = 6(4n - 3), \\
 IR1(HX_n) &= F(HX_n) - \frac{2m}{n}M_1(HX_n) = (648n^2 - 1560n + 906) - \frac{(2)(9n^2 - 15n + 6)(108n^2 - 228n + 114)}{3n^2 - 3n + 1} \\
 &= \frac{720n^3 - 2142n^2 + 1878n - 462}{3n^2 - 3n + 1}, \\
 IR2(HX_n) &= \sqrt{\frac{M_2(HX_n)}{m} - \frac{2m}{n}} = \sqrt{\frac{(324n^2 - 804n + 468)}{9n^2 - 15n + 6} - \frac{2(9n^2 - 15n + 6)}{3n^2 - 3n + 1}} \\
 &= \sqrt{\frac{108n^2 - 268n + 156}{3n^2 - 5n + 2} - \frac{18n^2 - 30n + 12}{3n^2 - 3n + 1}}, \\
 IRF(HX_n) &= F(HX_n) - 2M_2(HX_n) = (648n^2 - 1560n + 906) - 2(324n^2 - 804n + 468) = 6(8n - 5), \\
 IRFW(HX_n) &= \frac{IRF(HX_n)}{M_2(HX_n)} = \frac{(6)(8n - 5)}{324n^2 - 804n + 468} = \frac{(2)(8n - 5)}{(108n^2 - 268n + 156)}, \\
 IRA(HX_n) &= \sum_{uv \in E} (d_u^{-1/2} - d_v^{-1/2})^2 = \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}}\right)^2 (12) + \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{6}}\right)^2 (6) \\
 &\quad + \left(\frac{1}{\sqrt{4}} - \frac{1}{\sqrt{4}}\right)^2 (6n - 18) + \left(\frac{1}{\sqrt{4}} - \frac{1}{\sqrt{6}}\right)^2 (12n - 24) + \left(\frac{1}{\sqrt{6}} - \frac{1}{\sqrt{6}}\right)^2 (9n^2 - 33n + 30) = 0.1008n - 0.0418, \\
 IRB(HX_n) &= \sum_{uv \in E} (d_u^{1/2} - d_v^{1/2})^2 = (\sqrt{3} - \sqrt{4})^2 (12) + (\sqrt{3} - \sqrt{6})^2 (6) + (\sqrt{4} - \sqrt{6})^2 (12n - 24) = 2.4240n - 0.8981, \\
 IRC(HX_n) &= \sum_{uv \in E} \frac{\sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{\sqrt{12}(12) + (6)\sqrt{18} + (6n - 18)\sqrt{16} + (12n - 24)\sqrt{24} + (9n^2 - 33n + 30)(6)}{9n^2 - 15n + 6} \\
 &\quad - \frac{(2)(9n^2 - 15n + 6)}{3n^2 - 3n + 1} = \frac{4\sqrt{12} + \sqrt{18}(2) + (8n - 24) + (4n - 8)\sqrt{24} + (2)(9n^2 - 33n + 30)}{3n^2 - 5n + 2} - \frac{18n^2 - 30n + 12}{3n^2 - 3n + 1}, \\
 IRDIF(HX_n) &= \sum_{uv \in E} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| = (12) \left| \frac{3}{4} - \frac{4}{3} \right| + (6) \left| \frac{3}{6} - \frac{6}{3} \right| + (12n - 24) \left| \frac{4}{6} - \frac{6}{4} \right| = 10n - 4, \\
 IRL(HX_n) &= \sum_{uv \in E} |\ln d_u - \ln d_v| = (12)|\ln(3) - \ln(4)| + (6)|\ln(3) - \ln(6)| + (12n - 24)|\ln(4) - \ln(6)| = 4.8656n + 2.1201, \\
 IRLU(HX_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\min(d_u, d_v)} = (12) \frac{|3 - 4|}{3} + (6) \frac{|3 - 6|}{3} + (12n - 24) \frac{|4 - 6|}{4} = 6n - 2, \\
 IRLF(HX_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\sqrt{d_u d_v}} = (12) \frac{|3 - 4|}{\sqrt{12}} + (6) \frac{|6 - 3|}{\sqrt{18}} + (12n - 24) \frac{|6 - 4|}{\sqrt{24}} = 4.8990n - 20.0913, \\
 IRLA(HX_n) &= 2 \sum_{uv \in E} \frac{|d_u - d_v|}{d_u + d_v} = 2(12) \frac{|4 - 3|}{7} + 2(6) \frac{|6 - 3|}{9} + 2(12n - 24) \frac{|6 - 4|}{10} = \frac{168n - 76}{35}.
 \end{aligned} \tag{15}$$

Specific values of irregularities of the hexagonal network for different values of involved parameters are given in Table 8.

TABLE 8: Values of irregularity indices for  $HX_n$ .

Irregularity indices	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
$VAR(HX_n)$	300	31.5918	11.4848	6.2279	4.0633
$AL(HX_n)$	6	30	54	78	102
$IR1(HX_n)$	-6.0000	69.4286	280.7368	509.6757	743.9016
$IR2(HX_n)$	$\infty$	0.1770	0.3896	0.3377	0.2846
$IRF(HX_n)$	18	66	114	162	210
$IRFW(HX_n)$	-1.5000	0.4231	0.1173	0.0665	0.0462
$IRA(HX_n)$	0.0590	0.1598	0.2606	0.3614	0.4622
$IRB(HX_n)$	1.5259	3.9499	6.3739	8.7979	11.2219
$IRC(HX_n)$	$\infty$	0.1568	0.2887	0.2529	0.2146
$IRDIF(HX_n)$	6	16	26	36	46
$IRL(HX_n)$	2.7455	7.6111	12.4767	17.3423	22.2079
$IRLU(HX_n)$	4	10	16	22	28
$IRLF(HX_n)$	2.8077	7.7067	12.6057	17.5047	22.4037
$IRLA(HX_n)$	2.6286	7.4286	12.2286	17.0286	21.8286
$IRD1(HX_n)$	3.4522	16.6355	29.8188	43.0021	56.1854
$IRGA(HX_n)$	0.2323	0.4772	0.7221	0.9670	1.2119

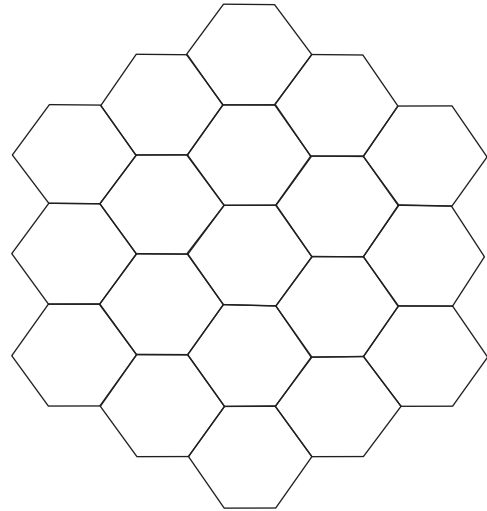


FIGURE 5: Structure of the honeycomb network.

$$\begin{aligned}
 IRD1(HX_n) &= \sum_{uv \in E} \ln\{1 + |d_u - d_v|\} = \ln\{1 + |3 - 4|\} (12) \\
 &\quad + \ln\{1 + |3 - 6|\} (6) + \ln\{1 + |4 - 6|\} (12n - 24) \\
 &= 13.1833n - 9.7311, \\
 IRGA(HX_n) &= \sum_{uv \in E} \ln \frac{d_u + d_v}{(2)\sqrt{d_u d_v}} = (12) \ln \frac{3+4}{(2)\sqrt{12}} \\
 &\quad + (6) \ln \frac{3+6}{(2)\sqrt{18}} + (12n-24) \ln \frac{4+6}{(2)\sqrt{24}} \\
 &= 0.2449n - 0.0126.
 \end{aligned} \tag{16}$$

**3.5. Honeycomb Network.** Honeycomb networks are commonly used as an illustration of benzoid hydrocarbons in chemistry in digital effects, cell phone transmitters, and image analysis. If we use dynamically hexagonal tiling in a specific pattern, honeycomb networks will be formed. An  $n$ -dimensional honeycomb network is referred to as  $HC_n$ , where  $n$  is the number of hexagons between the hexagon core and the boundary. Honeycomb network  $HC_n$  is built from  $HC_{n-1}$  by adding a hexagon sheet across the boundary of  $HC_{n-1}$  (refer Figure 5).

Table 9 shows the edge partition of the honeycomb network ( $HC_n$ ). The total number of vertices and edges in the honeycomb network ( $HC_n$ ) is  $6n^2$  and  $9n^2 - 3n$ .

**Theorem 5.** *The irregularity indices of the honeycomb network ( $HC_n$ ) for  $n > 1$  are*

$$\begin{aligned}
 VAR(HC_n) &= \frac{n-1}{n^2}, \\
 AL(HC_n) &= 12(n-1), \\
 IR1(HC_n) &= 30(n-1), \\
 IR2(HC_n) &= \sqrt{\frac{(27n^2 - 21n + 2)}{n(3n-1)}} - \frac{3n-1}{n}, \\
 IRF(HC_n) &= 12(n-1), \\
 IRFW(HC_n) &= \frac{14(n-1)}{(27n^2 - 21n + 2)}, \\
 IRA(HC_n) &= 0.2016(n-1), \\
 IRB(HC_n) &= 1.2120(n-1), \\
 IRC(HC_n) &= \frac{4 + 4\sqrt{6}(n-1) + 9n^2 - 15n + 6}{n(3n-1)} - \frac{3n-1}{n}, \\
 IRDIF(HC_n) &= 9.9996(n-1), \\
 IRL(HC_n) &= 4.8660(n-1), \\
 IRLU(HC_n) &= 6(n-1), \\
 IRLF(HC_n) &= 4.8990(n-1), \\
 IRLA(HC_n) &= 4.8000(n-1), \\
 IRD1(HC_n) &= 8.3172(n-1), \\
 IRGA(HC_n) &= 0.2448(n-1).
 \end{aligned} \tag{17}$$

*Proof.* By using the edge partition based on degrees of end vertices of each edge of the honeycomb network ( $HC_n$ )

given in Table 9, we have the following computations for the irregularity indices of the honeycomb network ( $HC_n$ ):

$$\begin{aligned}
 VAR(HC_n) &= \frac{M_1(HC_n)}{n} - \left(\frac{2m}{n}\right)^2 = \frac{3n(18n-10)}{6n^2} - \left(\frac{2(9n^2-3n)}{6n^2}\right)^2 = \frac{36n^2(n-1)}{36n^4} = \frac{n-1}{n^2}, \\
 AL(HC_n) &= \sum_{uv \in E} |d_u - d_v| = |2-2|(6) + |2-3|(12(n-1)) + |3-3|(9n^2-15n+6) = 12(n-1), \\
 IR1(HC_n) &= F(HC_n) - \frac{2m}{n}M_1(HC_n) = (162n^2 - 114n) - \frac{(2)(9n^2-3n)(54n^2-30)}{6n^2} = 30(n-1), \\
 IR2(HC_n) &= \sqrt{\frac{M_2(HC_n)}{m}} - \frac{2m}{n} = \sqrt{\frac{3(27n^2-21n+2)}{3n(3n-1)}} - \frac{2(3n)(3n-1)}{6n^2} = \sqrt{\frac{(27n^2-21n+2)}{n(3n-1)}} - \frac{3n-1}{n}, \\
 IRF(HC_n) &= F(HC_n) - 2M_2(HC_n) = (162n^2 - 114n) - 2(81n^2 - 63n + 6) = 12(n-1), \\
 IRFW(HC_n) &= \frac{IRF(HC_n)}{M_2(HC_n)} = \frac{12(n-1)}{81n^2 - 63n + 6} = \frac{14(n-1)}{(27n^2 - 21n + 2)}, \\
 IRA(HC_n) &= \sum_{uv \in E} (d_u^{-1/2} - d_v^{-1/2})^2 = \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right)^2 (6) + \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{3}}\right)^2 (12(n-1)) \\
 &\quad + \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}}\right)^2 (9^2 - 15n + 6) = 0.2016(n-1), \\
 IRB(HC_n) &= \sum_{uv \in E} (d_u^{1/2} - d_v^{1/2})^2 = (\sqrt{2} - \sqrt{2})^2 (6) + (\sqrt{2} - \sqrt{3})^2 (12n - 12) + (\sqrt{3} - \sqrt{3})^2 (9n^2 - 15n + 6) = 1.2120(n-1), \\
 IRC(HC_n) &= \sum_{uv \in E} \frac{\sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{\sqrt{4}(6) + (12n-12)\sqrt{6} + (9n^2-15n+6)\sqrt{9}}{3n(3n-1)} - \frac{2(3n)(3n-1)}{6n^2} \\
 &= \frac{4 + 4\sqrt{6}(n-1) + 9n^2 - 15n + 6}{n(3n-1)} - \frac{3n-1}{n}, \\
 IRDIF(HC_n) &= \sum_{uv \in E} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| = (6) \left| \frac{2}{2} - \frac{2}{2} \right| + (12n-12) \left| \frac{2}{3} - \frac{3}{2} \right| + (9n^2-15n+6) \left| \frac{3}{3} - \frac{3}{3} \right| = 9.9996(n-1), \\
 IRL(HC_n) &= \sum_{uv \in E} |\ln d_u - \ln d_v| = (6)|\ln(2) - \ln(2)| + (12n-12)|\ln(2) - \ln(3)| + (9n^2-15n+6)|\ln(3) - \ln(3)| = 4.8660(n-1), \\
 IRLU(HC_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\min(d_u, d_v)} = (6) \frac{|2-2|}{2} + (12n-12) \frac{|2-3|}{2} + (9n^2-15n+6) \frac{|3-3|}{3} = 6(n-1), \\
 IRLF(HC_n) &= \sum_{uv \in E} \frac{|d_u - d_v|}{\sqrt{d_u d_v}} = (6) \frac{|2-2|}{\sqrt{4}} + (12n-12) \frac{|3-2|}{\sqrt{6}} + (9n^2-15n+6) \frac{|3-3|}{\sqrt{9}} = 4.8990(n-1), \\
 IRLA(HC_n) &= 2 \sum_{uv \in E} \frac{|d_u - d_v|}{d_u + d_v} = 2(6) \frac{|2-2|}{4} + 2(12n-12) \frac{|3-2|}{5} + 2(9n^2-15n+6) \frac{|3-3|}{6} = 4.8000(n-1), \\
 IRD1(HC_n) &= \sum_{uv \in E} \ln\{1 + |d_u - d_v|\} = \ln\{1 + |2-2|\}(6) + \ln\{1 + |2-3|\}(12n-12) + \ln\{1 + |3-3|\}(9n^2-15n+6) = 8.3172(n-1), \\
 IRGA(HC_n) &= \sum_{uv \in E} \ln \frac{d_u + d_v}{(2)\sqrt{d_u d_v}} = (6)\ln \frac{2+2}{(2)\sqrt{4}} + (12n-12)\ln \frac{2+3}{(2)\sqrt{6}} + (9n^2-15n+6)\ln \frac{3+3}{(2)\sqrt{9}} = 0.2448(n-1).
 \end{aligned}$$



TABLE 9: Edge partition of the honeycomb network.

$(d_u, d_v)$ , where $uv \in E(G)$	(2, 2)	(2, 3)	(3, 3)
Number of edges	6	$12(n - 1)$	$9n^2 - 15n + 6$

TABLE 10: Values of irregularity indices for  $HC_n$ .

Irregularity indices	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
$VAR(HC_n)$	0.0000	0.2500	0.2222	0.1875	0.1600
$AL(HC_n)$	0	12	24	36	48
$IR1(HC_n)$	0	30	60	90	120
$IR2(HC_n)$	0.0000	0.1077	0.0871	0.0704	0.0586
$IRF(HC_n)$	0	12	24	36	48
$IRFW(HC_n)$	0.0000	0.0588	0.0440	0.0343	0.0280
$IRA(HC_n)$	0.0000	0.2016	0.4032	0.6048	0.8064
$IRB(HC_n)$	0.0000	1.2120	2.4240	3.6360	4.8480
$IRC(HC_n)$	0.0000	0.0798	0.0665	0.0544	0.0456
$IRDIF(HC_n)$	0.0000	9.9996	19.9992	29.9988	39.9984
$IRL(HC_n)$	0.0000	4.8660	9.7320	14.5980	19.4640
$IRLU(HC_n)$	0	6	12	18	24
$IRLF(HC_n)$	0.0000	4.8990	9.7980	14.6970	19.5960
$IRLA(HC_n)$	0.0000	4.8000	9.6000	14.4000	19.2000
$IRD1(HC_n)$	0.0000	8.3172	16.6344	24.9516	33.2688
$IRGA(HC_n)$	0.0000	0.2448	0.2496	0.7344	0.9792

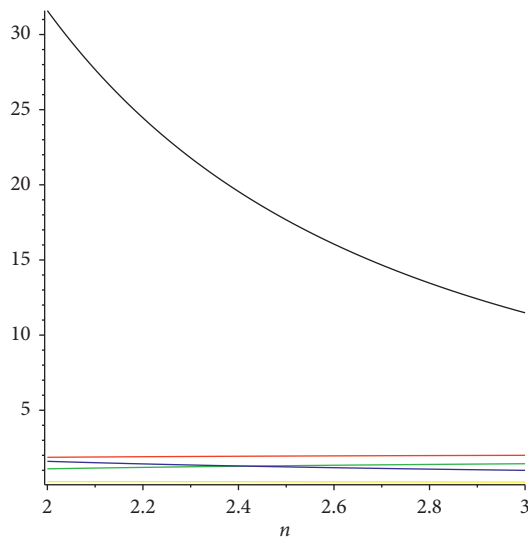


FIGURE 6: VAR of  $SL_n, CS_n, OX_n, HX_n$ , and  $HC_n$ .

Specific values for the irregularities of the honeycomb network for different values of involved parameters are given in Table 10.

### 4. Graphical Comparison

In this section, we will compare the results of irregularity indices of certain networks in graphical form. Different colors have been used to represent the behavior of irregularity indices for certain networks in the form of graphical lines. And these graphs have been generated by putting the

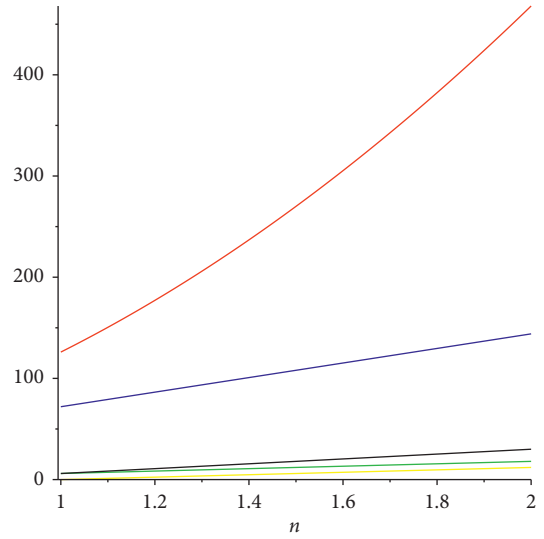


FIGURE 7: AL of  $SL_n, CS_n, OX_n, HX_n$ , and  $HC_n$ .

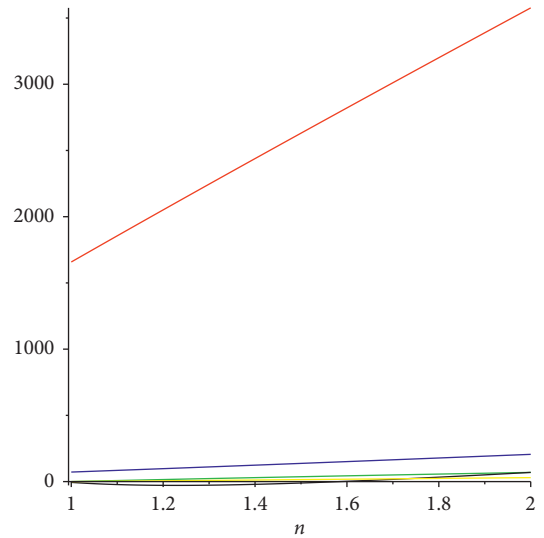


FIGURE 8: IR1 of  $SL_n, CS_n, OX_n, HX_n$ , and  $HC_n$ .

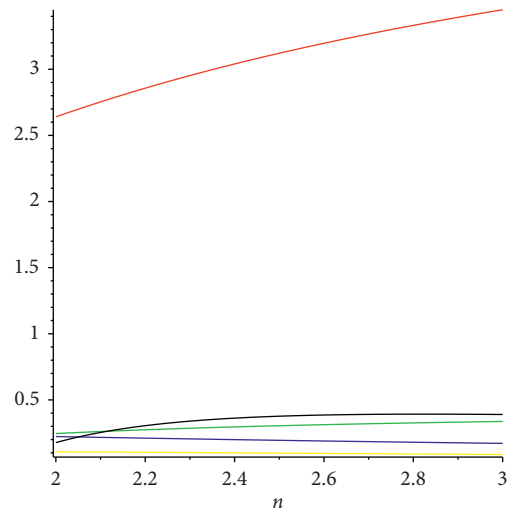


FIGURE 9: IR2 of  $SL_n, CS_n, OX_n, HX_n$ , and  $HC_n$ .

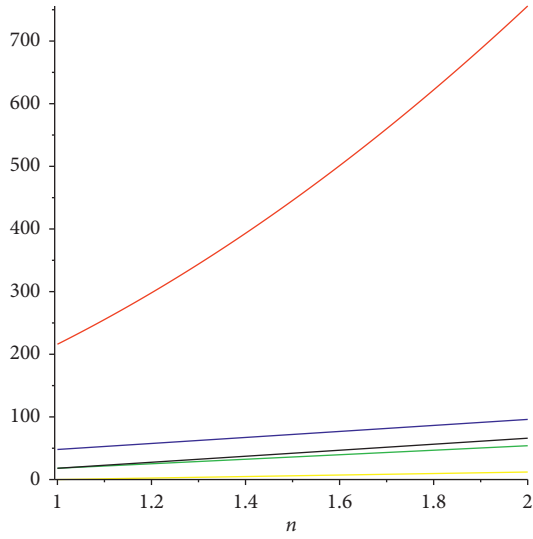


FIGURE 10: IRF of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

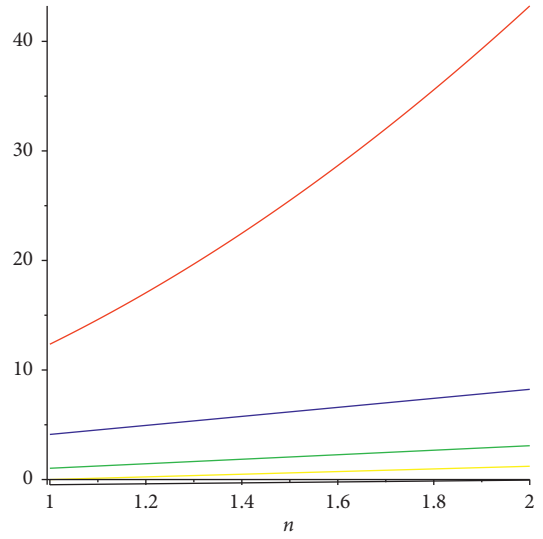


FIGURE 13: IRB of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

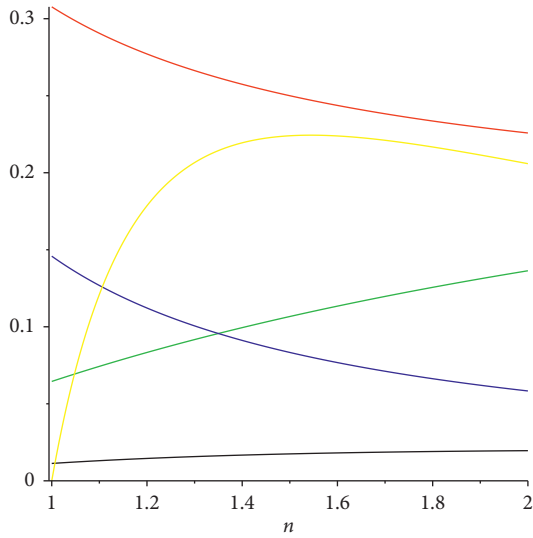


FIGURE 11: IRFW of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

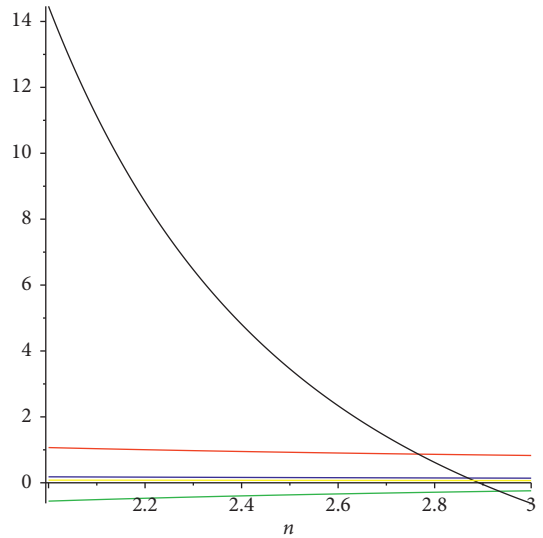


FIGURE 14: IRC of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

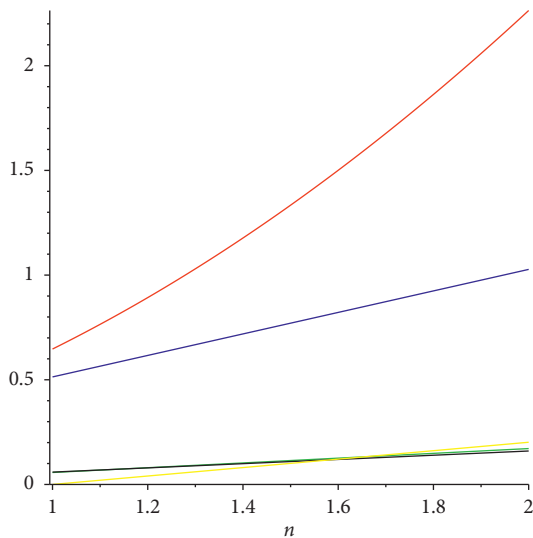


FIGURE 12: IRA of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

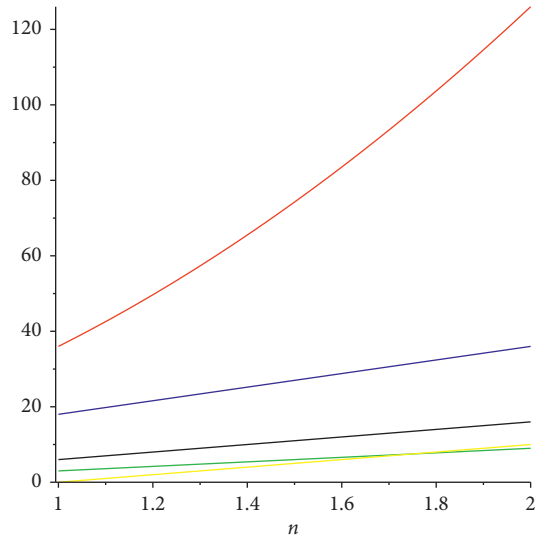


FIGURE 15: IRDIF of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

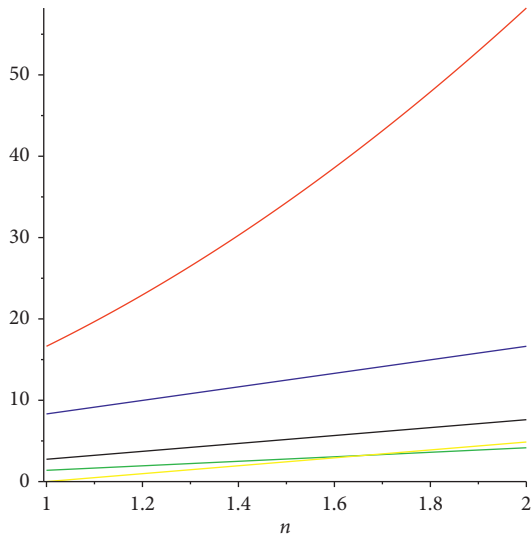


FIGURE 16: IRL of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

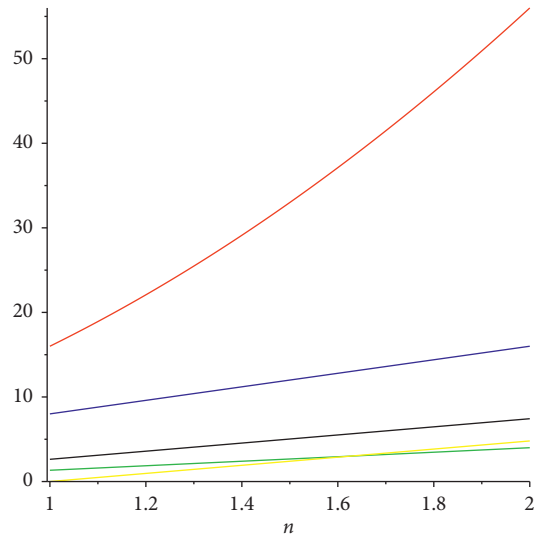


FIGURE 19: IRLA of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

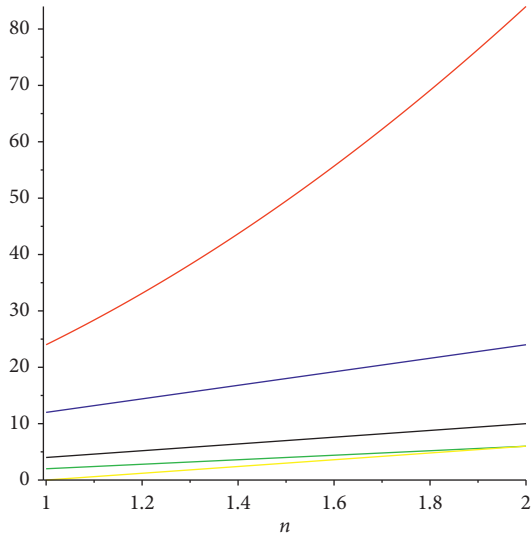


FIGURE 17: IRUL of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

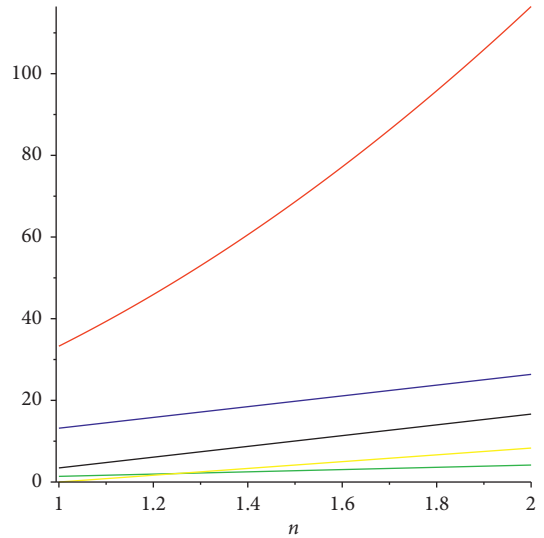


FIGURE 20: IRD1 of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

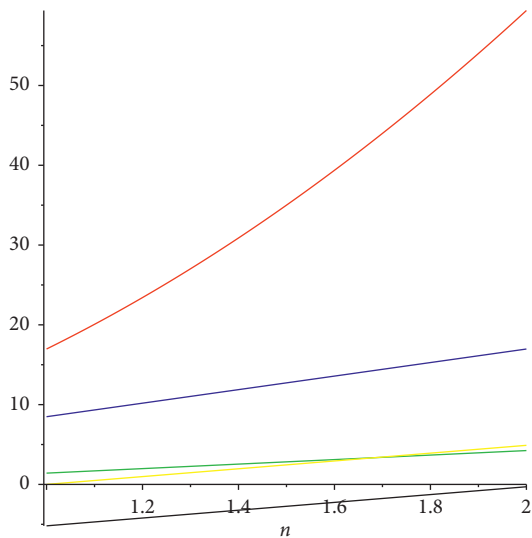


FIGURE 18: IRLF of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

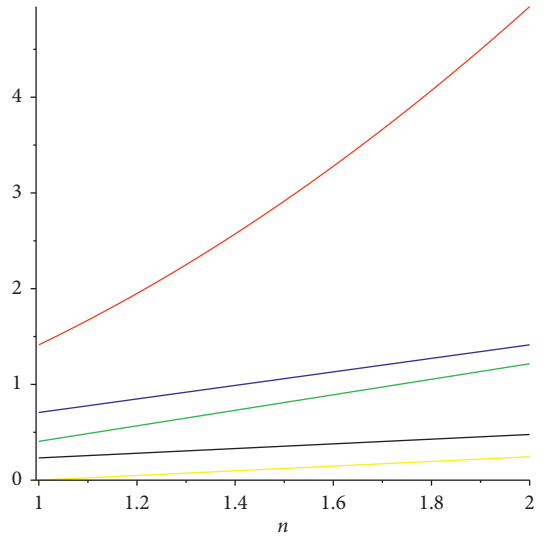


FIGURE 21: IRGA of  $SL_n, CS_n, OX_n, HX_n,$  and  $HC_n$ .

values of  $n$  along the  $X$ -axis in the results of irregularity indices along the  $Y$ -axis. Tables 1–5 represent such values of results which are obtained after putting values of  $n$ , and these values demonstrated in the tables will help us to generate graphical results.

In Figures 6–21, five different colors of graphical lines have been used. Red color represents irregular behavior in the silicate network ( $SL_n$ ), and blue color represents irregular behavior in the chain silicate network ( $CS_n$ ). Similarly, other colors which are green, purple, and black represent irregular behavior in the oxide network ( $OX_n$ ), hexagonal network ( $HX_n$ ), and honeycomb network ( $HC_n$ ), respectively. In Figures 6–21, values of  $n$  are plotted along the  $X$ -axis and behavior of all selected irregularity indices along the  $Y$ -axis.

## 5. Concluding Remarks

In this paper, we have computed irregularity indices of certain networks such as silicate network, chain silicate network, oxide network, hexagonal network, and honeycomb network. These results are valuable and helpful to understand deep irregular behavior of certain networks. These results are also useful for researchers to understand how these networks can be constructed with different irregular properties [16–20].

## Data Availability

All the data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

All authors contributed equally to this paper.

## Acknowledgments

This work was sponsored in part by Zhejiang Province Public Welfare Project (LGG20F030001) and Quzhou Science and Technology Project (2019K18).

## References

- [1] H. F. M. Salih, S. M. Mershkhan, and S. M. Mershkhan, "Generalized the Liouville's and Möbius functions of graph," *Open Journal of Mathematical Sciences*, vol. 4, no. 1, pp. 186–194, 2020.
- [2] F. Asif, Z. Zahid, and S. Zafar, "Leap Zagreb and leap hyper-Zagreb indices of Jahangir and Jahangir derived graphs," *Engineering and Applied Science Letter*, vol. 3, no. 2, pp. 1–8, 2020.
- [3] D. Wang, H. Ahmad, and W. Nazeer, "Hosoya and Harary polynomials of TUC 4 nanotube," *Mathematical Methods in the Applied Sciences*, vol. 58, 2020.
- [4] W. Gao, M. Younas, A. Farooq, A. Virk, and W. Nazeer, "Some reverse degree-based topological indices and polynomials of dendrimers," *Mathematics*, vol. 6, no. 10, p. 214, 2018.
- [5] U. Ali, Y. Ahmad, and M. S. Sardar, "On 3-total edge product cordial labeling of tadpole, book and flower graphs," 2020.
- [6] A. Tabassum, M. A. Lingua, M. A. Umar, M. Perveen, and A. Raheem, "Antimagicness of subdivided fans," *Open Journal of Mathematical Sciences*, vol. 4, no. 1, pp. 18–22, 2020.
- [7] H. Wiener, "Structural determination of paraffin boiling points," *Journal of the American Chemical Society*, vol. 69, no. 1, pp. 17–20, 1947.
- [8] M. Randić, "Characterization of molecular branching," *Journal of the American Chemical Society*, vol. 97, no. 23, pp. 6609–6615, 1975.
- [9] I. Gutman and N. Trinajstić, "Graph theory and molecular orbitals. Total  $\phi$ -electron energy of alternant hydrocarbons," *Chemical Physics Letters*, vol. 17, no. 4, pp. 535–538, 1972.
- [10] I. Gutman and K. C. Das, "The first Zagreb indices 30 years after," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 50, pp. 83–92, 2004.
- [11] B. Zhou and I. Gutman, "Further properties of Zagreb indices," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 54, pp. 233–239, 2005.
- [12] A. Farooq, M. Habib, A. Mahboob, W. Nazeer, and S. M. Kang, "Zagreb polynomials and redefined zagreb indices of dendrimers and polyomino chains," *Open Chemistry*, vol. 17, no. 1, pp. 1374–1381, 2019.
- [13] Y. Kwun, A. Virk, W. Nazeer, M. Rehman, and S. Kang, "On the multiplicative degree-based topological indices of silicon-carbon  $Si_2C_3-I[p,q]$  and  $Si_2C_3-II[p,q]$ ," *Symmetry*, vol. 10, no. 8, p. 320, 2018.
- [14] M. S. Ahmad, W. Nazeer, S. M. Kang, M. Imran, and W. Gao, "Calculating degree-based topological indices of dominating David derived networks," *Open Physics*, vol. 15, no. 1, pp. 1015–1021, 2017.
- [15] T. Reti, R. Sharafzadi, A. D. Kiss, and H. Haghbin, "Graph irregularity indices used as a molecular descriptor in QSPR studies," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 79, pp. 509–524, 2018.
- [16] J.-B. Liu, C. Wang, S. Wang, and B. Wei, "Zagreb indices and multiplicative zagreb indices of eulerian graphs," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 42, no. 1, pp. 67–78, 2019.
- [17] J.-B. Liu, J. Zhao, and Z. X. Zhu, "On the number of spanning trees and normalized Laplacian of linear octagonal quadrilateral networks," *International Journal of Quantum Chemistry*, vol. 119, p. 25971, 2019.
- [18] J.-B. Liu, J. Zhao, J. Min, and J. D. Cao, *On the Hosoya Index of Graphs Formed by a Fractal Graph, Fractals-Complex Geometry Patterns and Scaling in Nature and Society*, 2019.
- [19] J.-B. Liu, J. Zhao, and Z.-Q. Cai, "On the generalized adjacency, Laplacian and signless Laplacian spectra of the weighted edge corona networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 540, p. 123073, 2020.
- [20] J.-B. Liu, J. Zhao, H. He, and Z. Shao, "Valency-based topological descriptors and structural property of the generalized sierpiński networks," *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.

## Research Article

# Theoretical Calculation and Application Test of Lift Force for Ideal Electric Asymmetric Capacitor

Jian-Guo Huang<sup>1</sup> and Xiang-Yu Cheng<sup>2,3</sup> 

<sup>1</sup>Anhui Technical College of Water Resources and Hydroelectric Power, Hefei 231603, China

<sup>2</sup>Hefei Institute of Physical Science, Chinese Academy of Science, Hefei 231603, China

<sup>3</sup>No. 38 Research Institute of China Electronics Technology Group Corporation, Hefei 230088, China

Correspondence should be addressed to Xiang-Yu Cheng; [chancellor@163.com](mailto:chancellor@163.com)

Received 9 June 2020; Accepted 3 July 2020; Published 5 August 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Jian-Guo Huang and Xiang-Yu Cheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The asymmetric capacitor's lift force formula can be obtained on the basis of literature review, which can almost cover all practical forms of asymmetric capacity forms. But there are still some problems we should solve. The first and foremost one is whether the formulas are correct and can they be verified in engineering practices? On the contrary, the parameter  $q$  in the formulas is normally unknown in the beginning of calculations, how can we get or reckon up it so as to use the formulas smoothly? In this paper, we set out to solve these questions.

## 1. Introduction

How can we solve lift force produced by a lifter formed of an asymmetric capacitor? Based on some hypothetical conditions, a formula was obtained through three methods in an ideal scenario [1, 2]. But an unknown parameter  $q$  is still contained so that numerical calculations are difficult to carry out. This paper intends to solve this problem by eliminating the unknown factor in the hope that the formula can be effortlessly put into practical application and engineering. Following that, experimental tests and practical estimations are provided to verify its validity.

In former papers [1, 2], the same result is acquired through three ways using the following equation of electric lift force of asymmetric capacitor loaded by high voltage in ideal condition:

$$f = \frac{q^2}{\varepsilon} \left( \frac{1}{S_1} - \frac{1}{S_2} \right). \quad (1)$$

An unknown variable  $q$  is still included in the above formula. In order to solve this problem thoroughly, the carried charge  $q$  should be figured out. Normally, carried charge  $q$  of capacitor is relevant to the voltage  $U$  and the capacitance  $C$ . The voltage  $U$  can be known. But the capacitance  $C$  is difficult to calculate when the capacitor is in irregular shape.

Nevertheless, the analysis of hypothetical predetermined conditions verifies that the capacitance  $C$  of the asymmetric capacity is calculable. When the small plate of asymmetric capacitor is in a slender cylinder form, its capacitance could be estimated at a cylindrical way. When the small plate of asymmetric capacitor [3–5] is in sphere form, its capacitance could be estimated at a spherical way. The result might not be ideal in the case of precision. It can still be applied to estimation in engineering assessment [6, 7]. Furthermore, the subsequent test data verified that the estimate result was fairly accurate unexpectedly.

## 2. Theoretical Derivation

Regarding the reason why the experimental result is more precise than expected, the analysis of the unique characteristics of the asymmetric capacitor has presented several objective reasons: (1) the distance  $d$  between two plates is more larger than the dimension of surface area  $S_1$  of plate 1 (small plate), that is,  $d \gg S_1/l$  or  $d \gg \sqrt{S_1}$ ; (2) the area of plate 2 (large plate) is larger than that of plate 1, that is,  $S_2 \gg S_1$ ; and (3) the voltage loaded between two plates is below the breakdown voltage that is relevant to the gap distance.

Under the initial condition, we begin to deduce capacitance of the asymmetric capacitor [8–10] and then to estimate its lift force [11, 12]. Deducing processes are as follows:

- (1) For  $S_2 \gg S_1$ , when high voltage is loaded on two plates, the electric field intensity on plate 1,  $E_1 = q/\epsilon S_1$ , is larger than that on plate 2,  $E_2 = q/\epsilon S_2$ , i.e.,  $E_1 \gg E_2$ . It leads to the voltage drop  $\Delta U = \Delta d \cdot E$ , which mainly centralizes around plate 1. So when calculating the capacitance  $C = q/\Delta U = q/\int_{U_1}^{U_2} dU$ , the field intensity near plate 1 should be taken into major consideration. That is to say, the capacitance calculation can be carried out by combining the following equations (2), (3), (4), (5), and (6):

$$E = \frac{q}{\epsilon S}, \quad (2)$$

$$du = -E \cdot dr, \quad (3)$$

$$C = \frac{q}{\Delta U}, \quad (4)$$

$$r \propto \frac{S}{l} \text{ (for thin wire and board plates),} \quad (5)$$

$$r \propto \sqrt{S} \text{ (for sphere point and board plates).} \quad (6)$$

Equations (5) or (6) can be also written as

$$r = k_{sh1} \cdot \frac{S}{l}, \quad (7)$$

$$r = k_{sh2} \cdot \sqrt{S},$$

where  $r$  is the nominal dimensional size of plate 1,  $l$  is the length of plate 1, and  $k_{sh1}$  and  $k_{sh2}$  are the shape coefficient relevant to the plates' structure size.

- (2) Because the distance  $d$  between two plates is far larger than the nominal size of plate 1  $r$ , to simplify the calculation, we assume that surface charge of plate 1 is uniformly distributed, and voltage drop of thin wire plate or spherical capacitor plate is integrated for estimating the capacitance in magnitudes. The details are shown as follows.

For a thin wire small plate capacitor, we can take

$$\begin{aligned} S_1 &= 2\pi R_1 l, \\ S_2 &= 2H_2 l, \end{aligned} \quad (8)$$

where  $H_2$  is the width of the board plate.

Because

$$E = \frac{q}{\epsilon \cdot S}, \quad (9)$$

referring to equation (3), we get

$$du = -\frac{q}{\epsilon \cdot S} \cdot dr. \quad (10)$$

Mainly considering the electric field variation beside the thin wire, we have

$$S_{cir} = 2\pi r l. \quad (11)$$

Considering the effective fan-shaped part, we have

$$\begin{aligned} S_{fan} &= S_{cir} \cdot \int_{R_1+d}^{R_1+d+H_2} \frac{2dr}{2\pi r} \\ &= S_{cir} \cdot \frac{1}{\pi} \ln \left( \frac{R_1+d+H_2}{R_1+d} \right) \\ &= 2rl \cdot \ln \left( 1 + \frac{H_2}{R_1+d} \right), \end{aligned} \quad (12)$$

$$\Rightarrow dS = dS_{fan}$$

$$\Rightarrow dr = \frac{1}{2l \ln \left( 1 + \frac{H_2}{R_1+d} \right)} dS.$$

Integrating both sides of equation (10), we obtain



$$\begin{aligned}
 \Delta U &= \int_{U_1}^{U_2} du = - \int_{r_1}^{r_2} \frac{q}{\epsilon S} dr \\
 &= -\frac{q}{\epsilon} \int_{S_1}^{S_2} \frac{1}{2Sl \ln(1 + (H_2/(R_1 + d)))} dS \\
 &= -\frac{q \ln(S_2/S_1)}{2\epsilon l \ln(1 + (H_2/(R_1 + d)))} \\
 &= (q \ln[2\pi H_2 l / (2\pi R_1 l \cdot (1/\pi) \ln((R_1 + d + H_2)/(R_1 + d)))] (2\epsilon l \ln(1 + (H_2/(R_1 + d))))^{-1}, \\
 &= -((q \ln[H_2/R_1 \ln(1 + (H_2/(R_1 + d)))] (2\epsilon l \ln(1 + (H_2/(R_1 + d))))^{-1}).
 \end{aligned} \tag{13}$$

So we get the capacitance

$$C = \frac{q}{-\Delta U} = \frac{2\epsilon l \ln(1 + (H_2/(R_1 + d)))}{\ln[H_2/R_1 \ln(1 + (H_2/(R_1 + d)))]}. \tag{14}$$

For a spherical small plate capacitor, we can take

$$S_1 = 4\pi r^2. \tag{15}$$

Combining equation (2), we get

$$E = \frac{q}{\epsilon \cdot 4\pi r^2}. \tag{16}$$

Referring to equation (3), we get

$$du = -\frac{q}{\epsilon \cdot 4\pi r^2} \cdot dr. \tag{17}$$

Integrating both sides, we obtain

$$\Delta U = \int_{U_1}^{U_2} du = - \int_{R_1}^{R_1+d} \frac{q}{\epsilon \cdot 4\pi r^2} dr. \tag{18}$$

For the distance  $d \gg R_1$ , we have

$$\begin{aligned}
 \Delta U &= U_2 - U_1 = - \int_{R_1}^{R_1+d} \frac{q}{4\pi r^2 \epsilon} dr \\
 &= -\frac{q}{4\pi \epsilon} \left( \frac{1}{R_1} - \frac{1}{R_1 + d} \right) \approx -\frac{q}{4\pi \epsilon R_1}.
 \end{aligned} \tag{19}$$

So we get the capacitance

$$C = \frac{q}{-\Delta U} = 4\pi \epsilon R_1. \tag{20}$$

(3) We can calculate the electric lift force of asymmetric capacitor loaded by high voltage with the capacitance  $C$ .

For thin wire small plate capacitor, using equation (14),

$$\begin{aligned}
 f &= \frac{q^2}{\epsilon} \cdot \left( \frac{1}{S_1} - \frac{1}{S_2} \right) = \frac{(UC)^2}{\epsilon} \left( \frac{1}{S_1} - \frac{1}{S_2} \right) \\
 &= \frac{1}{\epsilon} \left\{ U \cdot \frac{2\epsilon l \ln(1 + (H_2/(R_1 + d)))}{\ln[H_2/R_1 \ln(1 + (H_2/(R_1 + d)))]} \right\}^2 \left( \frac{1}{2\pi R_1 l} - \frac{1}{2H_2 l} \right) \\
 &= \frac{2\epsilon U^2 \ln^2(1 + (H_2/(R_1 + d)))}{\ln^2[H_2/R_1 \ln(1 + (H_2/(R_1 + d)))]} \left( \frac{1}{\pi R_1} - \frac{1}{H_2} \right).
 \end{aligned} \tag{21}$$

This is the lift force formula about a normal lifter in thin wire asymmetric capacitor form under high voltage loaded.

For spherical small plate capacitor, using equation (20),

$$\begin{aligned}
 f &= \frac{q^2}{\varepsilon} \cdot \left( \frac{1}{S_1} - \frac{1}{S_2} \right) \\
 &= \frac{(UC)^2}{\varepsilon} \cdot \left( \frac{1}{S_1} - \frac{1}{S_2} \right) \\
 &= \frac{(U \cdot 4\pi R\varepsilon)^2}{\varepsilon} \cdot \left( \frac{1}{S_1} - \frac{1}{S_2} \right).
 \end{aligned} \tag{22}$$

Considering the condition  $S_2 \gg S_1$ , we have

$$f = \frac{(U \cdot 4\pi R\varepsilon)^2}{\varepsilon} \cdot \frac{1}{S_1} = U^2 \cdot 4\pi\varepsilon \cdot 4\pi R^2 \cdot \frac{1}{S_1}. \tag{23}$$

If simplifying calculation as a spherical plate, the surface area of plate 1  $S_1 = 4\pi R^2$ , we can get

$$f = 4\pi\varepsilon U^2. \tag{24}$$

This is the concised formula that finally turned out, from which we can tell the maximum lift force produced by spherical asymmetric capacitor under high voltage loaded.

### 3. Formula Application

The formulas are exerted on two applications to test their validity. They are, respectively, lift force estimation of a electricity lifter [13, 14] and a high-voltage rising hair experiment [15, 16].

**3.1. Lift Force Estimation of a Lifter.** When a lifter loads with high voltage (Figure 1)  $U = 30$  kV, what is the lift force produced by the lifter? The physical sizes (Figure 2) of the lifter are, respectively, length of thin wire  $l = 15$  cm  $\times 3 = 0.45$  m, radius of thin wire  $R_1 = 0.08$  mm  $= 8 \times 10^{-5}$  m, width of large board  $H_2 = 30$  cm  $= 0.3$  m, and gap between 2 plates  $d = 40$  cm  $= 0.4$  m. We take permittivity of atmosphere  $\varepsilon = 8.85 \times 10^{-12}$  C<sup>2</sup>N<sup>-1</sup>m<sup>-2</sup>.

On the initial conditions, using equation (21), we have

$$\begin{aligned}
 f &= \left( (2\varepsilon U^2 \ln^2(1 + (H_2/(R_1 + d)))) (\ln^2[H_2/R_1 \ln(1 + (H_2/(R_1 + d)))]^{-1}) \cdot \left( \frac{1}{\pi R_1} - \frac{1}{H_2} \right) \right)^{-1} \\
 &= \left( 2 \times 8.85 \times 10^{-12} \times 0.45 \times 30000^2 \times \ln^2(1 + (0.3/(8 \times 10^{-5} + 0.4))) \right) \left( \ln^2[0.3/8 \times 10^{-5} \times \ln(1 + (0.3/(8 \times 10^{-5} + 0.4)))] \right)^{-1} \\
 &\quad \cdot \left( \frac{1}{\pi \times 8 \times 10^{-5}} - \frac{1}{0.3} \right) \\
 &= 0.115 \text{ N} \approx 11.7 \text{ gf}.
 \end{aligned} \tag{25}$$

That is to say, a lifter loaded with 30 kV voltage can produce a largest lift force of 11.7 gf.

**3.2. Lift Force Estimation of High-Voltage Charged Conducting Sphere.** As we know, when a high voltage loads on human body, our hair may be lifted up by the static electricity [17]. But there is no precise data or concrete calculating method of the length of the lifted hair by the high voltage. By equation (24), the mentioned problem can be solved. We can use the formula to quantitatively calculate the hair length lifted by the static electric field. The details are shown as follows.

**3.2.1. Initial Conditions.** Voltage loaded on the head  $U = 100$  kV, diameter of the human head  $D = 10$  cm  $= 1 \times 10^{-2}$  m, average diameter of the hair  $D_h = 70$   $\mu$ m  $= 70 \times 10^{-6}$  m, density of the hair  $\rho = 1.25$  g/cm<sup>3</sup>  $= 1.25 \times 10^3$  kg/m<sup>3</sup>, and permittivity of atmosphere  $\varepsilon = 8.85 \times 10^{-12}$  C<sup>2</sup>N<sup>-1</sup>m<sup>-2</sup>.

**3.2.2. Target Problem.** When the voltage is loaded on the hair under the above initial conditions, what is the maximum length of the hair ( $l_h = ?$ ) that can be lifted up?

**3.2.3. Solving Process.** In this case, the head and ground can be considered as the two plates of asymmetric, where the head may be regarded as a small plate and its area of sphere surface is  $S_1$  and the distant ground as a large plate and its area of flat surface  $S_2$ . However,  $S_2 \gg S_1$ , and the distance between the two plates  $d \gg \sqrt{S_1}$ . The voltage loaded between small plate (head) and large plate (ground) should not reach to breakdown threshold. Under this condition, we apply equation (24) to calculate the electrostatic lift force acted on hairs.

The sum of the electrostatic lift force acted on the hair is

$$\begin{aligned}
 f &= 4\pi\varepsilon U^2 \\
 &= 4 \times 3.14 \times 8.85 \times 10^{-12} \times (100 \times 10^3)^2 \\
 &= 1.11 \text{ N}.
 \end{aligned} \tag{26}$$

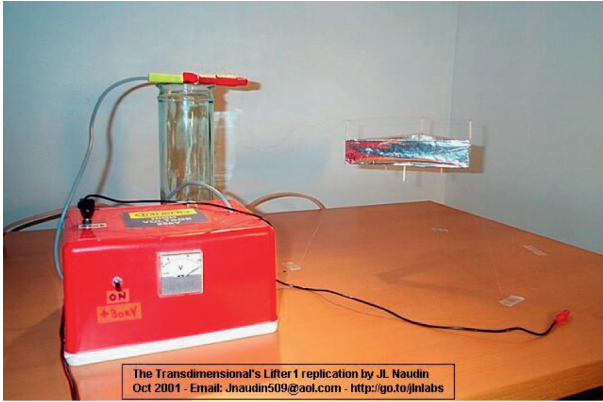


FIGURE 1: A lifter flying up loaded with high-voltage power.

The surface area of the head is

$$\begin{aligned} S_1 &= \pi D^2 \\ &= 3.14 \times (10^{-2})^2 \\ &= 3.14 \times 10^{-4} \text{m}^2. \end{aligned} \tag{27}$$

The intensity of pressure supported by electrostatic force is

$$p = \frac{f}{S_1} = \frac{1.11 \times 10^6}{3.14 \times 10^{-4}} = 3.54 \times 10^3 \text{Pa}. \tag{28}$$

The average transverse area of a hair is

$$\begin{aligned} S_h &= \frac{\pi D_h^2}{4} \\ &= \frac{3.14 \times (70 \times 10^{-6})^2}{4} \\ &= 3.8 \times 10^{-9} \text{m}^2. \end{aligned} \tag{29}$$

Head surface area occupied by a hair can support a mass by the electrostatic force:

$$\begin{aligned} m_h &= \frac{G_h}{g} \\ &= \frac{3.8 \times 10^{-9} \times 3.54 \times 10^3}{9.8} \\ &= 1.4 \times 10^{-6} \text{kg}. \end{aligned} \tag{30}$$

The mass is converted into length of a hair:

$$\begin{aligned} l_h &= \frac{V_h}{S_h} \\ &= \frac{1.4 \times 10^{-6}}{1.25 \times 10^3 \times 3.8 \times 10^{-9}} \text{m} \\ &= 0.29 \text{m} \end{aligned} \tag{31}$$

Therefore, the final result is obtained: when human being's hair is loaded by high-voltage static electricity of DC

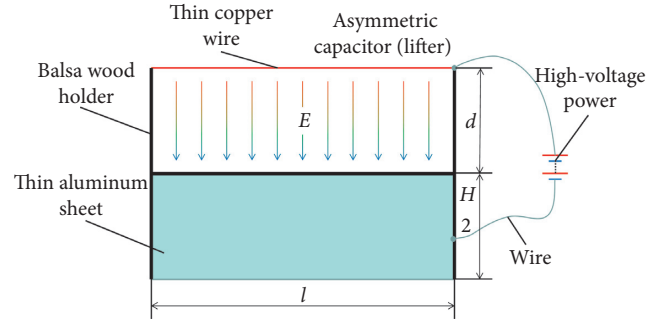


FIGURE 2: An asymmetric capacitor in lifter form.

100 kV through a conducting metallic ball, approximately 29 cm length hair floats up into air.

#### 4. Explanation and Conclusion

Based on some assumptions with simplified calculation, we derived lift force formula produced by an asymmetric capacitor in different conditions, with which the assess in certain survey and qualitative research can be undertaken in spite of unsatisfying precision. The method also provides a convenient way to calculate static electricity lift capacity produced by an asymmetric capacitor or lift force of lifters. It also contributes to the parameter optimization in designing [18] a larger load force of lifter formed by an asymmetric capacitor.

#### Data Availability

The data used to support the findings of this study are included within the article.

#### Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

#### Acknowledgments

The authors gratefully acknowledge the support of the Thirteenth Five-Year Plan of Hefei Institute of Physical Science of Chinese Academy of Science (Grant no. Y86CT21051, "Electric and Magnetic Propulsion System"), Research Activity Funding of Postdoctoral Fellow of Anhui Province (Grant no. 2018B250, "High-Energy Ions Accelerated Thruster"), and Natural Science Research Project of Anhui Education Department (Grant no. KJ2018A0725, "The Uniformity Optimization and Software Development for MRI Magnets"). A portion of this work was supported by the High Magnetic Field Laboratory of Anhui Province.

#### References

[1] X. Cheng, G. Kuang, Y. Zhang et al., "Theoretical calculation of lift force for ideal electric asymmetric capacitor loaded by high voltage," *Engineering*, vol. 12, no. 1, pp. 33-40, 2020.

- [2] Y. Zhang, X. Cheng, P. Huang et al., “Theoretical calculation of lift force for general electric asymmetric capacitor loaded by high voltage,” *Engineering*, vol. 12, no. 1, pp. 41–46, 2020.
- [3] W. G. Pell and B. E. Conway, “Peculiarities and requirements of asymmetric capacitor devices based on combination of capacitor and battery-type electrodes,” *Journal of Power Sources*, vol. 136, no. 2, pp. 334–345, 2004.
- [4] R. Ohyama, A. Watson, and J. S. Chang, “Electrical current conduction and electrohydrodynamically induced fluid flow in an AW type EHD pump,” *Journal of Electrostatics*, vol. 53, no. 2, pp. 147–158, 2001.
- [5] A. M. Alexandre and M. J. Pinheiro, “Modeling of an EHD corona flow in nitrogen gas using an asymmetric capacitor for propulsion,” *Journal of Electrostatics*, vol. 69, no. 2, pp. 133–138, 2011.
- [6] I. Celik, M. Klein, and J. Janicka, “Assessment measures for engineering LES applications,” *Journal of Fluids Engineering*, vol. 131, no. 3, Article ID 031102, 2009.
- [7] K. F. Long, “The role of speculative science in driving technology,” in *Deep Space Propulsion*, Springer, New York, NY, USA, 2012.
- [8] T. T. Brown, “A method of and an apparatus or machine for producing force or motion,” UK Patent No. 300311, 1928.
- [9] J. Primas, M. Malík, D. Jašíková, and V. Kopecký, “Force on high voltage capacitor with asymmetrical electrodes,” in *Proceedings of the of the WASET 2010 Conference*, pp. 335–339, Amsterdam, Netherlands, June 2010.
- [10] T. Bahder and Ch. Fazi, “Force on an asymmetrical capacitor,” Army Report ARL-TR-3005, Army Research Laboratory, Adelphi, MA, USA, 2003.
- [11] M. Cattani, A. Vannucci, and V. G. Souza, “Lifter—high voltage plasma levitation device,” *Revista Brasileira de Ensino de Física*, vol. 37, no. 3, pp. 3307–3311, 2015.
- [12] J. Rincón, L. Martínez, and R. Correa, “Experiments with an electrodynamic lifter prototype,” *Dyna*, vol. 77, no. 164, pp. 167–177, 2010.
- [13] J.-B. Liu, J. Zhao, and Z.-Q. Cai, “On the generalized adjacency, Laplacian and signless Laplacian spectra of the weighted edge corona networks,” *Physica A: Statistical Mechanics and Its Applications*, vol. 540, Article ID 123073, 2020.
- [14] J.-B. Liu, J. Zhao, H. He, and Z. Shao, “Valency-based topological descriptors and structural property of the generalized sierpiński networks,” *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.
- [15] “What causes static hair?” <https://www.reference.com/beauty-fashion/causes-static-hair-c9e490c1c0aacfd4>.
- [16] K. Koike, O. Yoshida, A. Mamada et al., “Structural analysis of human hair fibers under the ultra-high voltage electron microscope,” *Journal of Cosmetic Science*, vol. 55, no. Suppl 2, pp. S25–S27, 2004.
- [17] V. K. Otero, *The Process of Learning about Static Electricity and the Role of the Computer Simulator*, University of California and San Diego State University, San Diego, CA, USA, 2001.
- [18] H. Eschenauer, J. Koski, and A. Osyczka, *Multicriteria Design Optimization*, Springer, Berlin, Germany, 1990.

## Research Article

# Irregularity Measures for Metal-Organic Networks

Xuan Guo,<sup>1</sup> Yu-Ming Chu ,<sup>2,3</sup> Muhammad Khalid Hashmi,<sup>4</sup> Abaid Ur Rehman Virk ,<sup>4</sup> and Jingjing Li<sup>5</sup>

<sup>1</sup>Automation School of Wuhan University of Technology, Wuhan 430070, China

<sup>2</sup>Department of Mathematics, Huzhou University, Huzhou 313000, China

<sup>3</sup>Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science & Technology, Changsha 410114, China

<sup>4</sup>Department of Mathematics, University of Lahore, Lahore, Pakistan

<sup>5</sup>China Ship Development and Design Center, Wuhan 430064, China

Correspondence should be addressed to Yu-Ming Chu; chuyuming@zjhu.edu.cn

Received 30 April 2020; Accepted 16 June 2020; Published 1 August 2020

Guest Editor: Shaohui Wang

Copyright © 2020 Xuan Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Topological index plays an important role in predicting physicochemical properties of a molecular structure. With the help of the topological index, we can associate a single number with a molecular graph. Drugs and other chemical compounds are frequently demonstrated as different polygonal shapes, trees, graphs, etc. In this paper, we will compute irregularity indices for metal-organic networks.

## 1. Introduction

Every planet is a blend of various kinds of parts, and every part has a significant commitment in the composition of the earth. The most significant parts in the earth are hydrogen, oxygen, and nitrogen. Subatomic hydrogen is one of the segments that has cordial situation and is a cutting edge wellspring of vitality [1, 2]. As a gas, it can likewise be used in energy units to control engines. Among the various kinds of gases, hydrogen needs smell that makes any hole acknowledgment practically difficult to human creatures. The ongoing standards settled by the United States Vitality Sector put emphasis on the speed proficient of the gadget that can detect one percent by volume of drab and also scent-free atomic hydrogen in climate in less than sixty seconds in particular [3–6].

Jang et al. [7] presented an exceptionally quick hydrogen recognizing device comprising organic ligands and metals recognized with the assistance of palladium nanowires which can perceive hydrogen stages lower than 1 percent in just seven seconds. Moreover, other than distinguishing and detecting, the MONs appear exceptionally valuable with synthetic and physical properties, for example, grafting active groups [8], changing natural ligands [9], impregnating

reasonable active materials [10], postsynthetic ligands, ion trade [11], and getting ready composites with useable substance [12].

Graph theory provides the interesting appliance in mathematical chemistry that is used to compute the various kinds of chemical compounds by means of graph theory and predict their various properties [13–19]. One of the most important tools in the chemical graph theory is a topological index, which is useful in predicting the chemical and physical properties of the underlying chemical compound, such as boiling point, strain energy, rigidity, heat of evaporation, and tension [20, 21]. A graph having no loop or multiple edge is known as a simple graph. A molecular graph is a simple graph in which atoms and bounds are represented by the vertex and edge sets, respectively. The degree of the vertex is the number of edges attached with that vertex. These properties of various objects is of primary interest. Winner, in 1947, introduced the concept of the first topological index while finding the boiling point. In 1975, Gutman gave a remarkable identity [22] about Zagreb indices. Hence, these two indices are among the oldest degree-based descriptors, and their properties are extensively investigated. The mathematical formulae of these indices are



$$M_1(G) = \sum_{uv \in E(G)} (d_u + d_v), \quad (1)$$

$$M_2(G) = \sum_{uv \in E(G)} (d_u \times d_v).$$

A topological index is known as the irregularity index [23] if the value of the topological index of the graph is greater than or equal to zero, and the topological index of the graph is equal to zero if and only if the graph is regular. The irregularity indices are given in Table 1. Most of the irregularity indices are from the family of degree-based topological indices and are used in quantitative structure activity relationship modeling.

For more about topological indices, one can read [24–31].

## 2. Irregularity Indices for Metal-Organic Networks

In this section, we discuss metal-organic networks by means of a graph. The unit cell of the metal-organic network is given in Figure 1. We give computational results of irregularity indices for two types of metal-organic networks  $MON_1(p)$  and  $MON_2(p)$  in the following two sections.

**2.1. Irregularity Indices for Metal-Organic Network  $MON_1(p)$ .** This section is about irregularity indices of the metal-organic network  $MON_1(p)$ . The molecular graph of the metal-organic network  $MON_1(p)$  for  $p = 2$  is given in Figure 2. We can observe from Figure 2 that there are four types of vertices present in the molecular graph of  $MON_1(p)$ , i.e., 2, 3, 4, and 6. The cardinality of vertices 2, 3, 4, and 6 is  $30p$ ,  $12$ ,  $12p - 6$ , and  $6p - 6$ , respectively. The cardinality of the vertex set of  $MON_1(p)$  is  $48p$ , i.e.,  $|V(MON_1(p))| = 48p$ . There are four different types of edges present in the molecular graph of  $MON_1(p)$ , i.e.,  $\{2, 3\}$ ,  $\{2, 4\}$ ,  $\{2, 6\}$ , and  $\{4, 6\}$ . Their cardinalities are  $36$ ,  $36p - 12$ ,  $24p - 24$ , and  $12p - 12$ , respectively. The cardinality of the edge set of  $MON_1(p)$  is  $72p - 12$ , i.e.,  $|E(MON_1(p))| = 72p - 12$ .

The edge partition of the metal-organic network  $MON_1(p)$  is given in Table 2.

**Theorem 1.** Let  $G$  be the metal-organic network  $E(MON_1(p))$ . The irregularity indices are

- (1)  $VAR(G) = (8p^2 - 5p - 1)/4p^2$
- (2)  $AL(G) = 192p - 108$
- (3)  $IR1(G) = 6(120p^2 - 80p - 17)/p$
- (4)  $IR2(G) = (\sqrt{(864p - 456)/48p}) - ((72p - 12)/24p)$
- (5)  $IRF(G) = 576p - 444$
- (6)  $IRFW(G) = 48p - 37/2(36p - 19)$
- (7)  $IRA(G) = 48p - 2(\sqrt{6} + 9\sqrt{2} + 4\sqrt{3})p - 10\sqrt{6} + 6\sqrt{2} + 8\sqrt{3}$

TABLE 1: Definitions of irregularity indices.

Irregularity index	Mathematical form
VAR	$\sum_{uv \in E(G)} (d_u - d_v)^2 = (M_1(G)/n) - (2m/n)^2$
AL	$\sum_{uv \in E(G)}  d_u - d_v $
IR1	$\sum_{uv \in E(G)} (d_u^3 + d_v^3) - (2m/n)\sum_{uv \in E(G)} (d_u + d_v) = F(G) - (2m/n)M_1(G)$
IR2	$\sqrt{\sum_{uv \in E(G)} d_u d_v / m} - (2m/n) = \sqrt{M_2(G)/m} - (2m/n)$
IRF	$\sum_{uv \in E(G)} (d_u - d_v)^2 = F(G) - 2M_2(G)$
IRFW	$IRF(G)/M_2(G)$
IRA	$\sum_{uv \in E(G)} (d_u^{-1/2} - d_v^{-1/2})^2 = n - 2R(G)$
IRB	$\sum_{uv \in E(G)} (d_u^{1/2} - d_v^{1/2})^2 = M_1(G) - 2RR(G)$
IRC	$(\sum_{uv \in E(G)} \sqrt{d_u d_v} / m) - (2m/n) = (RR(G)/m) - (2m/n)$
IRDIF	$\sum_{uv \in E(G)}  (d_u/d_v) - (d_v/d_u)  = \sum_{i < j} m_{i,j} ((j/i) - (i/j))$
IRL	$\sum_{uv \in E(G)}  \ln d_u - \ln d_v  = \sum_{i < j} m_{i,j} \ln(j/i)$
IRLU	$\sum_{uv \in E(G)}  d_u - d_v  / \min(d_u, d_v) = \sum_{i < j} m_{i,j} \ln((j - i)/i)$
IRLF	$\sum_{uv \in E(G)}  d_u - d_v  / \sqrt{d_u d_v} = \sum_{i < j} m_{i,j} ((j - i) / \sqrt{ij})$
IRLA	$2\sum_{uv \in E(G)}  d_u - d_v  / (d_u + d_v) = 2\sum_{i < j} m_{i,j} ((j - i) / (i + j))$
IRDI	$\sum_{uv \in E(G)} \ln  d_u - d_v  = \sum_{i < j} m_{i,j} \ln(i + j - 1)$
IRGA	$\sum_{uv \in E(G)} \ln((d_u + d_v) / 2\sqrt{d_u d_v}) = \sum_{i < j} m_{i,j} (i + j / 2\sqrt{ij})$

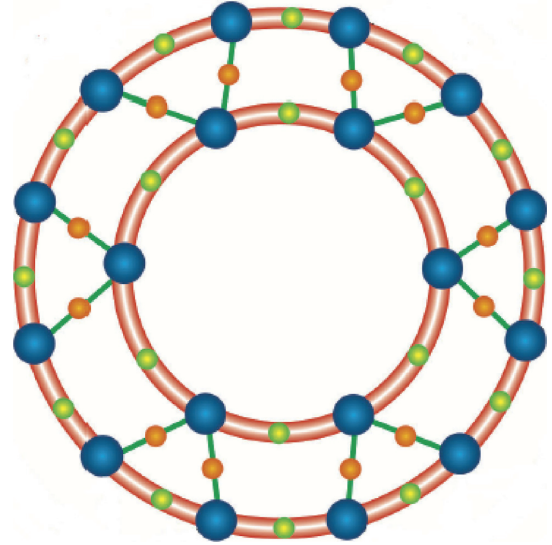


FIGURE 1: Basic metal-organic network.

- (8)  $IRB(G) = 528p - 204 - (2(24\sqrt{6} + 72\sqrt{2} + 48\sqrt{3}))p - 24\sqrt{6} + 48\sqrt{2} + 96\sqrt{3}$
- (9)  $IRC(G) = (4\sqrt{6}p^2 + 2\sqrt{6}p + 12\sqrt{2}p^2 + 8\sqrt{3}p^2 - 4\sqrt{2}p - 8\sqrt{3}p - 36p^2 + 12p - 1)/2p(6p - 1)$
- (10)  $IRDIF(G) = 127.9980p - 61.9992$
- (11)  $IRL(G) = 56.1840p - 24.9516$
- (12)  $IRLU(G) = 90p - 48$
- (13)  $IRLF(G) = 58.0668p - 26.4012$
- (14)  $IRLA(G) = 52.7976p - 22.3992$
- (15)  $IRDI(G) = 131.0064p - 42.4776$
- (16)  $IRGA(G) = 8.0172p - 5.8716$

*Proof.* Using definitions of irregularity indices given in Table 1 and edge partition given in Table 2, we have



$$\begin{aligned}\text{VAR}(G) &= \sum_{u \in V} \left( d_u - \frac{2m}{n} \right)^2 = \frac{M_1(G)}{n} - \left( \frac{2m}{n} \right)^2 \\ &= \left( \frac{528p - 204}{48p} \right) - \left( \frac{2(72p - 12)}{48p} \right)^2 \\ &= \frac{(8p^2 - 5p - 1)}{4p^2},\end{aligned}$$

$$\begin{aligned}\text{AL}(G) &= \sum_{uv \in E(G)} |d_u - d_v| \\ &= |2 - 3|(36) + |2 - 4|(36p - 12) + |2 - 6|(24p - 1) + |4 - 6|(12p - 1) \\ &= 192p - 108,\end{aligned}$$

$$\begin{aligned}\text{IR1}(G) &= \sum_{u \in V} d_u^3 - \frac{2m}{n} \sum_{u \in V} d_u^2 = F(G) - \left( \frac{2m}{n} \right) M_1(G) \\ &= (2304p - 1356) - \frac{2(72p - 12)}{48p} (528p - 204) \\ &= \frac{6(120p^2 - 80p - 17)}{p},\end{aligned}$$

$$\begin{aligned}\text{IR2}(G) &= \sqrt{\frac{\sum_{uv \in E(G)} d_u d_v}{m}} - \frac{2m}{n} = \sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n} \\ &= \sqrt{\frac{864p - 456}{72p - 12}} - \left( \frac{2(72p - 12)}{48p} \right) \\ &= \sqrt{\frac{864p - 456}{48p} - \frac{72p - 12}{24p}},\end{aligned}$$

$$\begin{aligned}\text{IRF}(G) &= \sum_{uv \in E(G)} (d_u - d_v)^2 \\ &= (2 - 3)^2(36) + (2 - 4)^2(36p - 12) + (2 - 6)^2(24p - 1) + (4 - 6)^2(12p - 1) \\ &= 576p - 444,\end{aligned}$$

$$\begin{aligned}\text{IRFW}(G) &= \frac{\text{IRF}(G)}{M_2(G)} \\ &= \frac{48p - 37}{2(36p - 19)},\end{aligned}$$

$$\begin{aligned}\text{IRA}(G) &= \sum_{uv \in E(G)} (d_u^{-1/2} - d_v^{-1/2})^2 = n - 2R(G) \\ &= (48p) - 2((9\sqrt{2} + \sqrt{3} + \sqrt{6})p - (3\sqrt{2} + 4\sqrt{3} - 5\sqrt{6})) \\ &= 48p - (2(\sqrt{6} + 9\sqrt{2} + 4\sqrt{3}))p - 10\sqrt{6} + 6\sqrt{2} + 8\sqrt{3},\end{aligned}$$

$$\begin{aligned}\text{IRB}(G) &= \sum_{uv \in E(G)} (d_u^{1/2} - d_v^{1/2})^2 = M_1(D_n P_n) - 2RR(G) \\ &= (528p - 204) - 2((72\sqrt{2} + 48\sqrt{3} + 24\sqrt{6})p - (24\sqrt{2} + 48\sqrt{3} - 12\sqrt{6}))\end{aligned}$$

$$\begin{aligned}
&= 528p - 204 - (2(24\sqrt{6} + 72\sqrt{2} + 48\sqrt{3}))p - 24\sqrt{6} + 48\sqrt{2} + 96\sqrt{3}, \\
\text{IRC}(G) &= \frac{\sum_{uv \in E(G)} \sqrt{d_u d_v} - 2m}{m} = \frac{\text{RR}(G) - 2m}{m} - \frac{2m}{n} \\
&= \frac{(72\sqrt{2} + 48\sqrt{3} + 24\sqrt{6})p - (24\sqrt{2} + 48\sqrt{3} - 12\sqrt{6})}{72p - 12} - \frac{2(72p - 12)}{48p} \\
&= \frac{4\sqrt{6}p^2 + 2\sqrt{6}p + 12\sqrt{2}p^2 + 8\sqrt{3}p^2 - 4\sqrt{2}p - 8\sqrt{3}p - 36p^2 + 12p - 1}{2p(6p - 1)}, \\
\text{IRDIF}(G) &= \sum_{uv \in E(G)} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| \\
&= \left| \frac{2}{3} - \frac{3}{2} \right| (36) + \left| \frac{2}{4} - \frac{4}{2} \right| (36p - 12) + \left| \frac{2}{6} - \frac{6}{2} \right| (24p - 1) + \left| \frac{4}{6} - \frac{6}{4} \right| (12p - 1) \\
&= 127.9980p - 61.9992, \\
\text{IRL}(G) &= \sum_{uv \in E(G)} |\ln d_u - \ln d_v| \\
&= |\ln 2 - \ln 3| (36) + |\ln 2 - \ln 4| (36p - 12) \\
&\quad + |\ln 2 - \ln 6| (24p - 1) + |\ln 4 - \ln 6| (12p - 1) \\
&= 56.1840p - 24.9516, \\
\text{IRLU}(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\min(d_u, d_v)} \\
&= \frac{|2 - 3|}{2} (36) + \frac{|2 - 4|}{2} (36p - 12) + \frac{|2 - 6|}{2} (24p - 1) + \frac{|4 - 6|}{4} (12p - 1) \\
&= 90p - 48, \\
\text{IRLF}(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\sqrt{d_u d_v}} \\
&= \frac{|2 - 3|}{\sqrt{6}} (36) + \frac{|2 - 4|}{\sqrt{8}} (36p - 12) + \frac{|2 - 6|}{\sqrt{12}} (24p - 1) + \frac{|4 - 6|}{\sqrt{24}} (12p - 1) \\
&= 58.0668p - 26.4012, \\
\text{IRLA}(G) &= \sum_{uv \in E(G)} 2 \frac{|d_u - d_v|}{(d_u + d_v)} \\
&= 2 \frac{|2 - 3|}{5} (36) + 2 \frac{|2 - 4|}{6} (36p - 12) + 2 \frac{|2 - 6|}{8} (24p - 1) + 2 \frac{|4 - 6|}{10} (12p - 1) \\
&= 52.7976p - 22.3992, \\
\text{IRD1}(G) &= \sum_{uv \in E(G)} \ln\{1 + |d_u - d_v|\} \\
&= \ln\{1 + |2 - 3|\} (36) + \ln\{1 + |2 - 4|\} (36p - 12) + \ln\{1 + |2 - 6|\} (24p - 1) + \ln\{1 + |4 - 6|\} (12p - 1) \\
&= 131.0064p - 42.4776, \\
\text{IRGA}(G) &= \sum_{uv \in E(G)} \ln\left(\frac{d_u + d_v}{2\sqrt{d_u d_v}}\right) \\
&= \ln\left(\frac{2 + 3}{2\sqrt{2 \times 3}}\right) (36) + \ln\left(\frac{2 + 4}{2\sqrt{2 \times 4}}\right) (36p - 1) + \ln\left(\frac{2 + 6}{2\sqrt{2 \times 6}}\right) (24p - 1) + \ln\left(\frac{4 + 6}{2\sqrt{4 \times 6}}\right) (12p - 1) \\
&= 8.0172p - 5.8716.
\end{aligned} \tag{2}$$

2.2. Irregularity Indices for Metal-Organic Network  $\text{MON}_2(p)$ . In this section, we will compute irregularity indices for the metal-organic network  $\text{MON}_2(p)$ . The

molecular graph of the metal-organic network  $\text{MON}_2(p)$  for  $p = 2$  is given in Figure 3. It is clear from Figure 3 that there are three types of vertices in the molecular graph of

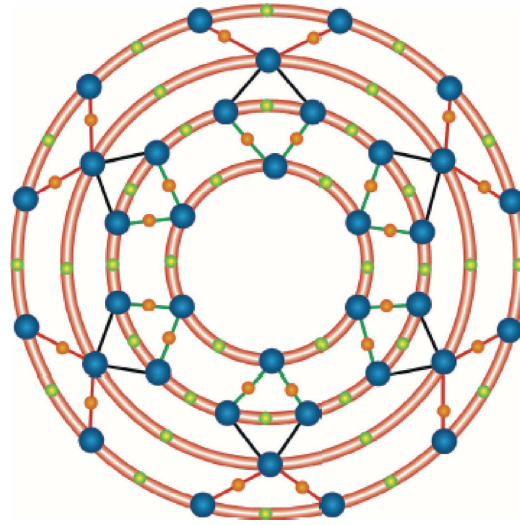


FIGURE 2:  $MON_1(p)$  for  $p = 2$ .

TABLE 2: Partition of  $E(MON_1(p))$ .

$(d_u, d_v)$	Frequency
(2, 3)	36
(2, 4)	$12(3p - 1)$
(2, 6)	$24(p - 1)$
(4, 6)	$12(p - 1)$

TABLE 3: Partition of  $E(MON_2(p))$ .

$(d_u, d_v)$	Frequency
(2, 3)	$12(p + 2)$
(2, 4)	$12(p + 1)$
(3, 3)	$24(p - 1)$
(3, 4)	$12(p - 1)$
(4, 4)	$12(p - 1)$

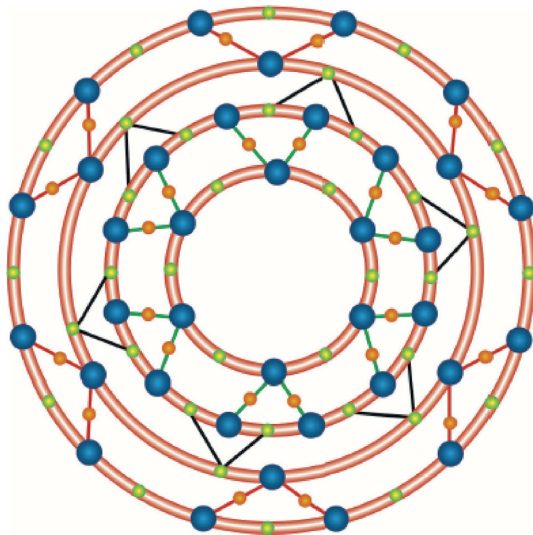


FIGURE 3:  $MON_2(p)$  for  $p = 2$ .

$MON_2(p)$ , i.e., 2, 3, and 4. The cardinality of vertices 2, 3, and 4 is  $12p + 18$ ,  $24p - 12$ , and  $12p - 6$ , respectively. The cardinality of the vertex set of  $MON_2(p)$  is  $48p$ , i.e.,  $|V(MON_2(p))| = 48p$ . There are five different types of edges present in the molecular graph of  $MON_2(p)$ , i.e.,  $\{2, 3\}$ ,  $\{2, 4\}$ ,  $\{3, 3\}$ ,  $\{3, 4\}$ , and  $\{4, 4\}$ . Their cardinalities are  $12p + 24$ ,  $12p + 12$ ,  $24p - 24$ ,  $12p - 12$ , and  $12p - 12$ , respectively. The cardinality of the edge set of  $MON_2(p)$  is  $72p - 12$ , i.e.,  $|E(MON_2(p))| = 72p - 12$ .

The edge partition of the metal-organic network  $MON_2(p)$  is given in Table 3.

**Theorem 2.**  $E(MON_2(p))$  Let  $G$  be the metal-organic network  $MON_2(p)$ . The irregularity indices are

- (1)  $VAR(G) = (2p^2 + p - 1)/4p^2$
- (2)  $AL(G) = 48p + 36$
- (3)  $IR1(G) = 6(24p^2 + 10p - 11)/p$
- (4)  $IR2(G) = \sqrt{(720p - 312)/48p} - ((72p - 12)/24p)$
- (5)  $IRF(G) = 72p + 60$
- (6)  $IRFW(G) = (72p + 60)/(720p - 312)$
- (7)  $IRA(G) = 32p + 16 - 2(2\sqrt{6} + 4\sqrt{3} + 3\sqrt{2})p - 8\sqrt{6} - 6\sqrt{2} + 8\sqrt{3}$
- (8)  $IRB(G) = 216p + 108 - (2(12\sqrt{6} + 24\sqrt{2} + 24\sqrt{3}))p - 48\sqrt{6} - 48\sqrt{2} + 48\sqrt{3}$
- (9)  $IRC(G) = (2\sqrt{6}p^2 + 4\sqrt{6}p + 4\sqrt{3}p^2 + 4\sqrt{2}p^2 - 4\sqrt{3}p + 4\sqrt{2}p - 16p^2 - 8p - 1)/2p(6p - 1)$
- (10)  $IRDIF(G) = 34.9992p + 30.9996$
- (11)  $IRL(G) = 16.8660p + 14.3664$
- (12)  $IRLU(G) = 21.9996p + 20.0004$
- (13)  $IRLF(G) = 16.8468p + 14.8188$
- (14)  $IRLA(G) = 16.2276p + 14.1708$
- (15)  $IRD1(G) = 29.8176p + 21.5004$
- (16)  $IRGA(G) = 1.0740p + 1.0716$

*Proof.* Using definitions of irregularity indices given in Table 1 and edge partition given in Table 3, we have

$$\begin{aligned}\text{VAR}(G) &= \sum_{u \in V} \left( d_u - \frac{2m}{n} \right)^2 = \frac{M_1(G)}{n} - \left( \frac{2m}{n} \right)^2 \\ &= \left( \frac{456p - 132}{48p} \right) - \left( \frac{2(72p - 12)}{48p} \right)^2 \\ &= \frac{(2p^2 + p - 1)}{4p^2},\end{aligned}$$

$$\begin{aligned}\text{AL}(G) &= \sum_{u \in E(G)} |d_u - d_v| \\ &= |2 - 3|(12p + 24) + |2 - 4|(12p + 12) + |3 - 3|(24p - 24) + |3 - 4|(12p - 12) + |4 - 4|(12p - 12) \\ &= 48p + 36,\end{aligned}$$

$$\begin{aligned}\text{IR1}(G) &= \sum_{u \in V} d_u^3 - \frac{2m}{n} \sum_{u \in V} d_u^2 = F(G) - \left( \frac{2m}{n} \right) M_1(G) \\ &= (1512p - 564) - \frac{2(72p - 12)}{48p} (456p - 132) \\ &= \frac{6(24p^2 + 10p - 11)}{p},\end{aligned}$$

$$\begin{aligned}\text{IR2}(G) &= \sqrt{\frac{\sum_{u \in E(G)} d_u d_v}{m}} - \frac{2m}{n} = \sqrt{\frac{M_2(G)}{m}} - \frac{2m}{n} \\ &= \sqrt{\frac{720p - 312}{72p - 12}} - \left( \frac{272p - 12}{48p} \right) \\ &= \sqrt{\frac{720p - 312}{48p}} - \frac{72p - 12}{24p},\end{aligned}$$

$$\begin{aligned}\text{IRF}(G) &= \sum_{u \in E(G)} (d_u - d_v)^2 \\ &= (2 - 3)^2(12p + 24) + (2 - 4)^2(12p + 12) + (3 - 3)^2(24p - 24) + (3 - 4)^2(12p - 12) + (4 - 4)^2(12p - 12) \\ &= 72p + 60,\end{aligned}$$

$$\begin{aligned}\text{IRFW}(G) &= \frac{\text{IRF}(G)}{M_2(G)} \\ &= \frac{72p + 60}{720p - 312},\end{aligned}$$

$$\begin{aligned}\text{IRA}(G) &= \sum_{u \in E(G)} (d_u^{-1/2} - d_v^{-1/2})^2 = n - 2R(G) = (48p) - 2(8p - 8 + (3\sqrt{2} + 4\sqrt{3} + 2\sqrt{6})p + (3\sqrt{2} - 4\sqrt{3} + 4\sqrt{6})) \\ &= 32p + 16 - 2(2\sqrt{6} + 4\sqrt{3} + 3\sqrt{2})p - 8\sqrt{6} - 6\sqrt{2} + 8\sqrt{3},\end{aligned}$$

$$\begin{aligned}\text{IRB}(G) &= \sum_{u \in E(G)} (d_u^{1/2} - d_v^{1/2})^2 = M_1(G) - 2RR(G) \\ &= (456p - 132) - 2(120p - 120 + (24\sqrt{2} + 24\sqrt{3} + 12\sqrt{6})p + (24\sqrt{6} - 24\sqrt{3} + 24\sqrt{2})) \\ &= 216p + 108 - (2(12\sqrt{6} + 24\sqrt{2} + 24\sqrt{3}))p - 48\sqrt{6} - 48\sqrt{2} + 48\sqrt{3},\end{aligned}$$

$$\begin{aligned}
\text{IRC}(G) &= \frac{\sum_{uv \in E(G)} \sqrt{d_u d_v}}{m} - \frac{2m}{n} = \frac{\text{RR}(G)}{m} - \frac{2m}{n} \\
&= \frac{120p - 120 + (24\sqrt{2} + 24\sqrt{3} + 12\sqrt{6})p + (24\sqrt{6} - 24\sqrt{3} + 24\sqrt{2})}{72p - 12} - \frac{2(72p - 12)}{48p} \\
&= \frac{2\sqrt{6}p^2 + 4\sqrt{6}p + 4\sqrt{3}p^2 + 4\sqrt{2}p^2 - 4\sqrt{3}p + 4\sqrt{2}p - 16p^2 - 8p - 1}{2p(6p - 1)},
\end{aligned}$$

$$\begin{aligned}
\text{IRDIF}(G) &= \sum_{uv \in E(G)} \left| \frac{d_u}{d_v} - \frac{d_v}{d_u} \right| \\
&= \left| \frac{2}{3} - \frac{3}{2} \right| (12p + 24) + \left| \frac{2}{4} - \frac{4}{2} \right| (12p + 12) + \left| \frac{3}{3} - \frac{3}{3} \right| (24p - 24) + \left| \frac{3}{4} - \frac{4}{3} \right| (12p - 12) + \left| \frac{4}{4} - \frac{4}{4} \right| (12p - 12), \\
&= 34.9992p + 30.9996,
\end{aligned}$$

$$\begin{aligned}
\text{IRL}(G) &= \sum_{uv \in E(G)} |\ln d_u - \ln d_v| \\
&= |\ln 2 - \ln 3| (12p + 24) + |\ln 2 - \ln 4| (12p + 12) + |\ln 3 - \ln 3| (24p - 24) + |\ln 3 - \ln 4| (12p - 12) \\
&\quad + |\ln 4 - \ln 4| (12p - 12) \\
&= 16.8660p + 14.3664,
\end{aligned}$$

$$\begin{aligned}
\text{IRLU}(G) &= \sum_{uv \in E(G)} \frac{|d_u - d_v|}{\min(d_u, d_v)} \\
&= \frac{|2 - 3|}{2} (12p + 24) + \frac{|2 - 4|}{2} (12p + 12) + \frac{|3 - 3|}{3} (24p - 24) + \frac{|3 - 4|}{3} (12p - 12) + \frac{|4 - 4|}{4} (12p - 12) \\
&= 21.9996p + 20.0004,
\end{aligned}$$

$$\begin{aligned}
\text{IRLF}(G) &= \sum_{(uv) \in EG} \frac{|d_u - d_v|}{\sqrt{d_u \cdot d_v}} \\
&= \frac{|2 - 3|}{\sqrt{6}} (12p + 24) + \frac{|2 - 4|}{\sqrt{8}} (12p + 12) + \frac{|3 - 3|}{\sqrt{9}} (24p - 24) + \frac{|3 - 4|}{\sqrt{12}} (12p - 12) + \frac{|4 - 4|}{\sqrt{16}} (12p - 12) \\
&= 16.8468p + 14.8188,
\end{aligned}$$

$$\begin{aligned}
\text{IRLA}(G) &= \sum_{uv \in E(G)} 2 \frac{|d_u - d_v|}{(d_u + d_v)} \\
&= 2 \frac{|2 - 3|}{5} (12p + 24) + 2 \frac{|2 - 4|}{6} (12p + 12) + 2 \frac{|3 - 3|}{6} (24p - 24) + 2 \frac{|3 - 4|}{7} (12p - 12) + 2 \frac{|4 - 4|}{8} (12p - 12) \\
&= 16.2276p + 14.1708,
\end{aligned}$$

$$\begin{aligned}
\text{IRD1}(G) &= \sum_{uv \in E(G)} \ln\{1 + |d_u - d_v|\} \\
&= \ln\{1 + |2 - 3|\} (12p + 24) + \ln\{1 + |2 - 4|\} (12p + 12) + \ln\{1 + |3 - 3|\} (24p - 24) + \ln\{1 + |3 - 4|\} (12p - 12) \\
&\quad + \ln\{1 + |4 - 4|\} (12p - 12) \\
&= 29.8176p + 21.5004,
\end{aligned}$$

$$\begin{aligned}
\text{IRGA}(G) &= \sum_{uv \in E(G)} \ln \left( \frac{d_u + d_v}{2\sqrt{d_u d_v}} \right) \\
&= \ln \left( \frac{2 + 3}{2\sqrt{2 \times 3}} \right) (12p + 24) + \ln \left( \frac{2 + 4}{2\sqrt{2 \times 4}} \right) (12p + 12) + \ln \left( \frac{3 + 3}{2\sqrt{3 \times 3}} \right) (24p - 24) + \ln \left( \frac{3 + 4}{2\sqrt{3 \times 4}} \right) (12p - 12) \\
&\quad + \ln \left( \frac{4 + 4}{2\sqrt{4 \times 4}} \right) (12p - 12) \\
&= 1.0740p + 1.0716.
\end{aligned}$$

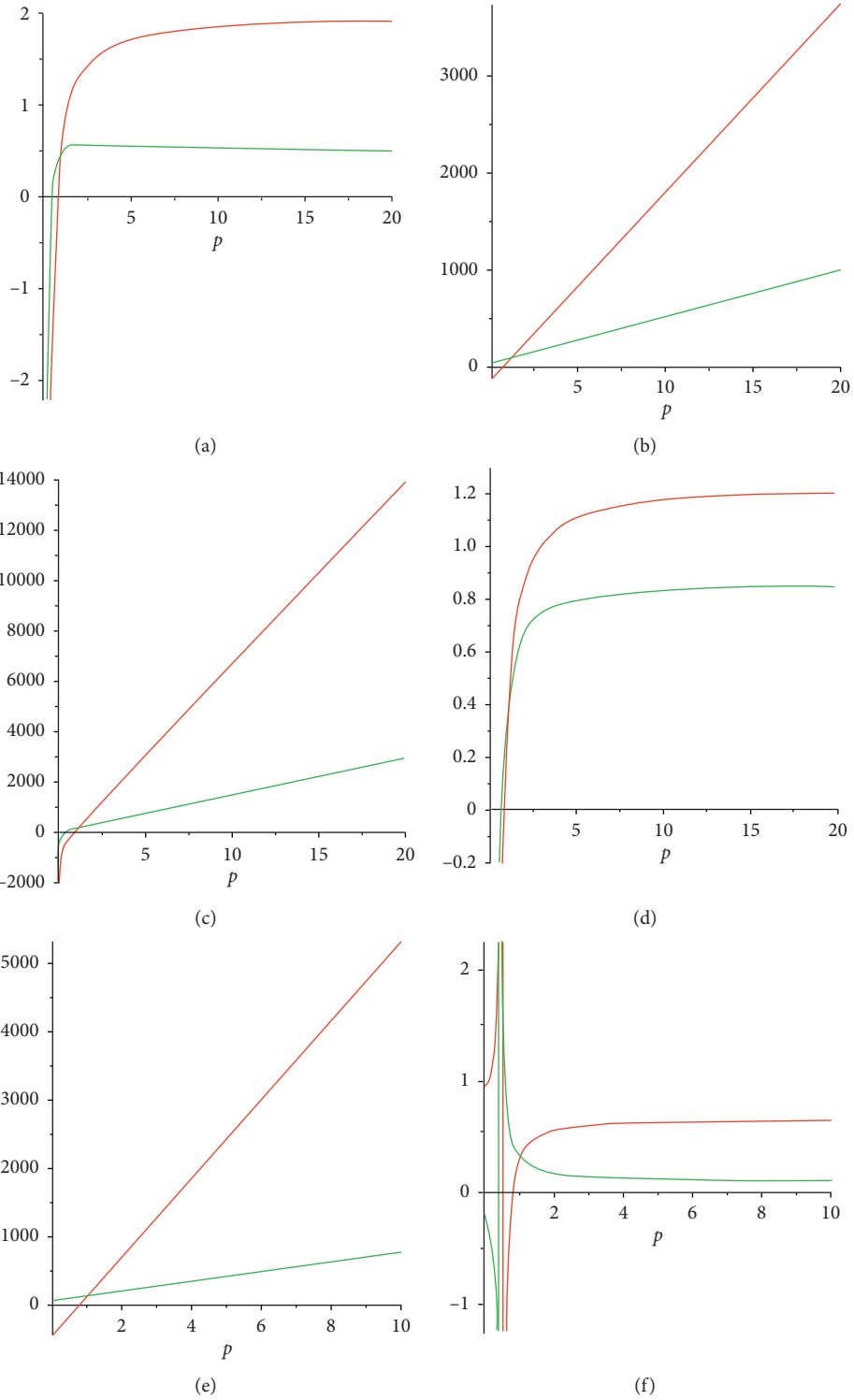
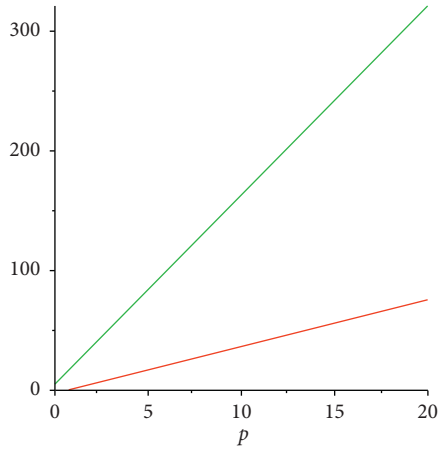
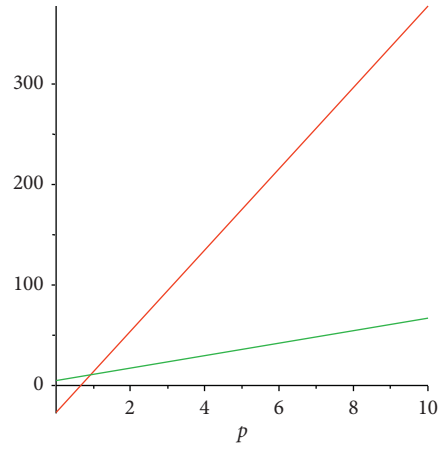


FIGURE 4: Continued.

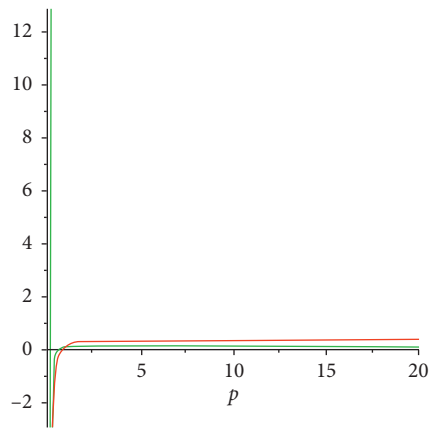




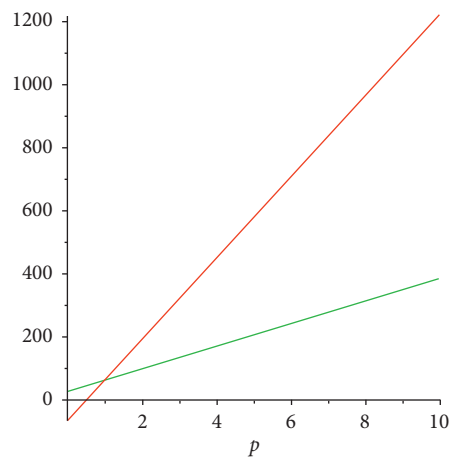
(g)



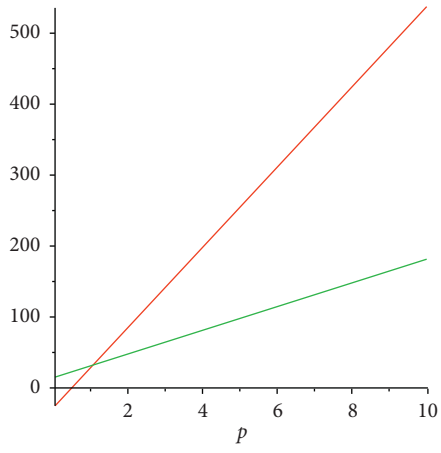
(h)



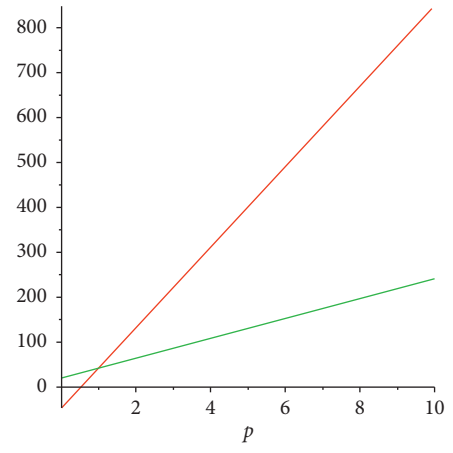
(i)



(j)



(k)



(l)

FIGURE 4: Continued.

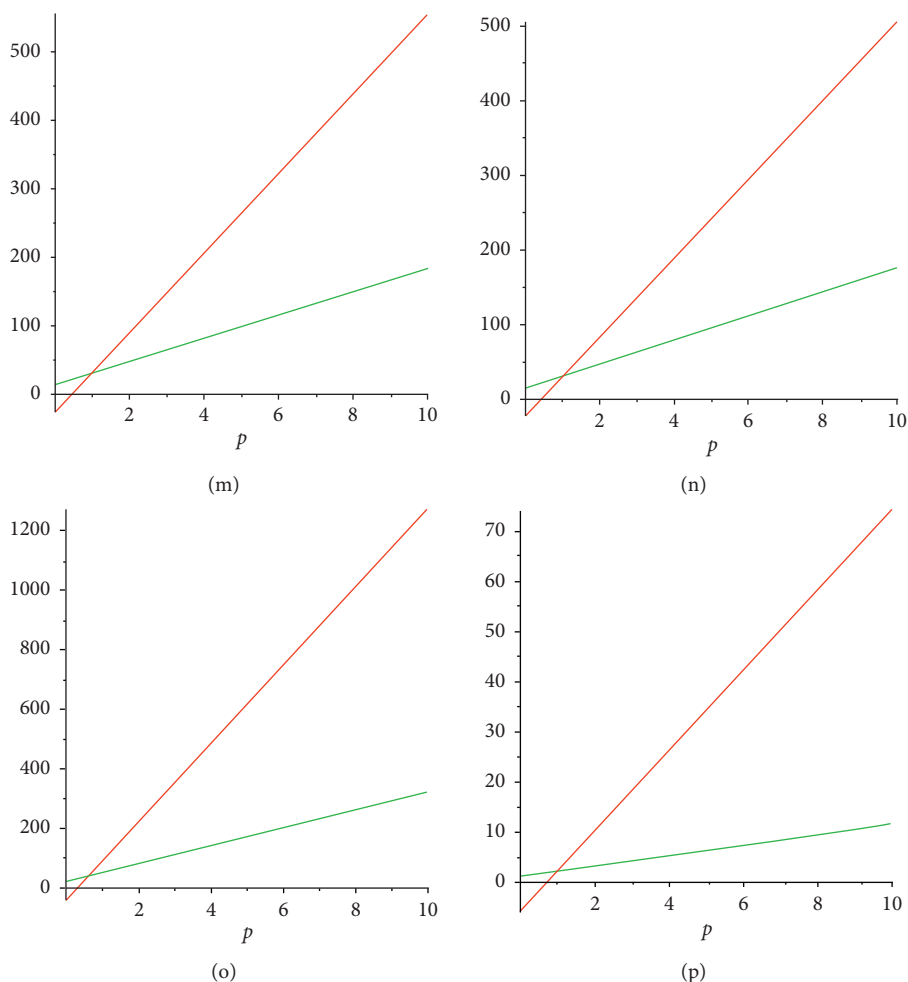


FIGURE 4: Graphical comparison of irregularity indices. (a) VAR. (b) AL. (c) IR1. (d) IR2. (e) IRF. (f) IRFW. (g) IRA. (h) IRB. (i) IRC. (j) IRDIF. (k) IRL. (l) IRLU. (m) IRLF. (n) IRLA. (o) IRD1. (p) IRGA.

### 3. Graphical Comparison

In this section, we give the graphical comparison of results of the metal-organic networks  $MON_1(p)$  and  $MON_2(p)$ . In Figure 4, the red color is fixed for  $MON_1(p)$  and the green color is fixed for  $MON_2(p)$ .

### 4. Conclusion

Topological indices associate a single number with a chemical structure. In quantitative structure activity relationship, knowledge of topological indices plays an important role. In this article, we computed sixteen irregularity indices for metal-organic networks  $MON_1(p)$  and  $MON_2(p)$ . Most of the calculated topological indices depend on degree-based indices. In the field of chemical graph theory, molecular topology, and mathematical chemistry, a degree-based index also known as a connectivity index is a type of a molecular descriptor that is calculated based on the molecular graph of a chemical compound. Topological indices are numerical parameters of a graph which characterize its topology and are usually graph invariant. Topological indices are used, for example, in the

development of quantitative structure activity relationships (QSARs) in which the biological activity or other properties of molecules are correlated with their chemical structure. Our results are helpful in drug delivery and computer engineering.

### Data Availability

All data required for this research are included within this paper.

### Conflicts of Interest

The authors do not have any conflicts of interest.

### Authors' Contributions

All authors contributed equally in this paper.

### Acknowledgments

The research was supported by the National Natural Science Foundation of China (Grant nos. 11971142, 11871202, 61673169, 11701176, 11626101, and 11601485).

## References

- [1] C. Petit and T. J. Bandoz, "MOF-graphite oxide composites: combining the uniqueness of graphene layers and metal-organic frameworks," *Advanced Materials*, vol. 21, no. 46, pp. 4753–4757, 2009.
- [2] S. J. Yang, J. Y. Choi, H. K. Chae, J. H. Cho, K. S. Nahm, and C. R. Park, "Preparation and enhanced hydrostability and hydrogen storage capacity of CNT@MOF-5 hybrid composite," *Chemistry of Materials*, vol. 21, no. 9, pp. 1893–1897, 2009.
- [3] X. Li and J. Zheng, "Extremal chemical trees with minimum or maximum general Randić index," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 55, no. 2, pp. 381–390, 2006.
- [4] M. C. Wasson, J. Lyu, T. Islamoglu, and O. K. Farha, "Linker competition within a metal-organic framework for topological insights," *Inorganic Chemistry*, vol. 58, no. 2, pp. 1513–1517, 2018.
- [5] I. Gutman, A. M. Naji, and N. D. Soner, "On leap Zagreb indices of graphs," *Communication in Combinatorics and Optimization*, vol. 2, no. 2, pp. 99–117, 2017.
- [6] J. Liu and Y. Lu, "Accelerated color change of gold nanoparticles assembled by DNAzymes for simple and fast colorimetric  $Pb^{2+}$  Detection," *Journal of the American Chemical Society*, vol. 126, no. 39, pp. 12298–12305, 2004.
- [7] J.-S. Jang, S. Qiao, S.-J. Choi et al., "Hollow Pd-Ag composite nanowires for fast responding and transparent hydrogen sensors," *ACS Applied Materials & Interfaces*, vol. 9, no. 45, pp. 39464–39474, 2017.
- [8] Y. K. Hwang, D.-Y. Hong, J.-S. Chang et al., "Amine grafting on coordinatively unsaturated metal centers of MOFs: consequences for catalysis and metal encapsulation," *Angewandte Chemie International Edition*, vol. 47, no. 22, pp. 4144–4148, 2008.
- [9] D. Bradshaw, A. Garai, and J. Huo, "Metal-organic framework growth at functional interfaces: thin films and composites for diverse applications," *Chemical Society Reviews*, vol. 41, no. 6, pp. 2344–2381, 2012.
- [10] A. W. Thornton, K. M. Nairn, J. M. Hill, A. J. Hill, and M. R. Hill, "Metal-organic frameworks impregnated with magnesium-decorated fullerenes for methane and hydrogen storage," *Journal of the American Chemical Society*, vol. 131, no. 30, pp. 10662–10669, 2009.
- [11] M. Azari and A. Iranmanesh, "Generalized zagreb index of graphs," *Studia Universitatis Babeş-Bolyai, Chimia*, vol. 56, no. 3, pp. 59–70, 2011.
- [12] A. Vasilyev, "Upper and lower bounds of symmetric division deg index," *Iranian Journal of Mathematical Chemistry*, vol. 5, no. 2, pp. 91–98, 2014.
- [13] N. Ali, M. A. Umar, and A. Tabassum, "Super  $(a, d)$ -C3-antimagicness of a corona graph," *Open Journal of Mathematical Sciences*, vol. 2, no. 1, pp. 371–378, 2018.
- [14] M. A. Umar, M. A. Javed, M. Hussain, and B. R. Ali, "Super  $(a, d)$ -C 4-antimagicness of book graphs," *Open Journal of Mathematical Sciences*, vol. 2, no. 1, pp. 115–121, 2018.
- [15] M. A. Umar, "Cyclic-antimagic construction of ladders," *Engineering and Applied Science Letters*, vol. 2, no. 2, pp. 43–47, 2019.
- [16] M. A. Umar, N. Ali, N. Ali, A. Tabassum, and B. R. Ali, "Book graphs are cycle antimagic," *Open Journal of Mathematical Sciences*, vol. 3, no. 1, pp. 184–190, 2019.
- [17] A. Tabassum, M. A. Lingua, M. A. Umar, M. Perveen, and A. Raheem, "Antimagicness of subdivided fans," *Open Journal of Mathematical Sciences*, vol. 4, no. 1, pp. 18–22, 2020.
- [18] F. Aslam, Z. Zahid, Z. Zahid, and S. Zafar, "3-total edge mean cordial labeling of some standard graphs," *Open Journal of Mathematical Sciences*, vol. 3(2019), no. 1, pp. 129–138, 2019.
- [19] F. Asif, Z. Zahid, and S. Zafar, "Leap Zagreb and leap hyper-Zagreb indices of Jahangir and Jahangir derived graphs," *Engineering and Applied Science Letter*, vol. 3, no. 2, pp. 1–8, 2020.
- [20] F. M. Brckler, T. Dolic, A. Graovac, and I. Gutman, "On a class of distance-based molecular structure descriptors," *Chemical Physics Letters*, vol. 503, no. 4–6, pp. 336–338, 2011.
- [21] H. Gonzalez-Diaz, S. Vilar, L. Santana, and E. Uriarte, "Medicinal chemistry and bioinformatics—current trends in drugs discovery with networks topological indices," *Current Topics in Medicinal Chemistry*, vol. 7, no. 10, pp. 1015–1029, 2007.
- [22] H. Hosoya, K. Hosoi, and I. Gutman, "A topological index for the total  $p$ -electron energy," *Theoretica Chimica Acta*, vol. 38, no. 1, 1975.
- [23] T. Réti, R. Sharafadini, A. Dregelyi-Kiss, and H. Haghbin, "Graph irregularity indices used as molecular descriptors in QSPR studies," *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 79, pp. 509–524, 2018.
- [24] E. Deutsch and S. Klavzar, "S. M-Polynomial and degree-based topological indices," *Iranian Journal of Mathematical Chemistry*, vol. 6, pp. 93–102, 2015.
- [25] M. Ajmal, W. Nazeer, M. Munir, S. M. Kang, and C. Y. Jung, "The M-polynomials and topological indices of generalized prism network," *International Journal of Mathematical Analysis*, vol. 11, no. 6, pp. 293–303, 2017.
- [26] M. Munir, W. Nazeer, Z. Shahzadi, and S. Kang, "Some invariants of circulant graphs," *Symmetry*, vol. 8, no. 11, p. 134, 2016.
- [27] H. Wiener, "Structural determination of paraffin boiling points," *Journal of the American Chemical Society*, vol. 69, no. 1, pp. 17–20, 1947.
- [28] A. A. Dobrynin, R. Entringer, and I. Gutman, "Wiener index of trees: theory and applications," *Acta Applicandae Mathematicae*, vol. 66, no. 3, pp. 211–249, 2001.
- [29] I. Gutman and O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer Science & Business Media, Berlin, Germany, 2012.
- [30] M. Randić, "Characterization of molecular branching," *Journal of the American Chemical Society*, vol. 97, no. 23, pp. 6609–6615, 1975.
- [31] B. Bollobás, P. Erdős, and A. Sarkar, "Graphs of extremal weights," *Ars Combinatoria*, vol. 50, pp. 225–233, 1998.
- [32] A. R. Virk, M. A. Rehman, and W. Nazeer, "New definition of atomic bond connectivity index to overcome deficiency of structure sensitivity and abruptness in existing definition," *Scientific Inquiry and Review*, vol. 3, no. 4, pp. 1–20, 2019.
- [33] J.-B. Liu, M. Younas, M. Habib, M. Yousaf, and W. Nazeer, "M-polynomials and degree-based topological indices of  $VC5C7 [p, q]$  and  $HC5C7 [p, q]$  nanotubes," *IEEE Access*, vol. 7, pp. 41125–41132, 2019.
- [34] Z. Shao, A. R. Virk, M. S. Javed, M. A. Rehman, and M. R. Farahani, "Degree based graph invariants for the molecular graph of Bismuth Tri-Iodide," *Engineering and Applied Science Letter*, vol. 2, no. 1, pp. 1–11, 2019.
- [35] A. u. R. Virk, M. N. Jhangeer, M. N. Jhangeer, and M. A. Rehman, "Reverse Zagreb and reverse hyper-zagreb indices for silicon carbide  $Si_2C_3 - I[r, s]$  and  $Si_2C_3 - II[r, s]$ ," *Engineering and Applied Science Letters*, vol. 1(2018), no. 2, pp. 37–50, 2018.

## Research Article

# Algebraic Connectivity and Disjoint Vertex Subsets of Graphs

Yan Sun<sup>1</sup> and Faxu Li<sup>2</sup> 

<sup>1</sup>School of Computer, Qinghai Nationalities University, Xining 810007, China

<sup>2</sup>School of Computer, Qinghai Normal University, Xining 810016, China

Correspondence should be addressed to Faxu Li; lifaxu@qhnu.edu.cn

Received 1 June 2020; Accepted 8 July 2020; Published 31 July 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Yan Sun and Faxu Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is well known that the algebraic connectivity of a graph is the second small eigenvalue of its Laplacian matrix. In this paper, we mainly research the relationships between the algebraic connectivity and the disjoint vertex subsets of graphs, which are presented through some upper bounds on algebraic connectivity.

## 1. Introduction

A graph  $G$  is often used to model a complex network. The vertex set and the edge set of graph  $G$  are denoted by  $\mathcal{N}$  and  $\mathcal{E}$ , respectively. A network is represented as an undirected graph  $G = (\mathcal{N}, \mathcal{E})$  consisting of  $N = |\mathcal{N}|$  nodes and  $E = |\mathcal{E}|$  links, respectively.

Graph theory has provided chemists with a variety of useful tools, such as in the topological structure [1–3]. The Laplacian matrix of a graph  $G$  is denoted by  $L$ , and  $L = D - A$ , where  $D$  is a diagonal matrix whose diagonal entries are its degrees and  $A$  is the adjacency matrix of  $G$ . The Laplacian eigenvalues of a graph  $G$  are the eigenvalues of  $L$ , denoted by  $0 = \mu_N \leq \mu_{N-1} \leq \dots \leq \mu_1$ , which are all real and nonnegative. The second smallest Laplacian eigenvalue  $\mu_{N-1}$  of a graph is well known as the algebraic connectivity, which was first studied by Fiedler [4]. The algebraic connectivity [5] of a graph is important for the connectivity of a graph [6], which can be used to measure the robustness of a graph. It has been emerged as an important parameter in many system problems [7–18]. Especially, the algebraic connectivity also plays an important role in the partitions of a complex network. For the literature on the algebraic connectivity of a graph [19], the reader is referred to [20, 21]. In this work, the relationships are researched between the algebraic connectivity and disjoint vertex subsets of graphs, which are presented through some upper bounds.

## 2. Preliminaries

Let  $x \in R^n$  be a vector. Let  $B$  be an incidence matrix of  $G$ . Then,  $x^T L x = \|B^T x\|_2^2 = \sum_{i,j \in E} (x_i - x_j)^2$ . For any vector  $x, y \in R^n$ , the inner product of  $x$  and  $y$  is defined as  $(x, y)$ . Two upper bounds on the algebraic connectivity are given as follows.

**Lemma 1** (see [20]). *For any vector  $f \in R^n$ , the Rayleigh inequality is as follows:*

$$\mu_{N-1} \leq \frac{(Lf, f)}{(f, f)}, \quad (1)$$

where  $(f, c) = 0$ ,  $c$  is a constant, and  $(Lf, f) = \sum_{v_i, v_j \in E} (f(v_i) - f(v_j))^2$ ,  $f(v_i)$  is the vector  $f$  for the node  $v_i$ .

**Lemma 2** (see [20]). *For any vector  $f \in R^n$ , we have*

$$\mu_{N-1} \leq \frac{\sum_{v_i, v_j \in E} (f(v_i) - f(v_j))^2}{\sum_{v_i \in \mathcal{N}} f^2(v_i)}, \quad (2)$$

$$\mu_{N-1} \leq \frac{\sum_{v_i, v_j \in \mathcal{E}} (f(v_i) - f(v_j))^2}{\sum_{v_i \in \mathcal{N}} \sum_{v_j \in \mathcal{N}} (f(v_i) - f(v_j))^2}, \quad (3)$$

where  $f(v_i)$  is the vector  $f$  for the node  $v_i$ .

Let  $A$  and  $B$  be two disjoint subsets of  $\mathcal{N}$ , respectively. The distance between two disjoint subsets  $A$  and  $B$  of  $\mathcal{N}$  is denoted by  $h(A, B)$ . For convenience,  $h$  takes the place of  $h(A, B)$ . Let  $h(u, A)$  be the distance between the node  $u$  and  $A$ , which is the shortest distance of the node  $u \in \mathcal{N}$  to a node of the set  $A$ . Suppose  $a = |A|/N$  and  $b = |B|/N$ . A result on the algebraic connectivity and two partitions of graphs is presented by Alon [22] and Milman [20] below.

**Lemma 3** (see [23]). *For any two disjoint subsets  $A$  and  $B$  of  $\mathcal{N}$ , it holds*

$$\mu_{N-1} \leq \frac{1}{Nh^2} \left( \frac{1}{a} + \frac{1}{b} \right) (E - E_A - E_B), \quad (4)$$

where  $E_A$  and  $E_B$  are the number of links in the sets  $A$  and  $B$ , respectively.

Moreover, the next step consider the case of three disjoint vertex subsets of graphs [24].

### 3. Main Result

Let  $A, B$ , and  $C$  be the subsets of  $\mathcal{N}$ , respectively, where their numbers of nodes are, respectively,  $|A|$ ,  $|B|$ , and  $|C|$ . Assume  $a = |A|/N$ ,  $b = |B|/N$ , and  $c = |C|/N$ . Let  $h(u, A), h(u, B)$ , and  $h(u, C)$  be the distances from the node  $u \in \mathcal{N}$  to subsets  $A, B$ , and  $C$  of  $\mathcal{N}$ , respectively. Suppose  $h_s = \min\{h(A, B), h(A, C), h(B, C)\}$ . Now, we construct a function  $g(u)$  related to node  $u$  as follows, where the constructed function is referred to the book [25]:

$$g(u) = \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \frac{1}{9} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \times \frac{\min\{h_s, h(u, A), h(u, B), h(u, C)\}}{h_s}. \quad (5)$$

Let  $f = g - \bar{g} \neq 0$ , where  $\bar{g} = 1/N \sum_{n \in \mathcal{N}} g(n)$ . It is easy to check that  $(f, c) = 0$ , where  $c$  is a constant function [26–30]. Meanwhile,  $(g, c) \neq 0$  can be checked. Thus, the following cases need to be discussed.

$$\sum_{u, v \in \mathcal{E}} (f(u) - f(v))^2 = \sum_{u, v \in \mathcal{E}} (g(u) - g(v))^2 = \sum_{u, v \in \mathcal{E} - (A \cup B \cup C)} (g(u) - g(v))^2 \leq \frac{1}{81h_s^2} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right)^2 (E - E_A - E_B - E_C). \quad (12)$$

From (2) and

$$\bar{g} = \frac{1}{N} \left( \sum_{n \in \mathcal{N}} g(n) \right), \quad \sum_{n \in \mathcal{N}} f^2(n) \geq \sum_{n \in (A \cup B \cup C)} (g(n) - \bar{g})^2, \quad (13)$$

*Case 1.* If the node  $u$  belongs to any one subset of  $\{A, B, C\}$ , then

$$0 = \min\{h_s, h(u, A), h(u, B), h(u, C)\}, \quad (6)$$

$$g(u) = \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) > 0. \quad (7)$$

*Case 2.* If the node  $u \in \mathcal{N} - \{A, B, C\}$ , then we can see that

$$\frac{\min\{h_s, h(u, A), h(u, B), h(u, C)\}}{h_s} \leq 1, \quad (8)$$

$$g(u) = \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \frac{1}{9} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \times \frac{\min\{h_s, h(u, A), h(u, B), h(u, C)\}}{h_s} > 0. \quad (9)$$

By Case 1 and Case 2,  $(g, c) \neq 0$  holds. In contrast, if  $(g, c) = 0$ , then  $g(u) = 0$  for each  $u \in \mathcal{N}$  and  $g - \bar{g} = 0$ , which is a contradiction with  $f = g - \bar{g} \neq 0$ . From the definition  $g(u)$ , for any two adjacent nodes  $u$  and  $v$ , we have

$$|g(u) - g(v)| \leq \frac{1}{9h_s} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right). \quad (10)$$

Our main result is as follows.

**Theorem 1.** *Let  $A, B$ , and  $C \in \mathcal{N}$  be three disjoint subsets of  $\mathcal{N}$ . Let  $E_A$  and  $E_B$  and  $E_C$  be the numbers of links in the sets  $A$  and  $B$  and  $C$ , respectively. Then,*

$$\begin{aligned} \mu_{N-1} &\leq \frac{1}{81N^2h_s^2} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) (E - E_A - E_B - E_C) \\ &= \frac{1}{81Nh_s^2} \left( \frac{1}{N_A} + \frac{1}{N_B} + \frac{1}{N_C} \right) (E - E_A - E_B - E_C). \end{aligned} \quad (11)$$

*Proof.* For subsets  $A, B$ , and  $C$ , by Lemma 2, we have

where  $\bar{g} = 0$ , and since the coordinates of the center of gravity of the three regions are the average of the triangle region, then the vectors  $g(u, A) + g(u, B) + g(u, C) = 0$ . The sum of the vectors of the center of gravity of the triangle to the vertices is equal to 0 [25]. The center of gravity is analogous to the mean or average from statistics [6, 31, 32]:

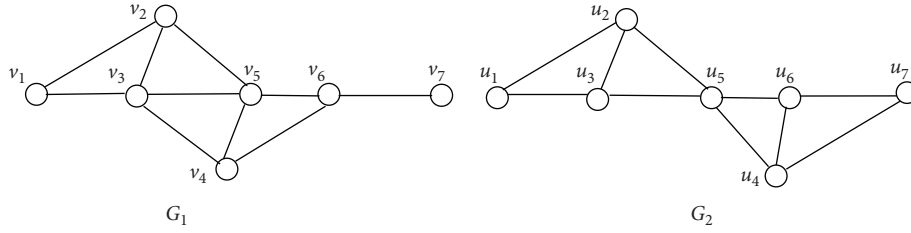


FIGURE 1: The graphs  $G_1$  and  $G_2$ .

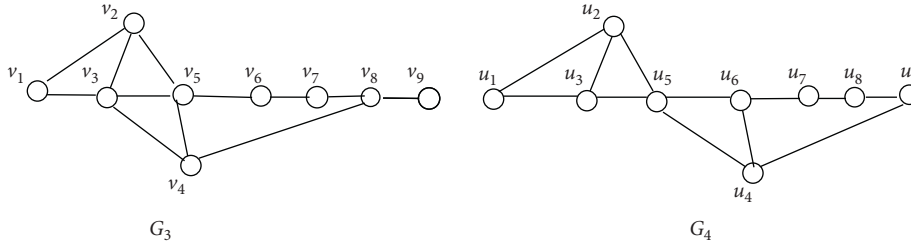


FIGURE 2: The graphs  $G_3$  and  $G_4$ .

$$\begin{aligned}
 \sum_{n \in \mathcal{N}} f^2(n) &\geq \sum_{n \in A \cup B \cup C} f^2(n) = \sum_{n \in A} (g(n) - \bar{g})^2 + \sum_{n \in B} (g(n) - \bar{g})^2 + \sum_{n \in C} (g(n) - \bar{g})^2 \\
 &= |A| \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right)^2 + |B| \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right)^2 + |C| \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right)^2 \\
 &= |A| \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \right)^2 + |B| \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \right)^2 + |C| \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \right)^2 \\
 &= (|A| + |B| + |C|) \left( \frac{1}{3} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \right)^2 \\
 &= \frac{1}{9} \left( |A| \frac{1}{a} + |B| \frac{1}{b} + |C| \frac{1}{c} \right)^2 \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \\
 &= N^2 \frac{1}{9} (1 + 1 + 1)^2 \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \geq N^2 \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right).
 \end{aligned} \tag{14}$$

By the above inequalities and Lemma 2, it arrives that

$$\begin{aligned}
 \mu_{N-1} &\leq \frac{1}{81N^2h_s^2} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) (E - E_A - E_B - E_C) \\
 &= \frac{1}{81Nh_s^2} \left( \frac{1}{N_C} + \frac{1}{N_B} + \frac{1}{N_C} \right) (E - E_A - E_B - E_C).
 \end{aligned} \tag{15}$$

*Example 1.* Figure 1 describes the graphs  $G_1$  and  $G_2$ , each with  $N = 7$  nodes,  $L = 10$  links, and a diameter  $\rho = 4$ . For  $G_1$  subsets,  $A = \{v_1, v_3, v_5, v_6, v_7\}$ ,  $B = \{v_2\}$ , and  $C = \{v_4\}$ . For  $G_2$

subsets,  $A = \{u_1, u_2, u_3\}$ ,  $B = \{u_4\}$ ,  $C = \{u_5, u_6, u_7\}$ , and  $h_s = 0.5$ . Their algebraic connectivity [33] and their upper bounds on (11) are as follows. For the  $G_1$  and  $G_2$  aplacian matrixes,



$$\begin{aligned}
 G_1 &= \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 \\ 1 & -1 & 4 & -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 & 0 \\ 0 & -1 & -1 & -1 & 4 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}, \\
 G_2 &= \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 \\ -1 & -1 & 3 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 3 & -1 & -1 & -1 \\ 0 & -1 & -1 & -1 & 4 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & 0 & -1 & 2 \end{bmatrix}.
 \end{aligned} \tag{16}$$

For  $G_1$ , the algebraic connectivity is 0.6338, and the algebraic connectivity of  $G_2$  is 0.5858. For  $G_1$  upper bounds on  $\mu_{N-1}(G_1) \leq (1/81N^2h_s^2)(1/a + 1/b + 1/c)(E - E_A - E_B - E_C) = 0.0931$ .

For  $G_2$  upper bounds on  $\mu_{N-1}(G_2) \leq (1/81N^2h_s^2)(1/a + 1/b + 1/c)(E - E_A - E_B - E_C) = 0.0587$ .

A graph with the second smallest Laplacian eigenvalue  $\mu_{N-1}$  is thus more robust, in the sense of being better connected.

**Proposition 1.** Let  $A, B,$  and  $C \in V$  be three disjoint subsets of  $V$ . Suppose  $h_s = 1$  and  $D = V - A - B - C$ . Let  $m_A, m_B, m_C,$  and  $m_D$  be the number of links in the sets  $A, B, C,$  and  $D,$  respectively. Then,

$$\mu_{N-1} \leq \frac{(1/81)((1/a) + (1/b) + (1/c))^2 (m - m_A - m_B - m_C - m_D)}{(|A| + |B| + |C|)[(1/a) + (1/b) + (1/c) - \bar{g}]^2 + |D|[2/9((1/a) + (1/b) + (1/c)) - \bar{g}]^2}, \tag{17}$$

where

$$\bar{g} = \frac{1}{n} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \left( |A| + |B| + |C| + \frac{2}{9}|D| \right), \tag{18}$$

in which  $\bar{g}$  is the average of the  $A, B, C, D$  field.

*Proof.* For subsets  $A, B,$  and  $C,$  by Lemma 2, we have

$$\sum_{u,v \in E} (f(u) - f(v))^2 = \sum_{u,v \in E} (g(u) - g(v))^2 = \sum_{u,v \in E - (E_A \cup E_B \cup E_C \cup E_D)} (g(u) - g(v))^2 \leq \frac{1}{81} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right)^2 (m - m_A - m_B - m_C - m_D), \tag{19}$$

where links of the sets in the node sets  $A, B, C,$  and  $D$  are  $E_A, E_B, E_C,$  and  $E_D,$  respectively. From (2), we obtain

$$\begin{aligned}
 \sum_{n \in \mathcal{V}} f^2(n) &= \sum_{n \in \{A \cup B \cup C \cup D\}} f^2(n) = \sum_{n \in A} (g(n) - \bar{g})^2 \\
 &\quad + \sum_{n \in B} (g(n) - \bar{g})^2 + \sum_{n \in C} (g(n) - \bar{g})^2 + \sum_{n \in D} (g(n) - \bar{g})^2 \\
 &= |A| \left[ \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right]^2 + |B| \left[ \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right]^2 + |C| \left[ \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right]^2 \\
 &\quad + |D| \left[ \frac{2}{9} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right]^2 \\
 &= (|A| + |B| + |C|) \left[ \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right]^2 + |D| \left[ \frac{2}{9} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) - \bar{g} \right]^2.
 \end{aligned} \tag{20}$$

By direct computation, we have

$$\bar{g} = \frac{1}{n} \left( \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \left( |A| + |B| + |C| + \frac{2}{9} |D| \right). \quad (21)$$

By the above equalities and Lemma 2, inequality (17) holds.

But, we note that the algebraic connectivity [34, 35],  $\mu_{N-1}$ , should not be seen as a strict disconnection or a robustness metric [36].

*Example 2.* For example, Figure 2 describes the graphs  $G_3$  and  $G_4$ , with  $n = 9$ ,  $m = 12$ , and diameter 6. By direct calculation, for  $G_3$  subsets,  $A = \{v_1 v_2 v_3\}$ ,  $B = \{v_4 v_5 v_6\}$ ,  $C = \{v_7\}$ , and  $D = \{v_8 v_9\}$ , and for  $G_4$  subsets,  $A = \{u_1 u_2 u_3\}$ ,  $B = \{u_4 u_5 u_6\}$ ,  $C = \{u_7\}$ , and  $D = \{u_8 u_9\}$ . Their algebraic connectivity  $G_3$  is 0.4798 and  $G_4$  is 0.4817, respectively. Their Laplacian matrices  $L(G_3)$  and  $L(G_4)$ , for  $G_3$  upper bounds on  $\mu_{N-1}(G_3) \geq 0.431$  and for  $G_4$  upper bounds on  $\mu_{N-1}(G_4) \geq 0.357$ .

Theorem 1 and Proposition 1 are two completely different situations. The theorem hypothesis is that  $A, B$ , and  $C \in \mathcal{N}$  be three disjoint subsets of  $\mathcal{N}$ . The proposition supposes that  $A, B$ , and  $C \in V$  be three disjoint subsets of  $V$  and  $h_s = 1$  and  $D = V - A - B - C$ . In other words, the proposition has constraints. Moreover, it is not the same as the four disjoint subsets of  $\mathcal{N}$ .

## Data Availability

All data, models, and codes generated or used during the study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the Chunhui project of the Ministry of Education, China (no. Z2017046) and the Qinghai Science and Technology Planning Project (Grant no. 2018-ZJ-718).

## References

- [1] R. García-Domenech, J. Gálvez, and L. Pogliani, "Some new trends in chemical graph theory," *Chemical Reviews*, vol. 108, no. 3, pp. 1127–1169, 2008.
- [2] J.-B. de Julián-Ortiz, C. Wang, S. Wang, and B. Wei, "Zagreb indices and multiplicative zagreb indices of eulerian graphs," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 42, no. 1, pp. 67–78, 2019.
- [3] J.-B. Liu, J. Zhao, J. Min, and J. D. Cao, "On the Hosoya index of graphs formed by a fractal graph," *Fractals-Complex Geometry Patterns and Scaling in Nature and Society*, vol. 27, no. 8, pp. 105–135, 2019.
- [4] M. Fiedler, "Algebraic connectivity of graphs, Czech," *Journal of Mathematics*, vol. 23, pp. 298–305, 1973.
- [5] N. L. Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge, UK, 2nd edition, 1993.
- [6] J.-B. Liu, J. Zhao, and Z.-Q. Cai, "On the generalized adjacency, Laplacian and signless Laplacian spectra of the weighted edge corona networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 540, p. 123073, 2020.
- [7] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley, Boston, MA, USA, 1999.
- [8] P. Baldi, P. Frascioni, and P. Smyth, *Modeling the Internet and the Web: Probabilistic Methods and Algorithms*, John Wiley & Sons, Hoboken, NJ, USA, 2003.
- [9] S. Barnett, *Matrices: Methods and Applications*, Oxford University Press, Oxford, UK, 1992.
- [10] T. F. Chan, P. Ciarlet, and W. K. Szeto, "On the optimality of the median cut spectral bisection graph partitioning method," *SIAM Journal on Scientific Computing*, vol. 18, no. 3, pp. 943–948, 1997.
- [11] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari, "The electrical resistance of a graph captures its commute and cover times," in *Proceedings of the twenty-first annual ACM symposium on Theory of computing—STOC '89*, pp. 574–586, Seattle WA, USA, May 1989.
- [12] P. Chebotarev and E. Shamis, "The matrix-forest theorem and measuring relations in small social groups," *Automation and Remote Control*, vol. 58, no. 9, pp. 1505–1514, 1997.
- [13] P. Chebotarev and E. Shamis, "On a duality between metrics and s-proximities," *Automation and Remote Control*, vol. 59, no. 4, pp. 608–612, 1998.
- [14] K.-W. Cheung, K.-C. Tsui, and J. Liu, "Extended latent class models for collaborative recommendation," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, vol. 34, no. 1, pp. 143–148, 2004.
- [15] D. Harel and Y. Koren, "On clustering using random walks," *FST TCS 2001: Foundations of Software Technology and Theoretical Computer Science*, vol. 22, no. 45, pp. 18–41, 2001.
- [16] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 5–53, 2004.
- [17] N.-D. Ho and P. Van Dooren, "On the pseudo-inverse of the laplacian of a bipartite graph," *Applied Mathematics Letters*, vol. 18, no. 8, pp. 917–922, 2005.
- [18] Z. Huang, H. Chen, and D. Zeng, "Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering," *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 116–142, 2004.
- [19] R. Brooks, "Spectral geometry and the Cheeger constant," in *Expanding Graphs*, J. Friedman, Ed., American Mathematical Society, Providence, RI, USA, pp. 5–19, 1993.
- [20] P. Van Mieghem, *Graph Spectra for Complex Networks*, Delft University of Technology, Delft, Netherlands, 2012.
- [21] N. M. M. de Abreu, "Old and new results on algebraic connectivity of graphs," *Linear Algebra and Its Applications*, vol. 423, no. 1, pp. 53–73, 2007.
- [22] N. Alon, "Eigenvalues and expanders," *Combinatorica*, vol. 6, no. 2, pp. 83–96, 1986.
- [23] N. Alon and V. D. Milman, " $\lambda_1$ , Isoperimetric inequalities for graphs, and superconcentrators," *Journal of Combinatorial Theory, Series B*, vol. 38, no. 1, pp. 73–88, 1985.
- [24] P. K. Chan, M. Schlag, and J. Zien, "Spectral k-way ratio cut partitioning and clustering," in *Proceedings of the Symposium on Research on integrated systems*, Warwick, UK, 1993.

- [25] A Court, Nathan College Geometry, *An Introduction to the Modern Geometry of the Triangle and the Circle*, Barnes & Noble, New York, NY, USA, 2nd edition, 1925.
- [26] F. Chung, *Spectral Graph Theory*, AMS Publications, Providence, RI, USA, 1997.
- [27] F. R. K. Chung, "Laplacians of graphs and Cheeger inequalities," in *Combinatorics, Paul Erdo is Eighty*, pp. 157–172, János Bolyai Mathematical Society, Budapest, Hungary, 1996.
- [28] P. Buser, *Geometry and Spectra of Compact Riemann Surfaces*, Birkhauser, Basel, Switzerland, 1992.
- [29] J. Cheeger, "A lower bound for the smallest eigenvalue of the Laplacian," in *Problems in Analysis*, R. C. Gunnig, Ed., pp. 195–199, Princeton University Press, Princeton, NJ, USA, 1970.
- [30] C. Delorme and S. Poljak, "Laplacian eigenvalues and the maximum cut problem," *Mathematical Programming*, vol. 62, no. 1-3, pp. 557–574, 1993.
- [31] J.-B. Liu, J. Zhao, H. He, and Z. Shao, "Valency-based topological descriptors and structural property of the generalized sierpiński networks," *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.
- [32] J.-B. Liu, Z.-Yu Shi, Y.-H. Pan, J. Cao, and M. Abdel-Aty, "Udai Al-juboori computing the laplacian spectrum of linear octagonal-quadrilateral networks and its applications," *Polycyclic Aromatic Compounds*, vol. 10, pp. 23–43, 2020.
- [33] D. M. Cvetkovic, M. Doob, I. Gutman et al., *Recent Results in the Theory of Graph Spectra, Annals of Discrete Mathematics*, Vol. 36, Elsevier, Amsterdam, Netherlands, 1988.
- [34] C. Godsil and G. Royle, "Algebraic graph theory," in *Graduate Texts in Mathematics*, p. 207, Springer-Verlag, New York, NY, USA, 2001.
- [35] S. J. Kirkland, J. J. Moliterno, M. Neumann, and B. L. Shader, "On graphs with equal algebraic and vertex connectivity," *Linear Algebra and Its Applications*, vol. 341, no. 1-3, pp. 45–56, 2002.
- [36] G. A. Edgar, D. H. Ullman, and D. B. West, "Problems and solutions," *The American Mathematical Monthly*, vol. 125, no. 1, pp. 81–89, 2018.

## Research Article

# Risk Factors Discovery for Cancer Survivability Analysis Using Graph-Rule Mining

Chaoyu Yang <sup>1</sup>, Jie Yang,<sup>2</sup> and Zhenyu Yang<sup>3</sup>

<sup>1</sup>*School of Economics and Management, Anhui University of Science and Technology, Huainan 232001, China*

<sup>2</sup>*School of Computing and Information Technology, Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, NSW 2522, Australia*

<sup>3</sup>*School of Computer Science and Engineering, The University of New South Wales, Sydney, NSW 2052, Australia*

Correspondence should be addressed to Chaoyu Yang; [yangchy@aust.edu.cn](mailto:yangchy@aust.edu.cn)

Received 20 May 2020; Accepted 13 July 2020; Published 31 July 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Chaoyu Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mining and understanding patients' disease-development pattern is a major healthcare need. A huge number of research studies have focused on medical resource allocation, survivability prediction, risk management of diagnosis, etc. In this article, we are specifically interested in discovering risk factors for patients with high probability of developing cancers. We propose a systematic and data-driven algorithm and build around the idea of association rule mining. More precisely, the rule-mining method is firstly applied on the target dataset to unpack the underlying relationship of cancer-risk factors, via generating a set of candidate rules. Later, this set is represented as a rule graph, where informative rules are identified and selected with the aim of enhancing the result interpretability. Compared to hundreds of rules generated from the standard rule-mining approach, the proposed algorithm benefits from a concise rule subset, without losing the information from the original rule set. The proposed algorithm is then evaluated using one of the largest cancer data resources. We found that our method outperforms existing approaches in terms of identifying informative rules and requires affordable computational time. Additionally, relevant information from the selected rules can also be used to inform health providers and authorities for cancer-risk management.

## 1. Introduction

Recent years have witnessed a significantly increasing amount of electronic health records (EHR), in addition to other data collected for the diagnosis and management purpose. The Surveillance, Epidemiology, and End Results (SEER) resource is one of the typical examples. As one comprehensive and authoritative resource in relation to cancer statistics, SEER is a publicly available dataset originating from the United State. This data repository aims to provide high-quality and comprehensive cancer information, in order to help institutions and laboratories worldwide performing their own research. As such, this SEER dataset has been used for a diverse range of research applications, which results in more than 1500 copies for the public use annually. In addition, this SEER repository has also been evolving and updated from time to time, by the increment of

new patient samples, the inclusion of more medical features/variables, the involvement of new types of cancers, etc.

Not surprisingly, numerous machine learning methods have been applied to the SEER dataset for monitoring patient status and facilitating a better understanding of cancer treatment and survivability. Prior research efforts include Expert Systems [1], Fuzzy Systems [2, 3], Evolutionary Computation [4, 5], Support Vector Machines [6], and Neural Networks and/or Deep Learning [7]. Yet, there are still open research questions remaining. Expert/Fuzzy Systems, for instance, are typically reliant on human knowledge to determine (semi)static decision strategies. Intuitively, the a priori knowledge may vary from experts to experts, thereby resulting in significantly different outcomes. In addition, knowledge/expertise acquisition could be very time-consuming and labor-expensive, particularly when the scale/dimension of the given problem is large. On the other hand,

Support Vector Machines and Neural Network based approaches are limited in their interpretability, and relevant results are always questionable for end users.

Alternatively, we consider adopting the association rule-mining (ARM) algorithm in this study. As one of the most popular data-mining algorithms, ARM is characterized by its capability of being data-driven (less dependency on the external knowledge, compared to Expert Systems) and interpretability (high transparency compared to Neural Networks). As such, ARM has attracted much research attention with its wide application in many areas, such as the analysis for smart-phone app usage [8], opinion leadership identification [9], and monitoring patients' disease-development behavior [10]. In addition, ARM-based applications in the medical domain can be found in the preliminary research [11, 12].

Yet, one major problem with ARM is the huge number of generated rules; that is, a typical result from ARM could be hundreds and thousands rules associated with different lengths. On the other hand, many rules are overlapping and/or repeating each other with minor changes, which leads to the issue of the rule redundancy. Obviously, a large number of rules is difficult to exam or interpret manually, not to mention its computational overhead, while applying a small set of rules may not be sufficient to capture the underlying pattern, due to the possibility of lacking of information. Consequently, how to control/manipulate the number of generated rules to accurately describe the given dataset becomes a critical process for any ARM-based applications.

Traditionally, there are two strategies in terms of optimizing generated rules: (i) the application of a priori domain knowledge and (ii) rule summarization technique. The former one usually works with predetermined conditions to filter rules, which relies on external resources, such as expert experience or domain knowledge. In this context, only certain items are permitted to be included in generated rules, while others will be cast as unnecessary items to remove. Intuitively, this strategy has two major drawbacks: firstly, identifying important items is time-consuming, particularly with the large number of available items; secondly, experts could impose their own bias via determining item importance, thereby resulting in questionable rules.

On the other hand, rule summarization is a data-driven and automate method, in which less domain knowledge is required. The basic concept of the summarization technique is to identify important rules automatically, from the entire rule set, without compromising the information loss. Some existing work has been reviewed in Section 2. Inspired by the general applicability of rule summarization, this paper explores the task of discovering patients' pattern using the association rule summarization method. To enhance the summarization capability, we further introduce a cluster-based strategy to identify important rules. More specifically, the proposed algorithm consists of three parts. To begin with, we establish a rule graph based on their similarity, in which rules are grouped into different clusters using the community-detection method. Eventually, significant rules are determined and selected across individual rules, which are cast as the output of the proposed summarization. To the

best of our knowledge, this is the first study of proposing a cluster-based rule summarization algorithm to reveal the relationship among cancer-related risk factors.

The remainder of the paper is organized as follows. Section 2 provides the brief review about the ground work, such as data-mining-based medical applications; we also discuss traditional techniques for the rule summarization and community-detection clustering approaches. Section 3 presents the proposed cluster-based summarization algorithm, where three major phases are discussed in terms of similarity graph construction, community-detection-based cluster, and applied summarization strategy within each individual cluster. Our proposed framework is then evaluated in Section 4 using the SEER dataset to explore patient risk factors, followed by concluding remarks in Section 5.

## 2. Related Work

This section offers a brief discussion of the state-of-the-art research work related to the analysis of patients' pattern. At first, we investigate the application of data-mining algorithms in the medical domain. We then discuss the basic concept behind association rule-mining and summarization methods. Finally, we focus on the clustering approach for community detection.

*2.1. Data-Mining-Based Medical Application.* Recent years have witnessed a vast amount of applications of data-mining techniques in the medical domain [1, 3, 6, 7, 13]. In [1], an expert system was proposed by integrating geographic information and Online Analytical Processing (OLAP) technologies to facilitate environmental health decision support. More precisely, this expert system aimed to investigate potential relations between health problems and environmental risk factors, such as neighborhood, industrial pollutants, and drinking water quality. Another research [6] was proposed to apply a number of supervised learning techniques to discover lung cancer patients in terms of their survivability. Experimental results suggested that the Gradient Boosting Machine led to the best prediction performance, while Support Vector Machine was the only model that generated a distinctive output. In addition, the work [7] investigated the combination of Neural Networks with adversarial domain adaptation. A couple of scenarios were considered in the experiment for the evaluation purpose, including the standard supervised classification, unsupervised domain adaptation, and supervised domain adaptation. Resultant outcome indicated that the hybrid model of Neural Networks and adversarial domain based adaptation achieved satisfactory performance to measure pathology reports. More recently, a Bidirectional Long Short-Term Memory Neural Network (BLSTM-NN) was employed to build an interaction monitoring system in [13]. In their study, ten volunteers were involved and their activities were recorded using a set of Kinect sensors. Then 3D skeletons from participants were detected and tracked using BLSTM-NN, which revealed the underlying activity pattern and interaction among patients. A more general survey was

represented [3] to discuss the various methodologies, such as Fuzzy Logic, Neural Networks, and Genetic Algorithm, and their various applications in medicine.

The majority of existing systems, however, generally are characterized as expert-defined or black-box style. For instance, Expert and Fuzzy Systems require the domain knowledge to setup prediction strategies, which could be very labor-expensive. Neural Network based approaches, on the other hand, are usually limited by their interpretability, which remain questionable for end users. To sum up, despite the general interest of applying data-mining techniques in the medial domain, discovering patient risk factors is still a difficult task. As an alternative, this paper explores the potential of applying association rule-mining-based methods. In particular, rule-based approaches benefit from their transparency, interpretability, and efficient computation, which have potential to overcome the aforementioned limitations from other approaches.

**2.2. Rules Mining and Summarization.** Association rules mining (ARM) is one of the most-common data-mining algorithms for the relationship analysis. Its goal is to extract rules of the form “IF-Then,” such that if a set of variable values is found, then another set of variables will generally have a specific value. A typical example from patients’ rule can be “AGE\_DX(1), MAR\_STAT(1)  $\rightarrow$  SEQ\_NUM(0), SRV (> 60),” which indicates if this patient is diagnosed less than 53 years old (i.e., AGE\_DX(1)) and is single (i.e., MAR\_STAT(1)), then to some extent she/he will have one primary only in the lifetime (i.e., SEQ\_NUM(0)) and survival months is more than 60 months (i.e., SRV (> 60)).

As such, the technique is very useful in terms of associating an immediate subsequence (i.e., consequent) given the previous condition (i.e., antecedent) and discovering patterns of interaction among different factors. On the other hand, the importance of a rule is usually estimated through critical indicators, such as “support” and “confidence.” Mathematically, given a rule of ( $\mathcal{A} \rightarrow \mathcal{C}$ ), its support is the proportion of records which contain all items from  $\mathcal{A}$ , which can be computed as follows:

$$\text{supp}(\mathcal{A} \rightarrow \mathcal{C}) = \text{supp}(\mathcal{A}) = \frac{|\mathcal{A}|}{N}, \quad (1)$$

where  $|\mathcal{A}|$  is the number of records containing  $\mathcal{A}$  and  $N$  is the total number of rules. The confidence of the rule  $\mathcal{A} \rightarrow \mathcal{C}$  is accordingly computed as

$$\text{conf}(\mathcal{A} \rightarrow \mathcal{C}) = \frac{\text{supp}(\mathcal{A}) \cup \text{supp}(\mathcal{C})}{\text{supp}(\mathcal{A})}. \quad (2)$$

Consequently, the “support” indicator is used to measure the extent to which the antecedent and consequent occurs simultaneously, while the “confidence” indicator estimates how often the consequent occurs given the antecedent.

Due to its high interpretability and efficiency, plenty of ARM-based applications have been applied for the analysis of smart-phone app usage [8], opinion leadership identification [9, 14], and monitoring patients’ disease-development behavior [10], to name a few. In their pilot work of [8],

the authors aimed to investigate how students use their smart-phone apps to support online learning. App data from 148 schools were collected accordingly, and the  $K$ -means algorithm was employed to separate students into five groups based on their app usage. By mining pattern rules from each cluster, results suggested that students’ online patterns showed a shifting ratio between educational and noneducational apps. In addition, generated rules also revealed unique emphases on different types of apps that could impact on student learning performance. The work in [14], on the other hand, investigated a niche subset of user-generated popular culture content on Douban, a well-known Chinese-language online social network. Built on a dataset comprised of 714,946 comments and 228,806 distinct users, a parallel rule-mining algorithm was proposed. The experimental results demonstrated the flexibility and applicability of the rule-based method for extracting useful relationship from complex social media data. In addition, another work to explore patient’s behavior in terms of disease complications and recurrences was reported in [10]. For this particular research, a database about colorectal cancer, with 1516 patients and 126 attributes, was considered. At its core, four heuristic operators and a complete methodology were proposed to implement the rule-mining process. From the experiments, the rule-based approach showed advantages over standard approaches, such as the associative classification methods to identify risk factors.

The major problem, however, with the traditional ARM is the huge number of generated rules, which is manually inefficient to exam them one-by-one. The large number of rules, on the other hand, also reduces the interpretability as a whole. To overcome this problem, one established approach is the rule summarization, i.e., to summarize rules based on their significance without degrading the relationship information from the expression of the entire rule set. There are a few implementations for summarizing important rules, including APRX-COLLECTION [15] and RPGlobal [16]. To begin with, APRX-COLLECTION introduces a concept of the false positive rate ( $\alpha$ ), which is used to control the level of wrong coverage. More precisely, APRX-COLLECTION firstly forms an aggregated rule set  $\mathcal{R}^*$  by enumerating all possible combinations of original rules  $\mathcal{R}$ . As such, additional rules, even though they might not exist in  $\mathcal{R}$ , could be created. Next, rules from  $\mathcal{R}^*$  are selected according to two criteria: (i) those rules cover most items from  $\mathcal{R}$  and (ii) the number of additional items should be less than the level of  $\alpha$ . On the other hand, RPGlobal adopts similar selection criteria by identifying rules that cover most items from  $\mathcal{R}$ . The major difference is that RPGlobal chooses rules from  $\mathcal{R}$  directly, instead of generating  $\mathcal{R}^*$ . In addition, to limit the increment of additional rules, RPGlobal further introduces a user-defined parameter  $\beta$  to control the number of selected rules each time.

Overall, rule-mining summarization techniques have been proposed to overcome the problem associated with the huge number of generated rules. The summarized rules make the subsequent interpretation process more efficient and easier, by filtering least-important rules and significantly reducing the scale of rules. Inspired by this insight, an



enhanced rule summarization method is proposed in this study, which is in light of clustering, in particular, the community-detection technique.

**2.3. Community Detection.** As a graph clustering technique, community detection has attracted a lot of attention in the past decade due to the increasing scale of social network. With the rapid growth of Internet infrastructure, more and more people utilize online resources in their daily life, such as Twitter and Facebook. The result is the generation of a huge social network, in which individual users play a role of nodes/vertices and their connections (e.g., friendship) become the edge in the network. As such, either the industry stakeholders or academia are interested in analyzing this giant social network to formulate better marketing and/or development strategy. In particular, identifying communities within complex networks is of great importance for many real scenarios. A typical example could be forming an online community in relation to a group of people who share the same interest.

A number of different methods have been proposed to implement the community detection. A pioneer work in [17] was proposed based on the concept of edge betweenness centrality (EBC). For each individual edge within the network, its EBC was measured by the total number of shortest paths (for any two vertices) passing this particular edge. As a result, an edge with higher EBC became a good indicator to separate among communities, while an edge with lower EBC was more likely to exist within a small community. By removing edges with high EBC, the entire network eventually could be split into small groups/communities. Another work was reported in [18] with a similar measurement, that is, edge clustering coefficient (ECC). This measurement was to count the number of triangles for a given edge, compared to the total number of such possible triangles. Compared to EBC, edges with low ECC were considered as the connections among communities. As such, disjoint subnetworks can be formed by eliminating those low-ECC edges.

In addition to edge-based measurements, the Walktrap algorithm from [19] considered the topological similarity between vertices. The main idea was to divide the network based on a distance between vertices such that distance in the same community was small but became larger in different groups. This vertex distance was formally defined by (i) the walking probability from one vertex to another and (ii) the vertex degree. Another vertex-based algorithm was proposed in [20], termed as Label Propagation. To begin with, every vertex was randomly initialized with a unique label. Later, during the iteration, each vertex adjusted its label based on neighbors; that is, new label will be made the same as its majority labels nearby. Finally, communities were formed by grouping vertices with the same labels. The Infomap algorithm, on the other hand, was proposed using the concept of random walks and information diffusion [21]. It started by performing a random walk within the network and calculated the information flows using the trajectory of the previous random walk. An information map was accordingly established, which differentiated communities

with a diverse range of map importance. One advantage with this Infomap algorithm is its nearly linear-computational time, thereby leading to a very efficient process.

**2.4. Summary.** In this section, we briefly discuss the existing work of applying data-mining techniques to address medical problems. We also review the basic concept of rule-mining, rules summarization, and the community-detection approaches. Although preliminary work has been conducted to identify rules in relation to patients risk management, traditional rule-mining algorithms suffer from a major problem associated with the huge number of generated rules. To cope with this issue, rules summarization techniques offer advantage to select important rules and minimize the information loss. Taking all these aspects into account, we propose an enhanced summarization algorithm using the community-detection approaches, which is detailed in the following section.

### 3. The Proposed Framework

In this section, we discuss a systematic and data-driven approach to discover risk-relevant factors. The main contribution of this study is the proposal of a novel cluster-based summarization algorithm. As illustrated in Figure 1, the proposed approach consists of three phases. To begin with, we apply the traditional rule-mining algorithm on the entire dataset to generate a comprehensive set of potential rules. Due to the large scale of this rule set, we then represent it as a rule-similarity graph; see Section 3.1. Secondly, the community-detection algorithm is employed to identify clusters from this rule graph; see Section 3.2. Finally, informative rules across clusters are summarized, as introduced in Section 3.3.

For convenience, Table 1 summarizes notations used throughout the paper.

**3.1. Similarity Graph.** The main purpose of this first phase is to generate a completed rule set that represents the entire transaction records and then to construct a rule-similarity graph. Towards this end, there are several steps we need to consider, including data discretization, rule-mining, similarity measurement, and graph construction.

**3.1.1. Data Discretization.** To begin with, the rule-mining algorithm works well with discrete data, rather than the continuous ones. However, in the real-world scenario, the majority of medical data is continuous and not operable by the rule-mining approaches. To quantify extracted features, a preprocess of data discretization is necessary. For simplicity, this study aims to split a continuous input into  $L$  groups (where  $L$  is a user-defined threshold). As such, samples belonging to the same group will be assigned with the same label, to convert the continuous data into discrete one. Note that a domain knowledge is required to decide the number of groups (i.e.,  $L$ ), while different business or operational requirements could result in a variety of discretization ranges.

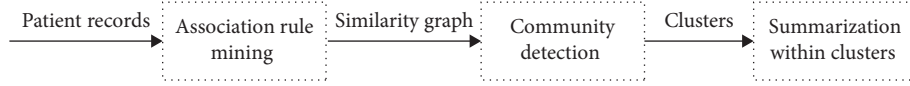


FIGURE 1: The pipeline of the proposed rule-summarization algorithm, including three phases: (i) applying association rule-mining algorithm to obtain the completed rule set, which is turned into the rule-similarity graph; (ii) employing the community-detection algorithm to cluster the rule graph; (iii) summarizing significant rules from individual clusters.

TABLE 1: List of adopted notations used in our study.

Notation	Description
$\mathcal{F}$	The complete itemset from the original data
$N$	Number of transaction records
$M$	Number of rules generated from these $N$ samples
$K$	Number of clusters to be found within the rule graph
$N_s$	Number of rules to be summarized/selected from one cluster

However, the advantage of the discretization is twofold: (i) continuous data is represented using discrete labels to facilitate the subsequent application of the rule-mining algorithm; (ii) the raw continuous dataset is represented by a smaller-sized but meaningful format, which is easy to be interpreted and also saves the computational cost.

**3.1.2. Rule Mining.** There exists a diverse range of implementations for rule mining, such as Apriori and FP-Growth. In particular, Apriori employs a “bottom-up” strategy to produce frequent-item sets, in which repeated scanning of the entire dataset is required. This typically leads to an expensive computational cost. Therefore, in this study, the FP-Growth algorithm is implemented, which adopts a “top-down” strategy to produce frequent-item sets. The main advantage is that it requires less scanning time to generate possible combinations of frequent sets.

**3.1.3. Similarity Measurement.** Before we construct the rule-similarity graph, it is essential to define the similarity measurement for any given rules. Consider the typical form of two rules  $r_1: (\mathcal{A}_1 \rightarrow \mathcal{C}_1)$  and  $r_2: (\mathcal{A}_2 \rightarrow \mathcal{C}_2)$ . The similarity function between two rules is accordingly defined in terms of the relative item coverage (RIC):

$$\text{RIC}(r_1, r_2) = \frac{(\|\mathcal{A}_1 \cap \mathcal{A}_2\|) \cup (\|\mathcal{C}_1 \cap \mathcal{C}_2\|)}{\|\mathcal{A}_1 \cup \mathcal{C}_1 \cup \mathcal{A}_2 \cup \mathcal{C}_2\|}. \quad (3)$$

As observed, the similarity is measured as the portion of the common items from both antecedents and consequences versus the portion of all items occurring within two rules. We then introduce the process of rule-graph construction based on the similarity measurement in equation (3).

**3.1.4. Graph Construction.** Graph is a very important data structure in computer science, while a large number of existing works have been proposed to demonstrate the solid application and success of graph-based techniques [2, 4]. Inspired by this insight, we also consider representing rules as a graph format. As such, the rule graph is represented as  $G = (V, E)$  in our study, where each vertex  $v_i$  ( $v_i \in V$ )

TABLE 2: Comparison of different community-detection implementations in terms of their computational complexity.

Algorithm	Cost	Reference
Edge betweenness centrality	$\mathcal{O}(m^2n)$	[17]
Edge clustering coefficient	$\mathcal{O}(m^4/n^2)$	[18]
Walktrap	$\mathcal{O}(n^2 \log(n))$	[19]
Label propagation	$\mathcal{O}(m+n)$	[20]
Infomap	$\mathcal{O}(m)$	[21]

Note that  $n$  and  $m$  represent the number of vertexes and edges, respectively.

denotes a rule  $r_i$ , and the edge  $e_{ij}$  is the connection between the  $i$ -th and  $j$ -th vertex. Furthermore,  $e_{ij}$  is associated with the similarity between the rule of  $r_i$  and  $r_j$ , i.e.,  $\text{RIC}(r_i, r_j)$ . Note that there are a diverse range of options to manipulate  $e_{ij}$ . For instance, we can introduce a user-defined threshold  $\epsilon$  to filter  $\text{RIC}(r_i, r_j)$  if  $\text{RIC}(r_i, r_j) < \epsilon$ . That is, two vertices are only connected if their similarity  $\text{RIC}(r_i, r_j)$  is larger than the value of  $\epsilon$ . Alternatively, we can also employ the concept of  $k$ -nearest neighbors, from which only the most similar  $(k-1)$  vertices to one specific vertex are connected. Without loss of generality, we consider the full-connect strategy; that is, all vertices will be connected to each other, while the edge  $e_{ij}$  equals their similarity  $\text{RIC}(r_i, r_j)$ .

**3.2. Rule Clustering.** This second phase intends to apply the community-detection algorithm to identify clusters from the rule-similarity graph. The general idea is to cluster vertices in a way that samples belonging to the same group are similar, while samples from different groups are dissimilar to each other. As mentioned in Section 3.3, there has been a great interest in developing implementations for detecting communities within the graph. Table 2 summarizes the computational complexity of the existing implementations for the community detection.

As observed, a variety of implementations may lead to different costs, as they focus on different optimization strategies on splitting vertices and/or edges. Considering our case of rule-based graph, there will be over  $10^4$  rules, indicating the final graph is with  $10^4$  vertices and approximately  $10^8$  edges. As such, we implement the work [21] in this study as our community-detection executor, due to its efficiency and affordable computation.

**3.3. Cluster-Based Summarization.** The final phase is used to perform rule summarization by determining or selecting important rules within each individual cluster. More precisely, after forming clusters within the rule graph, rules will be ranked according to their importance. As such, top- $N$  rules from each single cluster will be selected according to

**Input:**  $K$  clusters ( $C_1, C_2, \dots, C_K$ ), the number of selected rules  $N_s$ , and the penalty terms of  $\lambda_L, \lambda_A$ , and  $\lambda_R$ ;  
**Initialization:**  
 set the rule set  $S$  to empty:  $S = \emptyset$   
**for**  $k = 1$  **to**  $K$  **do**  
   **While** ( $\|C_k\| \neq 0$  or  $\|S\| \leq N_s$ ) **do**  
     Calculate the score of  $r_i$  ( $\forall r_i \in C_k$ ) according to equation (7)  
     Select one rule with the highest score and label it as  $r^*$   
      $S \leftarrow r^*$  and remove  $r^*$  from  $C_k$   
   **end**  
**end**  
**Output:** Return the optimal rules from  $S$ .

ALGORITHM 1: The proposed cluster-based algorithm for rule summarization.

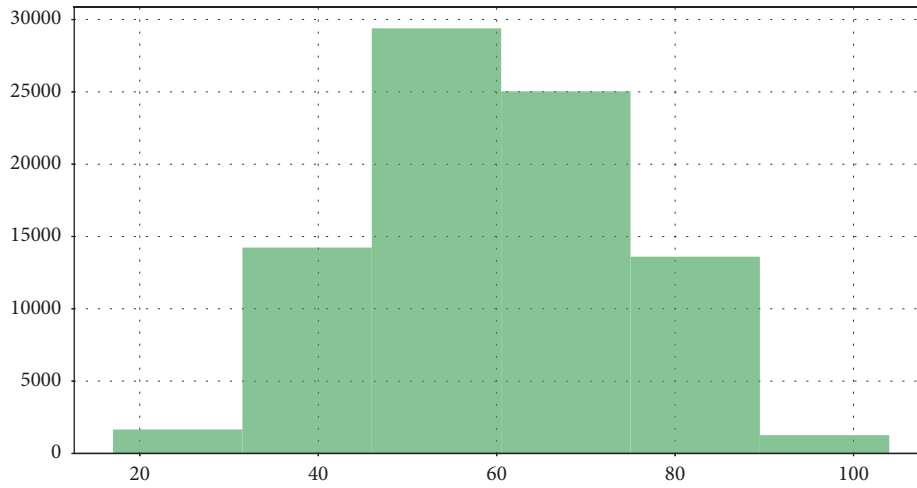


FIGURE 2: Data distribution of AGE\_DX.

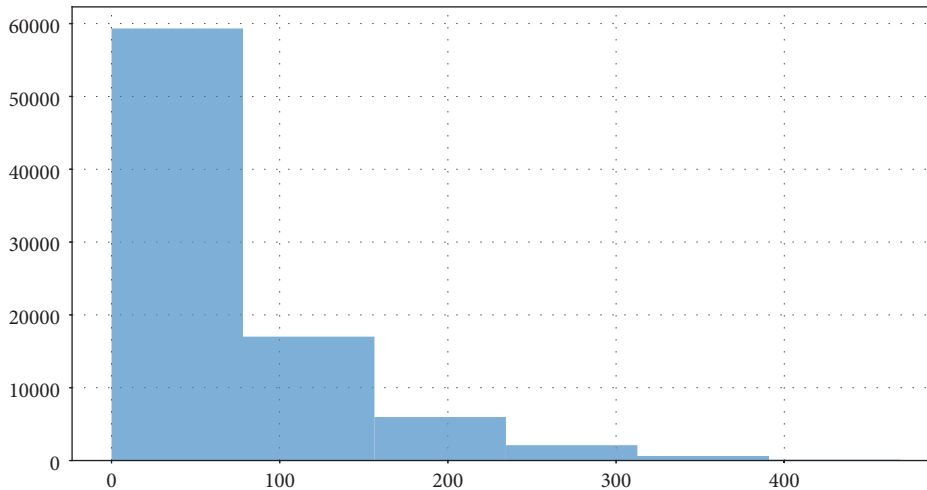


FIGURE 3: Data distribution of SRV\_TIME.

their importance, where  $N$  is a user-determined parameter, to produce the final summary. To begin with, we propose three statistical features before aggregating them to measure the rule importance:

- (i) Rule length: the first importance measurement is the rule length. This idea is inspired by the work of APRX-COLLECTION in [15], where a long rule has a better item coverage compared to shorter ones. Therefore, we use the rule length as a feature to

estimate the importance. Given a rule of  $(r_i: \mathcal{A} \rightarrow \mathcal{E})$ , the feature of length ( $\text{Length}(r_i)$ ) is computed as follows:

$$\text{Length}(r_i) = \frac{1}{\|\mathcal{S}\|} (\|\mathcal{A}\| + \|\mathcal{E}\|), \quad (4)$$

where  $\|\mathcal{S}\|$  represents the total number of distinct items from all rules.

- (ii) Aggregated similarity: we want to select informative rules that would represent the majority rules within a cluster. As such, rules should be selected if they are with a similar or close form to the rest. Therefore, the second measurement to the rule importance is the aggregated similarity. More specifically, the larger the aggregated similarity of a rule, the higher the rank. Given a cluster  $C$ , the aggregate similarity ( $\text{Aggregate}(r_i)$ ) for the  $i$ -th rule ( $r_i \in C$ ) is then given as follows:

$$\text{Aggregate}(r_i) = \frac{\sum_{j=1, j \neq i}^{\|C\|} \text{RIC}(r_i, r_j)}{\|C\|}, \quad (5)$$

where  $\text{RIC}(r_i, r_j)$  is the similarity measurement from (3) and  $\|C\|$  represents the total number of rules from the  $C$  cluster.

- (iii) Redundancy: another critical aspect is to consider the redundancy impact once a rule is selected. Note that each cluster is composed of similar rules. Therefore, if two rules contain similar items, then these two are more likely to convey the similar meaning. The redundancy feature is then employed to eliminate those semantically similar rules, without repeating the same rules all the time. In our study, the following estimation is proposed to measure the redundancy feature ( $\text{Redundancy}(r_i)$ ):

$$\text{Redundancy}(r_i) = \frac{\sum_{j=1}^{\|\mathcal{S}\|} \text{RIC}(r_i, r_j)}{\|\mathcal{S}\|}, \quad \text{if } \|\mathcal{S}\| \neq 0, \quad (6)$$

where  $\mathcal{S}$  is the set composed of selected rules and  $\|\mathcal{S}\|$  represents the number of rules from  $\mathcal{S}$ . Note that, at the beginning of the summarization,  $\mathcal{S} = \emptyset$  and  $\|\mathcal{S}\| = 0$ . In this case,  $\text{Redundancy}(r_i) = 0$  ( $\forall r_i \in \mathcal{S}$ ).

Apart from aforementioned measurements, we further leverage the support degree as an indicator for evaluating the rule importance. Consequently, the final score to one specific rule (i.e.,  $\text{Score}(r_i)$ ) is to involve four statistical features, including support degree, rule length, aggregated similarity, and redundancy, and the following equation is proposed to formulate this calculation:

TABLE 3: List of selected variables from the breast cancer data.

Variables	Description
MAR_STAT	Marital status
SEX	Gender
SEQ_NUM	Seq. of malignant
LATERAL	Laterality
NUMPRIMS	Number of primaries
SRV_TIME	Survival months
ORIGIN	Origin
AGE_DX	Age at diagnosis
SITEO2V	Primary site
RADIATN	Method of radiation therapy
HISTREC	Histology
ADJAJCCSTG	AJCC 6 <sup>th</sup> stage

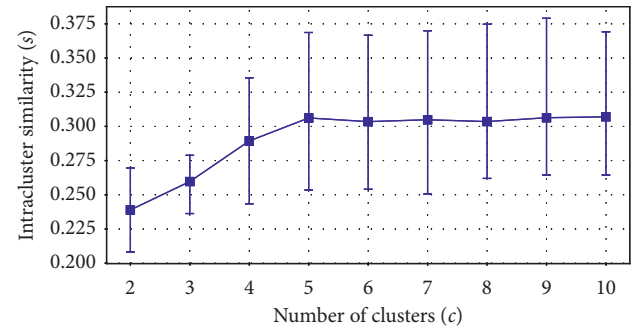


FIGURE 4: Intercluster similarity  $\mathcal{S}_{\text{inter}}$  based on varying values of the number of clusters ( $c$ ), while the upper and lower bound represent the maximal and minimal values of  $\mathcal{S}_{\text{inter}}$ . The related computational time is 426.07 s, 564.11 s, 710.78 s, 870.91 s, 1027.49 s, 1188.15 s, 1345.25 s, 1501.67 s, and 1658.97 s, respectively.

TABLE 4: The completed rule sets generated using 85,189 patient samples, with different settings of support degrees.

Support degrees	Number of rules	Support degrees	Number of rules
(0.0, 0.1]	9096	(0.1, 0.2]	2043
(0.2, 0.3]	287	(0.3, 0.4]	196
(0.4, 0.5]	125	(0.5, 0.6]	93
(0.6, 0.7]	20	(0.7, 0.8]	16
(0.8, 0.9]	10	(0.9, 1.0]	1

TABLE 5: Statistical information from the result of community-detection clustering when  $K = 5$ .

Cluster index	Number of rules	Average RIC within one cluster
0	1477	0.2535
1	2640	0.3135
2	2534	0.2949
3	3056	0.2903
4	2180	0.3687

TABLE 6: List of selected rules from the proposed algorithm, showing the top 15 rules from five clusters.

Index	Rule	Support	Confidence
0	(SEQ_NUM(0), HISTREC(9), RADIATN(0)) $\rightarrow$ (SEX(2), NUMPRIMS(1), ORIGIN(0))	0.428	0.942
0	(NUMPRIMS(1), HISTREC(9), RADIATN(0)) $\rightarrow$ (SEX(2), SEQ_NUM(0), ORIGIN(0))	0.428	0.938
0	(NUMPRIMS(1), RADIATN(0)) $\rightarrow$ (SEX(2), ORIGIN(0))	0.509	0.944
1	(NUMPRIMS(1), HISTREC(9), ORIGIN(0)) $\rightarrow$ (SEX(2), SEQ_NUM(0))	0.695	0.987
1	(SEX(2), SEQ_NUM(0), HISTREC(9)) $\rightarrow$ (NUMPRIMS(1), ORIGIN(0))	0.695	0.947
1	(SEX(2), NUMPRIMS(1), HISTREC(9)) $\rightarrow$ (ORIGIN(0))	0.699	0.948
2	(NUMPRIMS(1), HISTREC(9)) $\rightarrow$ (SEX(2), SEQ_NUM(0))	0.733	0.987
2	(NUMPRIMS(1), HISTREC(9)) $\rightarrow$ (SEX(2))	0.737	0.993
2	(NUMPRIMS(1)) $\rightarrow$ (SEX(2), SEQ_NUM(0))	0.860	0.987
3	(SEQ_NUM(0), ORIGIN(0)) $\rightarrow$ (SEX(2), NUMPRIMS(1))	0.816	0.991
3	(NUMPRIMS(1), ORIGIN(0)) $\rightarrow$ (SEX(2), SEQ_NUM(0))	0.816	0.987
3	(SRV(< 60)) $\rightarrow$ (ORIGIN(0))	0.574	0.948
4	(SRV(< 60)) $\rightarrow$ (SEX(2), ORIGIN(0))	0.569	0.940
4	(HISTREC(9), ORIGIN(0)) $\rightarrow$ (SEX(2))	0.807	0.993
4	(RADIATN(0)) $\rightarrow$ (SEX(2))	0.616	0.993

TABLE 7: List of selected rules based on their support/confidence degree.

Rule	Support	Confidence
(ORIGIN(0)) $\rightarrow$ (SEX(2))	0.943	0.992
(SEQ_NUM(0)) $\rightarrow$ (NUMPRIMS(1))	0.866	0.999
(NUMPRIMS(1)) $\rightarrow$ (SEQ_NUM(0))	0.866	0.995
(NUMPRIMS(1)) $\rightarrow$ (SEX(2))	0.864	0.992
(SEQ_NUM(0)) $\rightarrow$ (SEX(2))	0.860	0.992
(NUMPRIMS(1)) $\rightarrow$ (SEX(2), SEQ_NUM(0))	0.860	0.987
(SEQ_NUM(0)) $\rightarrow$ (SEX(2), NUMPRIMS(1))	0.860	0.991
(SEX(2), NUMPRIMS(1)) $\rightarrow$ (SEQ_NUM(0))	0.860	0.995
(SEX(2), SEQ_NUM(0)) $\rightarrow$ (NUMPRIMS(1))	0.860	0.999
(NUMPRIMS(1), SEQ_NUM(0)) $\rightarrow$ (SEX(2))	0.860	0.992
(HISTREC(9)) $\rightarrow$ (SEX(2))	0.851	0.993
(ORIGIN(0), SEQ_NUM(0)) $\rightarrow$ (NUMPRIMS(1))	0.822	0.999
(ORIGIN(0), NUMPRIMS(1)) $\rightarrow$ (SEQ_NUM(0))	0.822	0.995
(ORIGIN(0), NUMPRIMS(1)) $\rightarrow$ (SEX(2))	0.820	0.992
(ORIGIN(0), SEQ_NUM(0)) $\rightarrow$ (SEX(2))	0.817	0.992

$$\begin{aligned} \text{Score}(r_i) = & \text{Supp}(r_i) + \lambda_L \text{Length}(r_i) + \lambda_A \text{Aggregate}(r_i) \\ & - \lambda_R \text{Redundancy}(r_i), \end{aligned} \quad (7)$$

where  $\lambda_L$ ,  $\lambda_A$ , and  $\lambda_R$  are the penalty term for balancing four statistical features, respectively. Eventually, the cluster-based summarization algorithm is proposed for determining informative rules, which is further shown in Algorithm 1.

**3.4. Summary.** The main contribution of this work is to formulate the problem of rule summarization as a graph clustering process. Next, we discuss the computational complexity of the proposed method. Given a dataset with  $N$  samples, the FP-Growth algorithm is employed to mining potential rules with the cost of  $\mathcal{O}(N)$ . Next the rule-similarity graph is constructed, which requires a complexity of  $\mathcal{O}(M^2)$  (where  $M$  is the total number of generated rules; note that it usually leads to  $N \ll M$ ). Note that with the established graph, there will be  $M$  vertex and  $M(M-1)/2$  edges as we consider the full-connect strategy. As such, applying the community-detection algorithm to cluster this

rule graph costs  $\mathcal{O}(M^2)$ . Finally, the cluster-based summarization algorithm needs to go through all  $K$  clusters to select top  $N_s$  rules. As such, for the  $k$ -th cluster, the time complexity could be  $\mathcal{O}(\|C_k\| \cdot \min(N_s, \|C_k\|) \approx \mathcal{O}(\|C_k\| \cdot N_s))$ , where  $\|C_k\|$  is the number of rules within the  $k$ -th cluster. In the worst case, we have  $\|C_k\| = M$ , thereby leading to the worst complexity of  $\mathcal{O}(M \cdot N_s)$ . Overall, the complexity order of the proposed algorithm is  $\mathcal{O}(N) + \mathcal{O}(M^2) + \mathcal{O}(M^2) + \mathcal{O}(K \cdot M \cdot N_s) \approx \max(\mathcal{O}(M^2), \mathcal{O}(K \cdot M \cdot N_s))$ .

Notice that the overall complexity for the proposed algorithm depends on either the total number of generated rules (i.e.,  $M$ ) or the number of available clusters and rules to be selected. In the worst case, the complexity could be  $\mathcal{O}(M^2)$  if we select all rules; by contrast, it will be  $\mathcal{O}(K \cdot M)$  as only one rule to be chosen from individual cluster.

## 4. Experimental Results and Analysis

This section discusses the experimental results by performing the application of our proposed algorithm to the SEER dataset. The details about the employed dataset are presented in the following section. The aim of the

TABLE 8: Results from traditional rule summarization techniques, using the APRX-COLLECTION method.

Rule from APRX-COLLECTION	Support	Confidence
(SEX(2), RADIATN(0), LATERAL(1), MAR_STAT(2), SEQ_NUM(0)) → (ORIGIN(0), NUMPRIMS(1), HISTREC(9))	0.108	0.828
(RADIATN(1), LATERAL(2), SEQ_NUM(0), HISTREC(9), NUMPRIMS(1)) → (ORIGIN(0))	0.119	0.945
(SEX(2), MAR_STAT(5), SRV(< 60)) → (NUMPRIMS(1), SEQ_NUM(0))	0.122	0.918
(SRV(< 60), C509, RADIATN(0)) → (ORIGIN(0))	0.109	0.943
(AGE_DX(2), MAR_STAT(2)) → (SEX(2))	0.106	0.992
(C508, LATERAL(2)) → (SEX(2))	0.101	0.998
(ADJACCSTG(70.0), ORIGIN(0)) → (SEX(2))	0.101	0.990
(NUMPRIMS(2), SEQ_NUM(1)) → (HISTREC(9))	0.101	0.883
(C504, LATERAL(1)) → (ORIGIN(0))	0.138	0.953
(SRV(> 60), AGE_DX(3)) → (ORIGIN(0))	0.103	0.964
(MAR_STAT(1), ORIGIN(0)) → (NUMPRIMS(1))	0.102	0.876
(AGE_DX(1), SRV(> 60)) → (ORIGIN(0))	0.137	0.944
(NUMPRIMS(2)) → (SEQ_NUM(1))	0.115	0.993
(MAR_STAT(5)) → (ORIGIN(0))	0.189	0.958
(ADJACCSTG(70.0)) → (NUMPRIMS(1))	0.102	0.940

experiments is to (i) evaluate the influence of key parameters to the clustering performance and (ii) compare the proposed algorithm with existing work for the rule summarization.

*4.1. Experiments Design.* As mentioned before, the SEER dataset consists of samples from a great number of cancer types. In this study, the breast cancer is explicitly employed as the main resource. Relevant samples and variables from SEER that are associated with patient survivability and tumor status are selected based on a set of inclusions (the selection criteria for variables can be found in our preliminary work [22]). As such, 85,189 patient samples (with 12 variables) are identified as the main experimental data, and detailed description for chosen variables can be found in Table 3.

Note that among these 12 variables, two of them are with the continuous type, i.e., AGE\_DX and SRV\_TIME, and Figures 2 and 3 illustrate their distribution, respectively. In relation to the continuous data, the data discretization process is considered to split them into  $L$  groups, while each group will be assigned to one unique label. In this study, we set  $L = 3$  and 2 for the variables of AGE\_DX and SRV\_TIME, respectively. As a result, the discrete results for AGE\_DX will be [17, 53], [53, 67], and [67, 104] due to an equal-size separation. Meanwhile, as for the variable of survival months, we are taking SRV\_TIME = 60 as the splitting threshold, as the majority of studies has categorized patients' survivability using a threshold of five years [10, 22].

*4.2. Results from Cluster-Based Summarization.* In this section, we focus on the results from the proposed algorithm, with respect to the rule summarization for patients' behavior. Towards this end, we start by examining the completed rule sets and then performing rules cluster, and more importantly we investigate the summarization results across different clusters. To begin with, we utilize the FP-Growth algorithm to generate the completed rule set, which leads to a total of 11,887 rules. Table 4 shows the generated rules as a function of the support degrees, while a higher

support is normally with a smaller number of rules. Note that the threshold for the confidence degree is fixed as 0.5 in all cases.

We then perform the community detection to cluster the generated rule set. A particular problem with the application of community-detection clustering is to determine the number of clusters ( $K$ ). Yet, the selection of an appropriate value for the number of clusters has crucial impact on the clustering performance. Having said that, a too big value for  $K$  would make it difficult to interpret the result, not to mention the computational cost; while a smaller value of  $K$  could fail to group similar samples and result in poor clustering performance. To identify an optimal value for the number of clusters, the measurement of intercluster similarity ( $\mathcal{S}_{inter}$ ) is introduced that is estimated as follows:

$$\mathcal{S}_{inter} = \frac{1}{K} \sum_{k=1}^K \left( \frac{1}{N_k^2} \sum_{i=1}^{N_k} \sum_{j=1, j \neq i}^{N_k} \text{RIC}(r_i, r_j) \right), \quad (8)$$

where RIC represents the relative item coverage (defined in equation (3)),  $N_k$  represents the number of rules in the  $k$ -th cluster ( $k = [1, 2, \dots, K]$ ), and  $K$  is the number of clusters.

Using this measurement, our aim is to select the value of  $K$  that leads to the biggest value of  $\mathcal{S}_{inter}$ , so that similar rules are grouped together to maximize the intercluster similarity. Towards this end, Figure 4 plots the results of intercluster similarity for varying values of  $K$ , where  $K \in [2, 10]$ . At first, we confirm that the computational time associated with different sizes of clusters is increasing with respect to  $K$ . For instance, the minimal and maximal time has been found from  $K = 2$  (with 426.07 seconds) and  $K = 10$  (with 1658.97 seconds), respectively. On the other hand, we observe that the intercluster similarity ( $\mathcal{S}_{inter}$ ) performs stably after  $K \geq 5$ .

Given the expensive computation with a bigger value of  $K$ , we decide to take the optimal value of  $K = 5$  in the following study. As such, we perform the community-detection algorithm to cluster rules into five groups, and the statistical information about the particular clustering result is summarized in Table 5.



TABLE 9: Results from traditional rule summarization techniques, using the RPGlobal method.

Rule from RPGlobal	Support	Confidence
(RADIATN(0), LATERAL(2), SEQ_NUM(0), SRV(< 60), ORIGIN(0), HISTREC(9), NUMPRIMS(1)) $\rightarrow$ (SEX(2))	0.135	0.993
(SEX(2), RADIATN(0), LATERAL(2), SRV(< 60), ORIGIN(0), HISTREC(9), NUMPRIMS(1)) $\rightarrow$ (SEQ_NUM(0))	0.135	0.998
(RADIATN(1), MAR_STAT(2), SEQ_NUM(0), ORIGIN(0), HISTREC(9), NUMPRIMS(1)) $\rightarrow$ (SEX(2))	0.128	0.991
(SEX(2), LATERAL(1), MAR_STAT(2), SRV(< 60), ORIGIN(0), HISTREC(9)) $\rightarrow$ (SEQ_NUM(0))	0.104	0.923
(SEX(2), AGE_DX(2), RADIATN(0), SEQ_NUM(0), SRV(< 60)) $\rightarrow$ (NUMPRIMS(1))	0.101	0.999
(SEQ_NUM(0), AGE_DX(3), HISTREC(9), MAR_STAT(5), NUMPRIMS(1)) $\rightarrow$ (SEX(2))	0.101	0.996
(SEX(2), SEQ_NUM(0), AGE_DX(3), HISTREC(9), MAR_STAT(5)) $\rightarrow$ (NUMPRIMS(1))	0.101	0.998
(SEX(2), C504, ORIGIN(0), MAR_STAT(2)) $\rightarrow$ (HISTREC(9))	0.135	0.905
(SRV(> 60), NUMPRIMS(1), HISTREC(9)) $\rightarrow$ (ORIGIN(0))	0.127	0.962
(AGE_DX(1), NUMPRIMS(1), SEQ_NUM(0), RADIATN(1)) $\rightarrow$ (SEX(2))	0.107	0.995
(SEX(2), SRV(> 60), RADIATN(0), LATERAL(2)) $\rightarrow$ (HISTREC(9))	0.112	0.893
(SEX(2), C509, ORIGIN(0), LATERAL(1)) $\rightarrow$ (NUMPRIMS(1))	0.126	0.881
(NUMPRIMS(1), MAR_STAT(1), ORIGIN(0), SEQ_NUM(0)) $\rightarrow$ (SEX(2))	0.123	0.991
(MAR_STAT(1)) $\rightarrow$ (SEX(2), NUMPRIMS(1), SEQ_NUM(0))	0.106	0.864
(SEX(2), C508, RADIATN(0)) $\rightarrow$ (ORIGIN(0))	0.116	0.953

TABLE 10: Selected item sets and their descriptions (ordered alphabetically) from our study.

Item	Description
AGE_DX(1)	Age at diagnosis less than 53
AGE_DX(2)	Age at diagnosis more than or equal to 53 and less than 67
AGE_DX(3)	Age at diagnosis more than or equal to 67
ADJACCSTG(70.0)	Breast adjusted AJCC 6 <sup>th</sup> stage (1988+)–IV
C50(4,8,9)	Breast
HISTREC(9)	Histology recode—broad groupings (8500–8549)
LATERAL(1)	Not a paired site
LATERAL(2)	Right: origin of primary
MAR_STAT(1)	Single (never married)
MAR_STAT(2)	Married (including common law)
MAR_STAT(5)	Widowed
NUMPRIMS(1)	One primary
NUMPRIMS(2)	Two primaries
ORIGIN(0)	Non-Spanish/Non-Hispanic
RADIATN(0)	None or diagnosed at autopsy
RADIATN(1)	Beam radiation
SEQ_NUM(0)	One primary only in the patient’s lifetime
SEQ_NUM(1)	First of two or more primaries
SEX(2)	Female
SRV(< 60)	Survival months less than 60 months or equal
SRV(> 60)	Survival months more than 60 months

Next, the proposed summarization algorithm is performed on these five clusters to identify informative rules. In this study, we are setting key parameters for summarization as follows: the number of selected rules  $N_s = 3$ , and the penalty terms of  $\lambda_L = \lambda_A = \lambda_R = 0.35$ . As a result, summarized rules are listed in Table 6, which shows a diverse coverage of support degree and number of items. For instance, all selected rules, together, cover nearly 98% of patient samples (high support degree) while seven distinct items occur from the results that are identified as key items, including SEQ\_NUM(0), HISTREC(9), RADIATN(0), SEX(2), NUMPRIMS(1), ORIGIN(0), and SRV(< 60). We will then compare our proposed approach with others.

**4.3. Comparison Results.** In this section, we compare the proposed algorithm with the existing approaches in terms of mining risk factors associated with patients’ disease development. To begin with, we first extract top 15 rules from the completed rule set (without any summarization techniques), by simply ranking them based on their support degrees. As such, these rules are cast as the baseline results, and Table 7 illustrates this rule set with high support degree.

As observed, the majority of rules from the baseline results is overlapping each other. For instance, there are only five items observed from both the antecedent and consequent, including “SEX(2),” “NUMPRIMS(1),” “ORIGIN(0),” “SEQ\_NUM(0),” and “HISTREC(9),” respectively. That is, approximately 88.1% of the items have been repeated

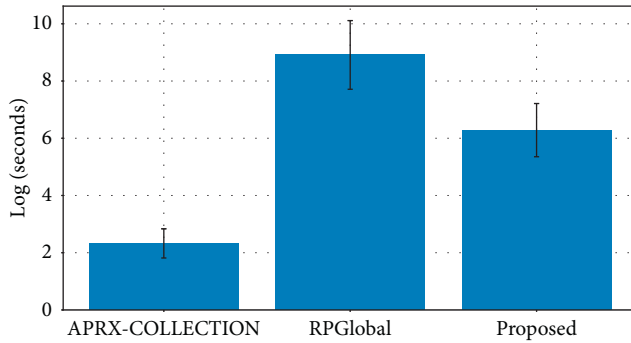


FIGURE 5: Comparison results from different summarization approaches in terms of their computational cost.

from this baseline result. On the other hand, rules from Table 7 are with relatively simple format (with no more than 3 items), which also indicates we could miss some complex or advanced rules. More importantly, although these top 15 rules are selected based on their support degree, their relevant data coverage is relatively low compared to our method (with 94.6% of the entire data). By contrast, the proposed method performs better than the baseline rules, by identifying more key items (seven) and covering a larger number of patients (98.3%).

Next, traditional rule summarization techniques are considered, including APRX-COLLECTION [15] and RPGlobal [16] method, while their results are shown in Tables 8 and 9, respectively. Again, those methods are applied to summarize rules by selecting the top 15 ones.

The results from the APRX-COLLECTION algorithm clearly indicate its preference of selecting long rules, regardless of their support. As mentioned before, the principle of APRX-COLLECTION is to choose rules with a large item coverage. As a result, the top two rules from APRX-COLLECTION, for instance, are with 8 and 6 items, respectively. However, the major problem with APRX-COLLECTION is the data coverage, that is, the support degree from rules. All selected rules are with support under the value of 0.2, which is associated with a very small population of patients. In other words, results from APRX-COLLECTION are insufficient and cover the majority patient's data, while they could lead to a misleading summarization result.

The similar problem is with the RPGlobal algorithm. Again, the long rules are still preferred and rules with more items are selected. However, RPGlobal also considers removing redundancy by encouraging rules that cover other population. As a result, we can see that the data coverage from RPGlobal is slightly better than that of APRX-COLLECTION, with the average support degrees of 0.1171 (RPGlobal) and 0.1154 (APRX-COLLECTION), respectively. Consequently, the major problem with traditional summarization technologies is that the selected rules are associated with a small data coverage, thereby reducing their generalization capability.

On the other hand, as mentioned before, our summarized rule set (illustrated in Table 6) shows a nice balance between the support degree (data coverage) and the number

of covered items. For instance, our algorithm leads to a total number of 7 distinct items, which are more than those of baseline (i.e., five). Therefore, more details or complex rules are allowed to be selected using our method. In addition, compared to traditional summarization methods, the proposed approach leads to a summarized rule set with approximately 98% support degree, which outperforms its peers that are with less than 25% support degree. In other words, our method is able to cover the majority of patient cases. Overall, the comparison results clearly show the summarization applicability of our method to represent an overlarge rule set by identifying important rules (with high support degree in terms of data coverage and less redundancy in terms of item overlapping).

At last, we investigate the computational cost from different approaches, and the comparison results are shown in Figure 5. From the experimental results, we notice that the proposed algorithm requires an affordable time for the rule summarization. For example, compared to the RPGlobal method, the proposed algorithm needs 535.8 seconds, which is much better than that of RPGlobal (with 7422.18 seconds). Although we notice that the APRX-COLLECTION approach comes with the least time of 10.23 seconds, its summarization result is the worst among three cases. As such, the satisfactory performance from our proposed algorithm compensate for its computational cost. More importantly, the computational time is accumulated with five clusters in our approach. Note that we can perform the summarization within individual clusters in a parallel way, which could reduce the total cost further. We will leave this work for our future study.

## 5. Conclusion

In this paper, we propose a novel rule summarization algorithm for identifying informative rules from a cancer-relevant data repository. Three phases are introduced that are capable of generating a comprehensive rule set and relevant rule-similarity graph, performing the community detection to cluster rules, and then selecting important rules to produce a fluent rule summary.

The proposed method is evaluated using the breast cancer dataset from the Surveillance, Epidemiology, and End Results (SEER) resource, which include 85,189 patient samples and 12 variables. The data leads to a completed rule set with over 11,887 rules. By applying the proposed method, we manage to identify the informative rules with high support degree in terms of data coverage and less redundancy in terms of item overlapping. Experimental results also demonstrate that the proposed method leads to competitive performance compared to existing approaches, in terms of the satisfactory summarization results and affordable computational cost. Overall, the proposed method offers a flexible capability and efficient applicability for processing a large amount of medical data that in turn can be utilized to facilitate patients' risk management.

Table 10 summaries the item sets and related medical information from rules within Tables 6–9, respectively.

## Data Availability

The Surveillance, Epidemiology, and End Results (SEER) data used in our manuscript to support the findings have been deposited in the publicly available, open-source data repository, which is accessible from <https://seer.cancer.gov/>.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (Grant no. 61873004) and the Humanities and Social Sciences Foundation of Anhui Department of Education, China (Grant no. SK2017A0098).

## References

- [1] Y. Bédard, P. Gosselin, S. Rivest et al., “Integrating gis components with knowledge discovery technology for environmental health decision support,” *International Journal of Medical Informatics*, vol. 70, no. 1, pp. 79–94, 2003.
- [2] J. B. Liu, Z. Y. Shi, Y. H. Pan, J. Cao, M. Abdel-Aty, and U. Al-Juboori, “Computing the laplacian spectrum of linear octagonal-quadrilateral networks and its applications,” *Polycyclic Aromatic Compounds*, pp. 1–12, 2020.
- [3] A. Yardimci, “Soft computing in medicine,” *Applied Soft Computing*, vol. 9, no. 3, pp. 1029–1043, 2009.
- [4] J.-B. Liu, J. Zhao, and Z.-Q. Cai, “On the generalized adjacency, laplacian and signless laplacian spectra of the weighted edge corona networks,” *Physica A: Statistical Mechanics and Its Applications*, vol. 540, Article ID 123073, 2020.
- [5] C. M. Lynch, B. Abdollahi, J. D. Fuqua et al., “Prediction of lung cancer patient survival via supervised machine learning classification techniques,” *International Journal of Medical Informatics*, vol. 108, pp. 1–8, 2017.
- [6] H. Lu, H. Wang, and S. W. Yoon, “A dynamic gradient boosting machine using genetic optimizer for practical breast cancer prognosis,” *Expert Systems with Applications*, vol. 116, pp. 340–350, 2019.
- [7] A. Rios, E. B. Durbin, I. Hands et al., “Cross-registry neural domain adaptation to extract mutational test results from pathology reports,” *Journal of Biomedical Informatics*, vol. 97, Article ID 103267, 2019.
- [8] J. Yang, J. Ma, and S. K. Howard, “Usage profiling from mobile applications: a case study of online activity for australian primary schools,” *Knowledge-Based Systems*, vol. 191, Article ID 105214, 2019.
- [9] J. Yang, B. Yecies, and P. Y. Zhong, “Characteristics of Chinese online movie reviews and opinion leadership identification,” *International Journal of Human-Computer Interaction*, vol. 36, no. 3, pp. 211–226, 2020.
- [10] J. A. Delgado-Osuna, C. García-Martínez, J. Gómez-Barbadillo, and S. Ventura, “Heuristics for interesting class association rule mining a colorectal cancer database,” *Information Processing & Management*, vol. 57, no. 3, Article ID 102207, 2020.
- [11] M. R. Nalluri, K. Kannan, X.-Z. Gao, and D. S. Roy, “Multiobjective hybrid monarch butterfly optimization for imbalanced disease classification problem,” *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 7, pp. 1423–1451, 2019.
- [12] D. Nguyen, W. Luo, D. Phung, and S. Venkatesh, “Ltarm: a novel temporal association rule mining method to understand toxicities in a routine cancer treatment,” *Knowledge-Based Systems*, vol. 161, pp. 313–328, 2018.
- [13] R. Saini, P. Kumar, B. Kaur, P. P. Roy, D. P. Dogra, and K. C. Santosh, “Kinect sensor-based interaction monitoring system using the blstm neural network in healthcare,” *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 9, pp. 2529–2540, 2019.
- [14] J. Yang and B. Yecies, “Mining Chinese social media UGC: a big-data framework for analyzing Douban movie reviews,” *Journal of Big Data*, vol. 3, no. 1, p. 3, 2016.
- [15] F. Afrati, A. Gionis, and H. Mannila, “Approximating a collection of frequent sets,” in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Seattle, WA, USA, pp. 12–19, August 2004.
- [16] D. Xin, J. Han, X. Yan, and H. Cheng, “Mining compressed frequent-pattern sets,” in *Proceedings of the 31st International Conference on Very Large Data Bases*, VLDB Endowment, Trondheim, Norway, pp. 709–720, August 2005.
- [17] M. Girvan and M. E. J. Newman, “Community structure in social and biological networks,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [18] F. Radicchi, C. Castellano, F. Ceconi, V. Loreto, and D. Parisi, “Defining and identifying communities in networks,” *Proceedings of the National Academy of Sciences*, vol. 101, no. 9, pp. 2658–2663, 2004.
- [19] P. Pons and M. Latapy, “Computing communities in large networks using random walks,” in *Computer and Information Sciences. Lecture Notes in Computer Science*, p. 3733, Springer, Berlin, Germany, 2005.
- [20] U. N. Raghavan, R. Albert, and S. Kumara, “Near linear time algorithm to detect community structures in large-scale networks,” *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 76, no. 3, Article ID 036106, 2007.
- [21] M. Rosvall and C. T. Bergstrom, “Maps of random walks on complex networks reveal community structure,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008.
- [22] N. Shukla, M. Hagenbuchner, K. T. Win, and J. Yang, “Breast cancer data analysis for survivability studies and prediction,” *Computer Methods and Programs in Biomedicine*, vol. 155, pp. 199–208, 2018.

## Research Article

# An Improved Mask R-CNN Model for Multiorgan Segmentation

Jian-Hua Shu <sup>1</sup>, Fu-Dong Nian <sup>2,3</sup>, Ming-Hui Yu,<sup>4</sup> and Xu Li <sup>5</sup>

<sup>1</sup>School of Medical Information Engineering, Anhui University of Chinese Medicine, Hefei 230012, China

<sup>2</sup>School of Advanced Manufacturing Engineering, Hefei University, Hefei 230601, China

<sup>3</sup>School of Computer Science and Technology, Anhui University, Hefei 230601, China

<sup>4</sup>School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China

<sup>5</sup>Sino-German Institute of Applied Mathematics, Hefei University, Hefei 230601, China

Correspondence should be addressed to Fu-Dong Nian; nianfd@hfu.edu.cn

Received 8 June 2020; Accepted 1 July 2020; Published 24 July 2020

Guest Editor: Shaohui Wang

Copyright © 2020 Jian-Hua Shu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical image segmentation is a key topic in image processing and computer vision. Existing literature mainly focuses on single-organ segmentation. However, since maximizing the concentration of radiotherapy drugs in the target area with protecting the surrounding organs is essential for making effective radiotherapy plan, multiorgan segmentation has won more and more attention. An improved Mask R-CNN (region-based convolutional neural network) model is proposed for multiorgan segmentation to aid esophageal radiation treatment. Due to the fact that organ boundaries may be fuzzy and organ shapes are various, original Mask R-CNN works well on natural image segmentation while leaves something to be desired on the multiorgan segmentation task. Addressing it, the advantages of this method are threefold: (1) a ROI (region of interest) generation method is presented in the RPN (region proposal network) which is able to utilize multiscale semantic features. (2) A prebackground classification subnetwork is integrated to the original mask generation branch to improve the precision of multiorgan segmentation. (3) 4341 CT images of 44 patients are collected and annotated to evaluate the proposed method. Additionally, extensive experiments on the collected dataset demonstrate that the proposed method can segment the heart, right lung, left lung, planning target volume (PTV), and clinical target volume (CTV) accurately and efficiently. Specifically, less than 5% of the cases were missed detection or false detection on the test set, which shows a great potential for real clinical usage.

## 1. Introduction

Diagnostic imaging plays an important role in modern medicine. Computed tomography (CT), magnetic resonance imaging (MRI), and other imaging modalities provide important assistance for diagnosis and treatment planning. Take esophageal cancer as an example; esophageal cancer is a primary malignant tumor of the esophagus. At least 200,000 people suffer from esophageal cancer every year [1], and radiotherapy is one of the main treatments in China. However, treatment planning of radiotherapy is highly dependent on planning target volume (PTV) and accurate description of the organs at risk. The accuracy of organ countersegmentation determines the quality of dose planning optimization in radiotherapy and thus affects the success or failure of radiotherapy or the incidence of complications [2].

With the increasing scale and quantity of medical images, organ segmentation via manual delineation by the clinical experience of radiologists is inefficient [3]. And it is necessary to use computers for processing and analyzing the medical images automatically. With the development of computer vision technology, many different automatic image segmentation and delineation algorithms have been developed. These algorithms are called medical image segmentation or organ segmentation [4] in the literature.

Conventional medical image segmentation/organ segmentation algorithms can be roughly divided into eight categories [4]: (a) thresholding approaches, (b) region-growing approaches, (c) classifiers, (d) clustering approaches, (e) Markov random field models, (f) deformable models, (g) artificial neural networks, and (h) atlas-guided approaches. Although these methods have made some progress, the accuracy is not sufficient.

Benefit from the continuous progress of deep learning technology, medical image segmentation/organ segmentation is currently dominated by the CNN (convolutional neural network) [5]. Similar to the object detection method, CNN-based organ segmentation can also be divided into two types: (a) one-stage algorithm, which deems the organ segmentation as a one-stage pixel classification task. The typical structure is fully convolutional networks (FCNs) [6]; (b) two-stage algorithm, which decouples the organ segmentation into organ localization and instance segmentation stages. The typical structure is region CNN (R-CNN) [7]. The most well-known one-stage CNN architecture for organ segmentation is U-Net, published by Ronneberger et al. [8]. Most state-of-the-art organ segmentation methods are the invariants of U-Net [9–11]. Although they have achieved encouraging performance, two shortcomings exist. On the one hand, many literature studies focus on single-organ segmentation, while only few works are made effort to address the multiorgan segmentation problem [12, 13]. On the other hand, two-stage segmentation methods work well for multiobject segmentation on the natural image segmentation dataset [14] but worse than the one-stage algorithm on medical image segmentation [15]. Therefore, mining the potential of the two-stage multiorgan segmentation algorithm has great research value.

In this paper, to address the shortcomings mentioned above, we present an improved Mask R-CNN framework for multiorgan segmentation. Original Mask R-CNN [16] is presented to address the multi-instance segmentation problem on the natural image. Although the original Mask R-CNN has achieved state-of-the-art instance segmentation performance on general image datasets, the latest research [15] shows that it is able to accurately find bounding boxes for organs, while its performance on segmentation is worse than U-Net on the medical image segmentation dataset. We think a major reason for this is that the semantic representation obtained from the original Mask R-CNN framework is too rough for organ segmentation because organ boundaries may blur and organ shapes are various. To address it, we have made two improvements to the original Mask R-CNN: (a) a ROI (region of interest) generation method is presented in the RPN which is able to utilize multiscale semantic features; (b) a prebackground classification subnetwork is integrated to improve the precision of multiorgan segmentation. Moreover, CT images of 44 esophageal cancer patients are collected and annotated as benchmark to evaluate the proposed method.

To sum up, our contributions are as follows:

- (1) We applied the Mask R-CNN to esophageal cancer medical image processing successfully. Most existing methods focus on single-organ segmentation, while this paper devotes to address the multiorgan segmentation problem.
- (2) To provide a better multiorgan segmentation model, we propose two improvements compared with the original Mask R-CNN framework.

- (3) We conduct extensive experiments and analysis on the collected real multiorgan dataset and demonstrate the excellent performance of our proposed method on the multiorgan segmentation task.

The rest of this paper is organized as follows. Section 2 reviews and discusses the related works. Section 3 describes the proposed improved Mask R-CNN model in detail. Experimental results and comparisons are discussed in Section 4, and conclusions with the future work are described in Section 5.

## 2. Related Work

Pham et al. [4] and Litjens et al. [5] reviewed the conventional and deep learning-based organ segmentation methods, respectively. In this section, we briefly review the previous methods which are most related to our work including the conventional medical image segmentation method, deep learning-based single-organ segmentation method, and deep learning-based multiorgan segmentation method.

### 2.1. Conventional Medical Image Segmentation Method.

Conventional medical image segmentation method can be roughly divided into eight categories: (a) thresholding approaches [17]: thresholding approaches first attempt to determine an intensity value (threshold), then group all pixels with intensity greater than the threshold into one class, and all other pixels into another class. (b) Region-growing approaches [18]: region-growing approaches utilize intensity information and/or edges in the medical image to predefine criteria for extracting a region of the image that is connected. (c) Classifiers [19, 20]: classifier methods convert the medical image from the image space to the feature space first and then train classifiers on the feature space to distinguish which class of the pixel they belong to. (d) Clustering approaches [21]: commonly used clustering approaches for medical image segmentation are K-means, fuzzy c-means, and expectation-maximization. Compared with the classifiers, the clustering approaches are unsupervised approaches. (e) Markov random field models: Markov random field (MRF) is a statistical model which can be used within segmentation methods by modeling model spatial interactions between neighboring or nearby pixels. (f) Deformable models: deformable models use closed parametric curves or surfaces to delineate region boundaries. (g) Artificial neural networks (ANNs) [22]: the most widely applied use of the ANN in conventional medical image processing is as a classifier. (h) Atlas-guided approaches [23, 24]: the atlas is generated by compiling information on the anatomy that requires segmenting. This atlas is then used as a reference frame for segmenting new images. In addition, level set optimization is also utilized for multiorgan segmentation [25]. Though the methods mentioned above have achieved some progress, the accuracy of organ segmentation is not too high because all conventional methods depend on manual feature representation.

**2.2. Deep Learning-Based Single-Organ Segmentation Method.** Ronneberger et al. [8] first presented a novel CNN architecture (U-Net) and became the most popular structure in medical image analysis. The main novelty in U-Net is the combination of an equal amount of upsampling and downsampling layers. Inspired by U-Net, Zhou et al. [26] presented U-Net++, a more powerful architecture for medical image segmentation. Milletari et al. [27] proposed V-Net (a 3D variant of U-Net architecture) performing 3D image segmentation using 3D convolutional layers with an objective function directly based on the Dice coefficient. Drozdal et al. [11] investigated the use of short ResNet-like skip connections in addition to the long skip connections in a regular U-Net. Besides CNN, Xie et al. [28], Stollenga et al. [29], Chen et al. [30], and Poudel et al. [31] utilized the recurrent neural network (RNN) for organ segmentation tasks. To combat spurious responses, few papers attempt to combine the CNN/RNN with graphical models like MRFs [32] and conditional random fields (CRFs) [33] to refine the segmentation output. Although these methods have achieved encouraging performance, they were presented to address the single-organ segmentation problem, which may not be suitable/optimal for multiorgan segmentation (It is difficult to segment multiple organs at the same time, which damages the clinical auxiliary effect.).

**2.3. Deep Learning-Based Multiorgan Segmentation Method.** The research on the deep learning-based multiorgan segmentation method is in its early phase. Tong et al. [34] introduced discriminative dictionary learning for abdominal multiorgan segmentation. Lay et al. [35] used context integration and discriminative models for rapid multiorgan segmentation. Roth et al. [36] and Chen et al. [37] adopted the 3D fully convolutional network. Recently, Dong et al. [38] presented a generative model (U-Net-GAN), and Wang et al. [39] proposed densely connected U-Net for multiorgan segmentation. Lei et al. [40] presented a review of deep learning in multiorgan segmentation. Different from these methods, the proposed method in this paper aims to improve the two-stage instance segmentation algorithm which is widely used in the natural image dataset, making it suitable for the multiorgan segmentation task.

### 3. Methods

In this section, we introduce the proposed method (which is named improved Mask R-CNN) for multiorgan segmentation. As shown in Figure 1, the proposed method is based on the existing well-known multi-instance segmentation method, Mask R-CNN. Compared with the original Mask R-CNN, we have made two improvements: (a) a ROI (region of interest) generation method is presented in the RPN which is able to utilize multiscale semantic features; (b) a prebackground classification subnetwork is integrated to improve the precision of multiorgan localization. The detailed proposed approach is presented in two sections: (a) the network structure and (b) loss function.

**3.1. The Network Structure.** The network of the proposed algorithm can be mainly divided into three modules. The first module is called feature extraction and ROI generation, which is mainly composed of ResNet50 + FPN + RPN. In this module, we generate multilayer feature maps first. Then, each point on the feature map is mapped into the original image to acquire the corresponding ROI.

The second module is named region of interest alignment, which pools the ROIs obtained from the first module to a fixed size and avoids quantization error. The third module is mask acquisition. In this module, the fixed-size ROIs obtained from the second module are sent to the organ region segmentation network for generating organ mask. And at the same time, they are also sent to the fully connected layer for organ-position rectangular bounding box regression and organ classification. The above three modules are detailed as follows.

**3.1.1. Feature Extraction and ROI Generation.** The purpose of this step is to extract the features of the input image and generate the ROI in the corresponding feature layer. First, a medical CT image containing multiple organs is input to the ResNet50 network. Res2, Res3, Res4, and Res5 are the feature output layers of the ResNet [15, 41], respectively. Then, feature pyramid network (FPN) [42] is adopted to fuse these multilayer features to obtain strong semantic information and improve the accuracy of organ detection. As shown in Figure 2, the specific approach is to conduct dimensionality reduction operation on the features above Res4 (that is, to add a layer of  $1 \times 1$  convolution layer) and upsampling operation on the features above P5 to make them have the same size. Then, addition operation (adding corresponding elements) is performed on the processed P5 and the processed Res4 to output the obtained results to P4, P2, P3, and so on. Then, the RPN network is used to predict in different output layers, P2, P3, P4, and P5, to obtain ROIs.

**3.1.2. Region of Interest Alignment.** This step aims to pool all ROIs remaining on the feature maps to a fixed size. Since the ROI position is usually obtained by the regression model, it is generally a floating-point number, while the pooled feature map requires a fixed size. In order to avoid quantization errors, the ROI align [15] (illustrated in Figure 3) layer is adopted. In the presented framework, we use the ROI align layer to traverse each ROI first, keeping the floating-point number boundary unquantized. Then, the ROI is divided into  $k \times k$  cells with the boundary of each cell not quantized. Then, the fixed four coordinate positions are calculated in each cell, the values of these four positions are calculated by bilinear interpolation, and the max-pooling operation is carried out finally. Through the above operations, the fixed size ROI can be obtained with no quantization error.

In the original Mask R-CNN segmentation algorithm, the ROI obtained by the RPN network is aligned to extract the ROI features. In this step, each ROI is aligned by a single-layer (single-scale) feature. In the presented method, as shown in Figure 4, we replace the single-layer features with multilayer features, that is to say, each ROI needs to do ROI



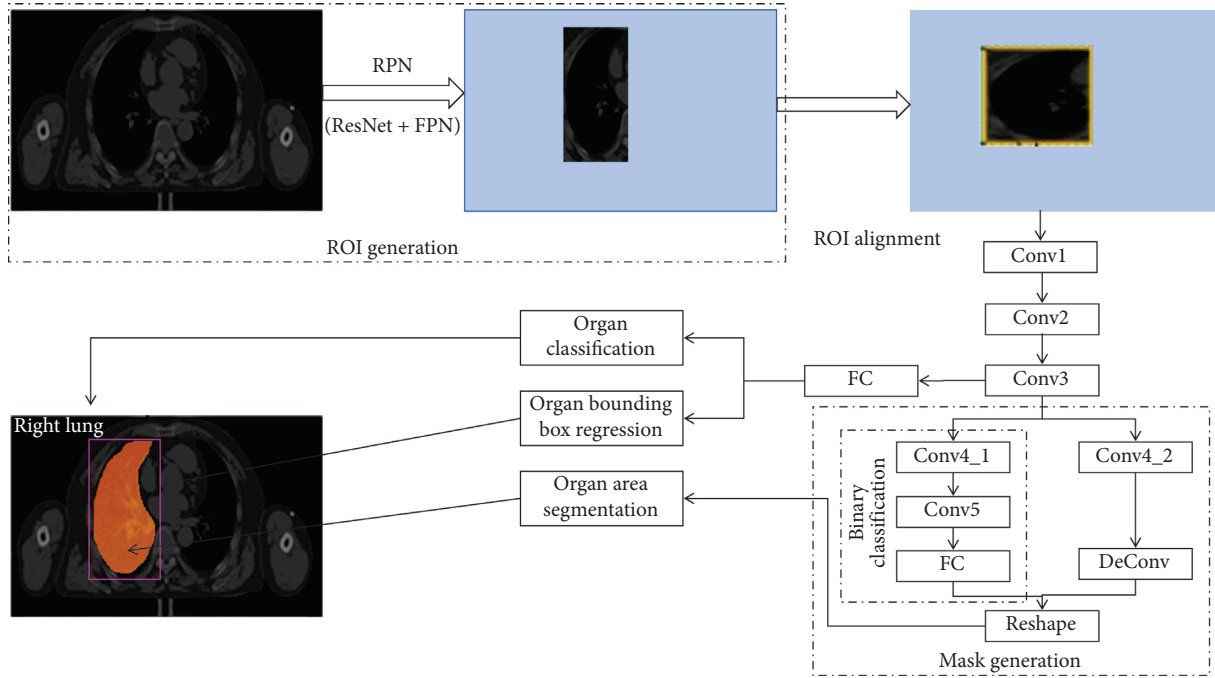


FIGURE 1: The framework of the proposed method.

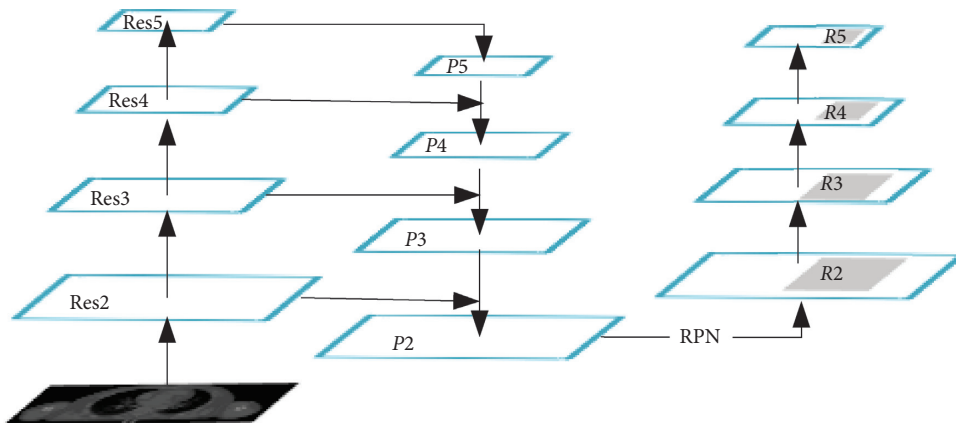


FIGURE 2: From left to right are the ResNet50 network, the feature pyramid network (FPN), and the region proposal network (RPN).

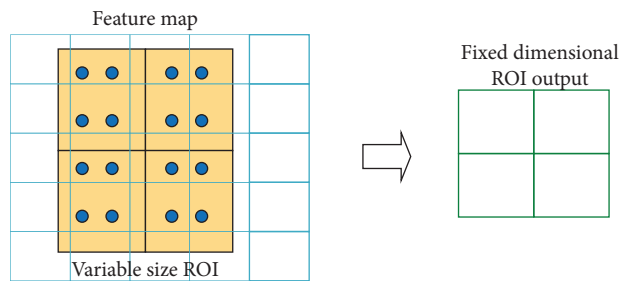


FIGURE 3: The illustration of ROI alignment operation. The blue-dotted box represents the feature map obtained after convolution, and the black solid box represents the ROI feature.

alignment operation with multilayer features, and then the ROI features of different layers will be fused together so that each ROI feature will have multilayer features.

**3.1.3. Mask Acquisition.** The goal of this step is to get the multiorgan segmentation result. ROI of pooling to a fixed size was sent to the fully connected layer for organ

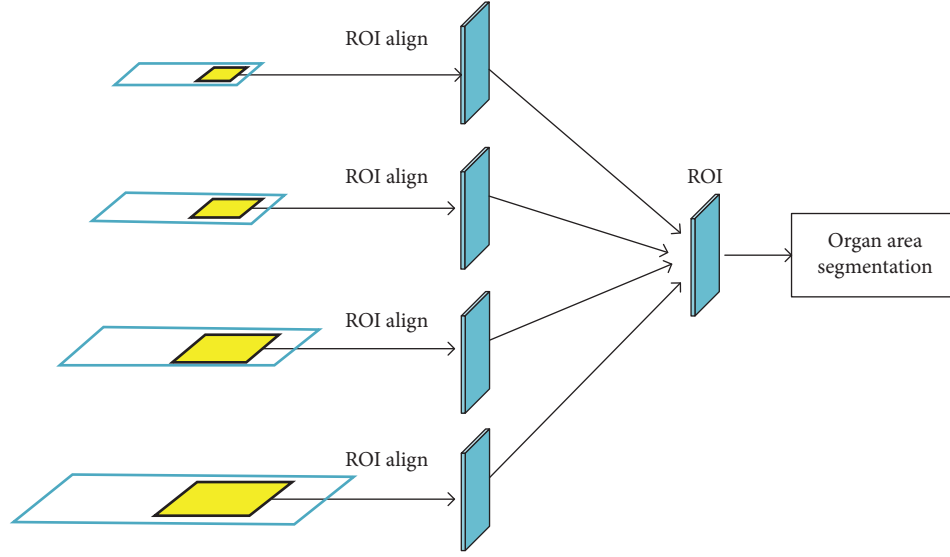


FIGURE 4: The multilayer ROI shall be merged after the operation of ROI alignment and then sent to the organ area segmentation network.

classification (6 categories including background) and organ-position rectangular bounding box regression. Meanwhile, ROI of pooling to a fixed size was also sent to a mask generation branch (i.e., fully convolution neural network operation in each ROI). Organ area segmentation is a parallel branch to organ classification and organ-position rectangular bounding box regression. As shown in Figure 5(a), the branch consists of four consecutive convolution layers and a deconvolution layer (with 2 times of upsampling). The kernel size and channels of each convolution layer are  $3 \times 3$  and 25, respectively. A binary classification branch is added to distinguish foreground and background before the original mask branch (illustrated in Figure 5(b)). The new branch contains two  $3 \times 3$  convolution layers and a fully connected layer. The dimension of the output of the new branch is the same as the original branch via a reshape operation. The output mask of these two branches was fused to get the final multiorgan segmentation result.

**3.2. Loss Function.** In terms of loss function, a third loss function, which is used to generate mask, is added on the basis of Fast R-CNN [43] so that the total loss function of our improved Mask R-CNN framework is

$$\mathcal{L} = \mathcal{L}_{\text{cls}} + \mathcal{L}_{\text{box}} + \mathcal{L}_{\text{mask}}. \quad (1)$$

Here, the classification and regression losses are defined as  $\mathcal{L}_{\text{cls}}$  and  $\mathcal{L}_{\text{box}}$ , respectively:

$$\mathcal{L}_{\text{cls}} = -\log P_u, \quad (2)$$

$$\mathcal{L}_{\text{box}} = \sum_{i=1}^4 \text{Smooth}_{L1}(t_i^u - v_i), \quad (3)$$

$$\text{Smooth}_{L1}(X) = \begin{cases} 0.5x^2, & |x| < 1, \\ |x| - 0.5, & \text{otherwise.} \end{cases} \quad (4)$$

$P$  is a  $(k+1)$ -dimensional vector representing the probability of a pixel belonging to the  $k$  class or background. For each ROI,  $P = (P_0, P_1, \dots, P_k)$ , and  $P_u$  represents the probability corresponding to class  $u$ .  $t_u = (t_x^u, t_y^u, t_w^u, t_h^u)$  represents the predicted translation scaling parameter of class  $u$ .  $t_x^u, t_y^u$  refer to the translation with the same scale as the object proposal, and  $t_w^u, t_h^u$  refer to the height and width of the logarithmic space relative to the object proposal.  $t_1, t_2, t_3$ , and  $t_4$  in equation (3) represent  $t_x, t_y, t_w$ , and  $t_h$ , respectively. Moreover,  $v_i$  represents the corresponding parameter of the ground-truth bounding box.

Note that the smooth  $L1$  loss is utilized in equation (3); the reasons are twofold: (a) compared with the widely used  $L2$  loss, smooth  $L1$  loss is robust for outlier points. (2) Many famous object detection frameworks use smooth  $L1$  loss, e.g., Faster-RCNN and Mask R-CNN. We utilize the same bounding loss function which can guarantee the fairness of algorithm comparison. Of course, some box regression loss functions which have been proposed recently (e.g., GIoU, DIoU, and CIoU) are also compatible with the proposed framework.

$\mathcal{L}_{\text{mask}}$  in equation (1) is the mask loss of the newly added background segmentation branch (as described in Section 3.1.3). In our improved Mask R-CNN framework, the output dimension of each ROI is  $K * m * m$  for the newly added mask branch, where  $m * m$  represents the size of the mask and  $K$  represents categories, so a total of  $K$ -binary masks were generated in here. After the predicted mask was obtained, the value of the sigmoid function was calculated for each pixel of the mask, and the obtained result was taken as one of the inputs of  $\mathcal{L}_{\text{mask}}$  (cross-entropy loss function). It should be noted that only positive sample ROI is used to calculate  $\mathcal{L}_{\text{mask}}$ . The definition of the positive sample is the same as that of general object detection algorithms, and IOU greater than 0.5 is defined as the positive sample. In fact,  $\mathcal{L}_{\text{mask}}$  is very similar to  $\mathcal{L}_{\text{cls}}$  except that the former is calculated on the basis of pixels and the latter on the basis of

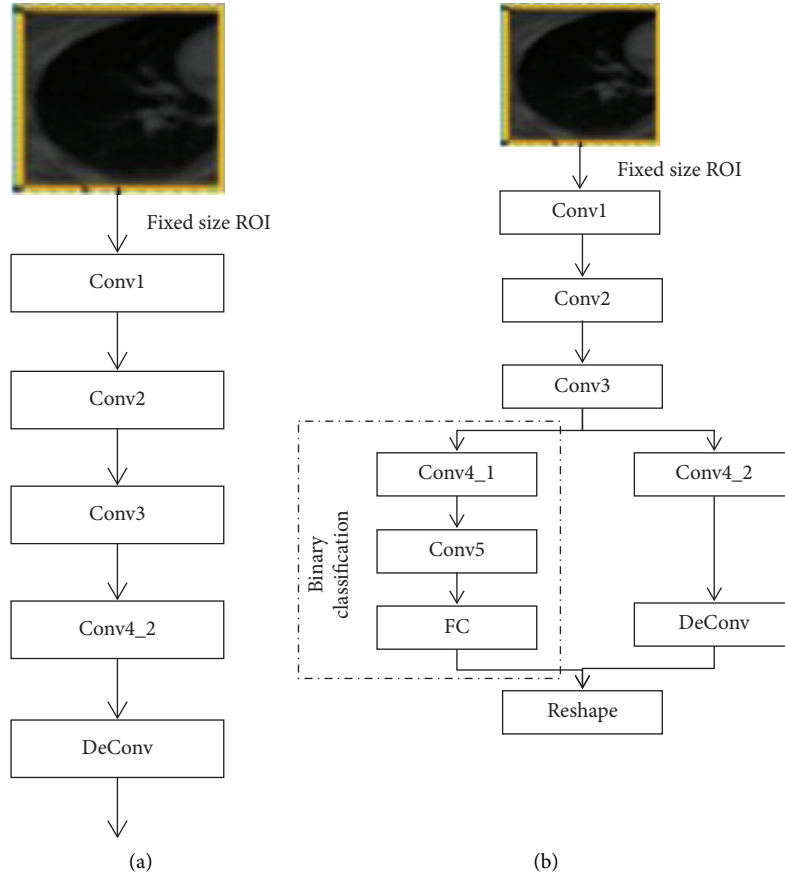


FIGURE 5: A new binary classification branch is added to the original mask generation branch. (a) The original mask generation branch. (b) The proposed mask generation branch.

images, so it is similar to  $\mathcal{L}_{cls}$  in that although  $K$  masks are given here, only the one corresponding to the ground truth is valid in calculating the cross-entropy loss function. A mask contains multiple pixels, so here,  $\mathcal{L}_{mask}$  is the average of the cross-entropy loss of each pixel:

$$\mathcal{L}_{mask} = -\frac{1}{K} \sum_{i=1}^K \sum_{j=1}^{m*m} (\log P_{i,j}^M). \quad (5)$$

Here,  $P_{i,j}^M$  is the  $j$ -th pixel of the  $i$ -th generated mask.

## 4. Experiments

In this section, we conduct extensive experiments to evaluate the proposed improved Mask R-CNN multiorgan segmentation framework. We first introduce the collected and annotated dataset in Section 4.1 followed by the evaluation criteria in Section 4.2. Then, Section 4.3 describes the implementation details. Finally, we discuss the comparison with state-of-the-art methods in Section 4.4.

**4.1. Dataset.** The utilized multiorgan segmentation dataset consists of all the slice information of 44 esophageal cancer patients, with a total of 4341 CT images. Each image was labeled with five areas (heart, right lung, left lung, PTV, and CTV) by the doctor. We use 80% of these CT images as the

training set, 5% as the validation set, and the remaining 15% as the test set.

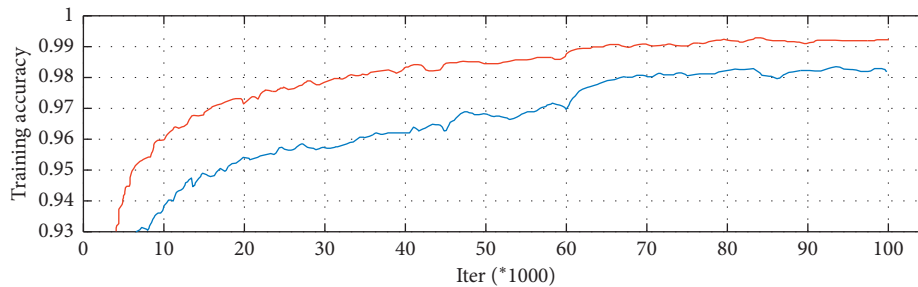
**4.2. Evaluation Criteria.** There are many evaluation criteria which are proposed to evaluate the image segmentation results, e.g., region overlap and boundary similarity [44]. Here, we select Dice coefficient (DICE) [45] and Jaccard index (JAC) [46] as criteria to evaluate the overlap between the prediction and the ground-truth organ regions. Suppose that  $x$  and  $y$  are the organ regions of the prediction and the ground truth, respectively; JAC and DICE are calculated as follows:

$$\begin{aligned} \text{JAC} &= \frac{|X \cap Y|}{|X \cup Y|}, \\ \text{DICE} &= \frac{2|X \cap Y|}{|X| + |Y|}. \end{aligned} \quad (6)$$

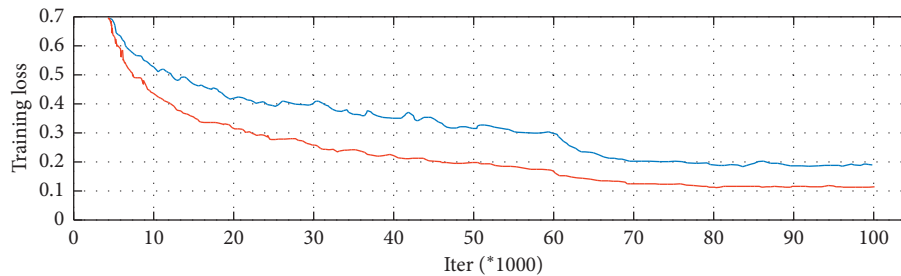
**4.3. Implementation Details.** We implement our improved Mask R-CNN model based on the framework of PyTorch. The backbone is the adjusted ResNet50 which is detailed in Section 3.1.1. We use the stochastic gradient descent (SGD) optimizer with the learning rate set to 0.01 initially, and the batch size is set to 64. The maximum number of iterations is

TABLE 1: Comparisons of the proposed and state-of-the-art multiorgan segmentation methods on the presented dataset.

Methods	Organ	JAC (%)	DICE (%)
Linguraru et al. [47]	Heart	74.9 ± 4.7	71.3 ± 3.1
	Right lung	86.5 ± 2.2	86.3 ± 1.2
	Left lung	85.1 ± 1.3	84.6 ± 0.9
	PTV	82.5 ± 1.6	81.3 ± 2.8
	CTV	80.5 ± 1.2	77.5 ± 1.9
He et al. [48]	Heart	87.5 ± 1.2	86.3 ± 0.7
	Right lung	89.3 ± 1.6	87.3 ± 2.4
	Left lung	90.2 ± 1.9	88.3 ± 1.1
	PTV	86.3 ± 1.7	84.1 ± 2.4
	CTV	85.6 ± 2.3	83.7 ± 3.1
Gauriau et al. [49]	Heart	88.8 ± 0.6	87.5 ± 1.2
	Right lung	91.5 ± 0.9	91.3 ± 0.8
	Left lung	90.3 ± 0.9	90.8 ± 1.5
	PTV	89.1 ± 1.4	86.2 ± 1.7
	CTV	88.7 ± 1.9	87.3 ± 1.5
Original mask R-CNN	Heart	95.1 ± 0.5	94.2 ± 0.7
	Right lung	97.0 ± 0.9	96.2 ± 1.2
	Left lung	96.3 ± 0.4	95.1 ± 0.6
	PTV	94.7 ± 1.1	93.2 ± 0.9
	CTV	94.3 ± 0.5	93.7 ± 0.8
Improved mask R-CNN (ours)	Heart	96.6 ± 1.3	95.1 ± 1.2
	Right lung	98.1 ± 0.5	97.8 ± 0.3
	Left lung	97.6 ± 0.4	96.2 ± 1.4
	PTV	95.3 ± 1.5	95.2 ± 0.7
	CTV	95.8 ± 0.7	94.4 ± 1.2



(a)



(b)

FIGURE 6: (a) The accuracy curves in the training stage. (b) The loss curves in the training stage. Red curve represents the proposed improved Mask R-CNN model, and blue represents the original Mask R-CNN model.

set to 100,000. When the number of iterations reached 50,000 and 80,000, the learning rate is reduced 10 times. All images are resized to  $800 \times 1000$ . The weight decay is set to 0.0001, and the momentum is set to 0.9 for all convolution and fully connected layers. It should be noted that all parameters in the proposed model are trained from scratch.

#### 4.4. Results and Discussion

**4.4.1. Quantitative Evaluation with State-of-the-Art Methods.** We compare our proposed methods against the current widely used multiorgan segmentation models (Linguraru et al. [47], He et al. [48], and Gauriau et al. [49]), and the

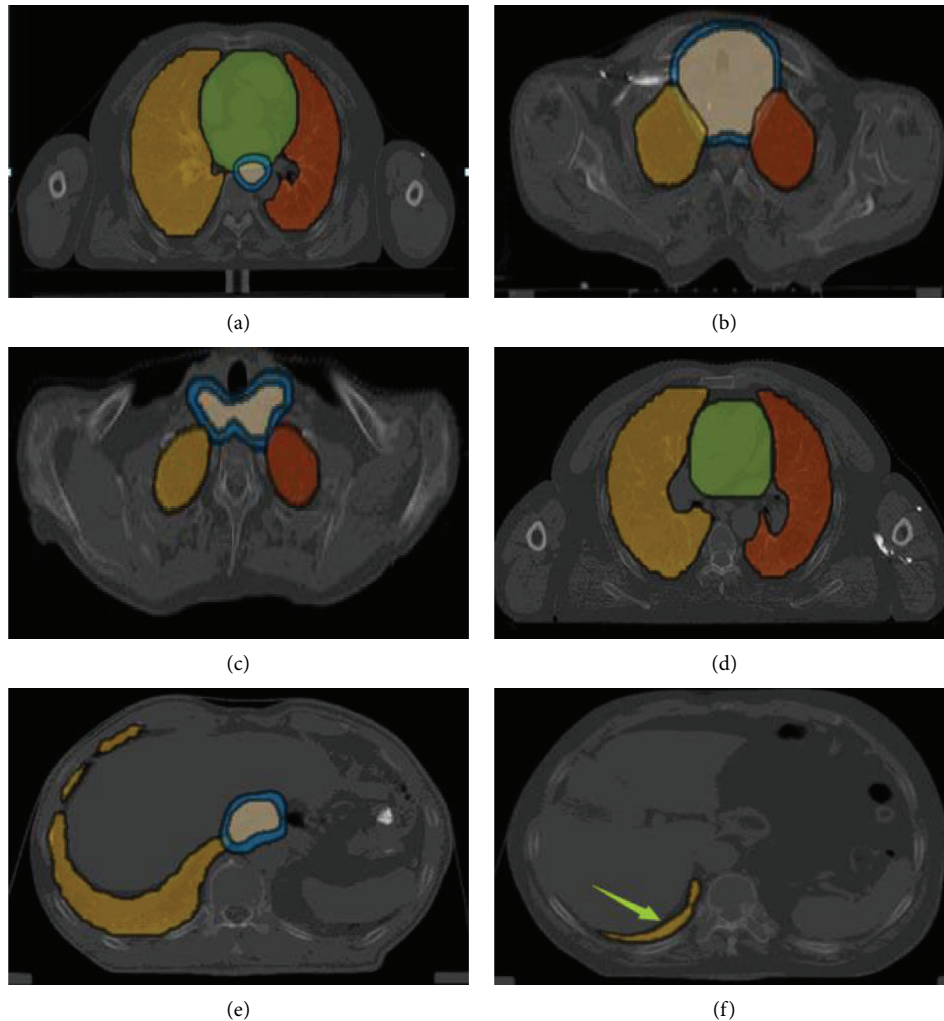


FIGURE 7: The visualization results. These six images are slices of the 38th, 60th, 66th, 71st, 89th, and 103rd layers of a patient with esophageal cancer (yellow represents the right lung, brown represents the left lung, cyan represents the heart, blue represents PTV, and gray represents CTV). (a) The slice of the 38th layer of a patient. (b) The slice of the 60th layer of a patient. (c) The slice of the 66th layer of a patient. (d) The slice of the 71st layer of a patient. (e) The slice of the 89th layer of a patient. (f) The slice of the 103rd layer of a patient.

comparison results are shown in Table 1. In general, we can observe that the proposed improved Mask R-CNN framework achieved the best performance. Moreover, Figure 6 shows the accuracy (JAC) and loss curves of the improved Mask R-CNN and original Mask R-CNN framework in the training stage. From Table 1 and Figure 6, we can conclude that the presented technique is able to improve the multiorgan segmentation performance of the original Mask R-CNN significantly and steadily.

**4.4.2. Qualitative Evaluation.** To illustrate the effectiveness of our method more visually, some multiorgan segmentation results are shown in Figure 7. The image we selected is distributed between 35 and 100 slices basically because in this range, each slice contains five organ regions that we need

basically, and the information of each organ region is relatively rich. We found that the area of some organs from the 60th to 80th layers of patients is very small, which is difficult to be observed by the naked eye due to the perspective. However, our improved mask R-CNN algorithm can also achieve good results (as shown in Figure 4, especially the area indicated by the arrow in the figure may be difficult for doctors to annotate).

Although the proposed method can achieve encouraging performance, there are still some shortcomings. Examples of false detection and missed detection segmentation are shown in Figure 8. After analyzing all failure results, we find that that the missed detection was mainly concentrated in the slices from the 1st to the 35th layer of the patient, while the missed detection was mainly concentrated in the slices from the 110th to the

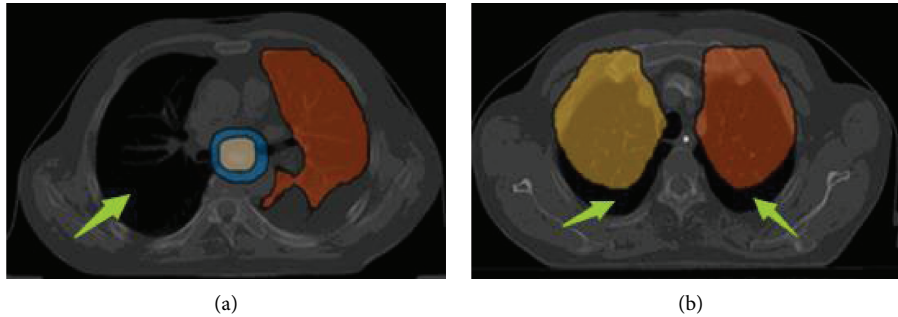


FIGURE 8: Failure cases. False detection and missed detection segmentation results. Left image is the segmentation result of the 10th slice of a patient, missed detection occurred on the right lung. Right image is the segmentation result of the 115th slice of a patient. The left lung and the right lung areas may be segmented partially (yellow represents the right lung, brown represents the left lung, cyan represents the heart, blue represents PTV, and gray represents CTV). (a) Missed detection. (b) False detection.

130th layer. By observing the constructed dataset, we find that the amount of data of the slice near the front and the slice near the back is relatively small, that is, the slice near the front layer contains relatively less target organ area, so the doctor's label information in these parts is less. Therefore, we believe the major reason for these failure cases is due to the fact that training data are insufficient and unbalanced.

## 5. Conclusion

In this paper, we present the improved Mask R-CNN segmentation framework for the medical domain that is able to work well on the multiorgan segmentation task. The proposed improved Mask R-CNN framework builds around the original Mask R-CNN framework [15]. Compared with the original Mask R-CNN framework, there are two major improvements on the improved Mask R-CNN: (a) a ROI (region of interest) generation method is presented in the RPN (region proposal network) which is able to utilize multiscale semantic features; (b) a prebackground classification subnetwork is integrated to the original mask generation branch to improve the precision of multiorgan segmentation. Additionally, extensive experiments on the collected and annotated esophageal cancer dataset demonstrate the effectiveness of the proposed framework, i.e., the improved Mask R-CNN framework can segment the heart, right lung, left lung, PTV, and CTV accurately and simultaneously. Since it is time consuming and laborious to label medical images, we will investigate semi-supervised and weakly supervised multiorgan segmentation techniques in the future.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation (NSF) of China (nos. 61702001 and 61902104), the Key Project of the Natural Science Foundation of Anhui University of Traditional Chinese Medicine (2019zrzd10), the Scientific Research Development Foundation of Hefei University (no. 19ZR15ZDA), the Talent Research Foundation of Hefei University (no. 18-19RC54) and Hefei University Annual Academy Research Development Fund Project (Natural Science) (No. 18ZR12ZDA).

## References

- [1] R. S. Holmes and T. L. Vaughan, "Epidemiology and pathogenesis of esophageal cancer," in *Seminars in Radiation Oncology*, vol. 17, pp. 2–9, Elsevier, Amsterdam, Netherlands, 2007.
- [2] Y.-l. Peng, L. Chen, G.-z. Shen et al., "Interobserver variations in the delineation of target volumes and organs at risk and their impact on dose distribution in intensity-modulated radiation therapy for nasopharyngeal carcinoma," *Oral Oncology*, vol. 82, pp. 1–7, 2018.
- [3] M. Feng, C. Demiroz, K. A. Vineberg, A. Eisbruch, and J. M. Balter, "Normal tissue anatomy for oropharyngeal cancer: contouring variability and its impact on optimization," *International Journal of Radiation Oncology \* Biology \* Physics*, vol. 84, no. 2, pp. e245–e249, 2012.
- [4] D. Pham, C. Xu, and J. Price, "A survey of current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, 2000.
- [5] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [6] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, June 2015.
- [7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in



- Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Munich, Germany, pp. 234–241, October 2015.
- [9] F. Isensee, J. Petersen, A. Klein et al., “nnu-net: selfadapting framework for u-net-based medical image segmentation,” <http://arxiv.org/abs/1809.10486>.
- [10] O. Cicek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: learning dense volumetric segmentation from sparse annotation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Springer, Berlin, Germany, 2016.
- [11] M. Drozdal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, “The importance of skip connections in biomedical image segmentation,” in *Deep Learning and Data Labeling for Medical Applications*, pp. 179–187, Springer, Berlin, Germany, 2016.
- [12] E. Gibson, F. Giganti, Y. Hu et al., “Automatic multi-organ segmentation on abdominal ct with dense v-networks,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 8, pp. 1822–1834, 2018.
- [13] M. P. Heinrich, O. Oktay, and N. Bouteldja, “Obelisk-net: fewer layers to solve 3d multi-organ segmentation with sparse deformable convolutions,” *Medical Image Analysis*, vol. 54, pp. 1–9, 2019.
- [14] X. Chen, H. Fang, T.-Y. Lin et al., “Microsoft coco captions: data collection and evaluation server,” <http://arxiv.org/abs/1504.00325>.
- [15] A. O. Vuola, S. U. Akram, and J. Kannala, “Mask-rcnn and u-net ensemble for nuclei segmentation,” in *Proceedings of the IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 208–212, Venice, Italy, April 2019.
- [16] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–2969, Venice, Italy, October 2017.
- [17] P. K. Sahoo, S. Soltani, and A. K. C. Wong, “A survey of thresholding techniques,” *Computer Vision, Graphics, and Image Processing*, vol. 41, no. 2, pp. 233–260, 1988.
- [18] P. Gibbs, D. L. Buckley, S. J. Blackband, and A. Horsman, “Tumour volume determination from mr images by morphological segmentation,” *Physics in Medicine and Biology*, vol. 41, no. 11, pp. 2437–2446, 1996.
- [19] T. Kapur, W. Eric, L. Grimson, R. Kikinis, and W. M. Wells, “Enhanced spatial priors for segmentation of magnetic resonance imagery,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cambridge, MA, USA, pp. 457–468, October 1998.
- [20] W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, “Adaptive segmentation of mri data,” *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, 1996.
- [21] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, “Fuzzy c-means clustering with spatial information for image segmentation,” *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 9–15, 2006.
- [22] L. O. Hall, A. M. Bensaid, L. P. Clarke, R. P. Velthuizen, M. S. Silbiger, and J. C. Bezdek, “A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain,” *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 672–682, 1992.
- [23] T. Okada, K. Yokota, M. Hori, M. Nakamoto, H. Nakamura, and Y. Sato, “Construction of hierarchical multi-organ statistical atlases and their application to multi-organ segmentation from ct images,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, New York, NY, USA, pp. 502–509, September 2008.
- [24] R. Wolz, C. Chu, K. Misawa, M. Fujiwara, K. Mori, and D. Rueckert, “Automated abdominal multi-organ segmentation with subject-specific atlas generation,” *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1723–1730, 2013.
- [25] T. Kohlberger, M. Sofka, J. Zhang et al., “Automatic multi-organ segmentation using learning based segmentation and level set optimization,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Toronto, Canada, pp. 338–345, September 2011.
- [26] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: a nested u-net architecture for medical image segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11, Springer, Berlin, Germany, 2018.
- [27] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, IEEE, Stanford, CA, USA, pp. 565–571, October 2016.
- [28] Y. Xie, Z. Zhang, M. Sapkota, and L. Yang, “Spatial clockwork recurrent neural network for muscle perimysium segmentation,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Athens, Greece, pp. 185–193, October 2016.
- [29] M. F. Stollenga, W. Byeon, M. Liwicki, and J. Schmidhuber, “Parallel multidimensional lstm, with application to fast biomedical volumetric image segmentation,” in *Advances in Neural Information Processing Systems*, pp. 2998–3006, MIT Press, Cambridge, MA, USA, 2015.
- [30] J. Chen, L. Yang, Y. Zhang, M. Alber, and D. Z. Chen, “Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation,” in *Advances in Neural Information Processing Systems*, pp. 3036–3044, MIT Press, Cambridge, MA, USA, 2016.
- [31] R. P. Poudel, P. Lamata, and G. Montana, “Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation,” in *Reconstruction, Segmentation, and Analysis of Medical Images*, pp. 83–94, Springer, Berlin, Germany, 2016.
- [32] M. Shakeri, S. Tsogkas, E. Ferrante et al., “Sub-cortical brain structure segmentation using f-cnn’s,” in *Proceedings of the IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Prague, Czech Republic, pp. 269–272, April 2016.
- [33] P. F. Christ, M. E. A. Elshaer, F. Ettlinger et al., “Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Athens, Greece, pp. 415–423, October 2016.
- [34] T. Tong, R. Wolz, Z. Wang et al., “Discriminative dictionary learning for abdominal multi-organ segmentation,” *Medical Image Analysis*, vol. 23, no. 1, pp. 92–104, 2015.
- [35] N. Lay, N. Birkbeck, J. Zhang, and S. K. Zhou, “Rapid multi-organ segmentation using context integration and discriminative models,” in *Proceedings of the International Conference on Information Processing in Medical Imaging*, Springer, Asilomar, CA, USA, pp. 450–462, June 2013.
- [36] H. R. Roth, H. Oda, Y. Hayashi et al., “Hierarchical 3D fully convolutional networks for multi-organ segmentation,” <http://arxiv.org/abs/1704.06382>.

- [37] S. Chen, H. Roth, S. Dorn et al., "Towards automatic abdominal multi-organ segmentation in dual energy ct using cascaded 3D fully convolutional network," <http://arxiv.org/abs/1710.05379>.
- [38] X. Dong, Y. Lei, T. Wang et al., "Automatic multiorgan segmentation in thoraxCTimages using U-net-GAN," *Medical Physics*, vol. 46, no. 5, pp. 2157–2168, 2019.
- [39] Z.-H. Wang, Z. Liu, Y.-Q. Song, and Y. Zhu, "Densely connected deep u-net for abdominal multi-organ segmentation," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, IEEE, Taipei, Taiwan, pp. 1415–1419, September 2019.
- [40] Y. Lei, Y. Fu, T. Wang et al., "Deep learning in multi-organ segmentation," <http://arxiv.org/abs/2001.10619>.
- [41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [42] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2117–2125, Honolulu, HI, USA, July 2017.
- [43] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems*, pp. 91–99, MIT Press, Cambridge, MA, USA, 2015.
- [44] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, p. 29, 2015.
- [45] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [46] P. Jaccard, "The distribution of the flora in the alpine Zone.1," *New Phytologist*, vol. 11, no. 2, pp. 37–50, 1912.
- [47] M. G. Linguraru, J. A. Pura, V. Pamulapati, and R. M. Summers, "Statistical 4d graphs for multi-organ abdominal segmentation from multiphase ct," *Medical Image Analysis*, vol. 16, no. 4, pp. 904–914, 2012.
- [48] B. He, C. Huang, and F. Jia, "Fully automatic multi-organ segmentation based on multi-boost learning and statistical shape model search," in *VISCERAL Challenge*, pp. 18–21, ISBI, New York, NY, USA, 2015.
- [49] R. Gauriau, R. Cuingnet, D. Lesage, and I. Bloch, "Multi-organ localization with cascaded global-to-local regression and shape prior," *Medical Image Analysis*, vol. 23, no. 1, pp. 70–83, 2015.

## Research Article

# Construction for the Sequences of Q-Borderenergetic Graphs

Bo Deng <sup>1,2,3,4</sup>, Caibing Chang,<sup>1</sup> Haixing Zhao <sup>2,3,4</sup> and Kinkar Chandra Das <sup>5</sup>

<sup>1</sup>School of Mathematics and Statistics, Qinghai Normal University, Xining 810001, China

<sup>2</sup>Academy of Plateau Science and Sustainability, Xining 810016, China

<sup>3</sup>Key Laboratory of Tibetan Information Processing, Ministry of Education, Xining 810008, China

<sup>4</sup>Tibetan Intelligent Information Processing and Machine Translation Key Laboratory, Qinghai, Xining 810008, China

<sup>5</sup>Department of Mathematics, Sungkyunkwan University, Suwon 16419, Republic of Korea

Correspondence should be addressed to Kinkar Chandra Das; [kinkardas2003@googlemail.com](mailto:kinkardas2003@googlemail.com)

Received 31 May 2020; Accepted 27 June 2020; Published 18 July 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Bo Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This research intends to construct a signless Laplacian spectrum of the complement of any  $k$ -regular graph  $G$  with order  $n$ . Through application of the join of two arbitrary graphs, a new class of Q-borderenergetic graphs is determined with proof. As indicated in the research, with a regular Q-borderenergetic graph, sequences of regular Q-borderenergetic graphs can be constructed. The procedures for such a construction are determined and demonstrated. Significantly, all the possible regular Q-borderenergetic graphs of order  $7 < n \leq 10$  are determined.

## 1. Introduction

All graphs considered in this paper are simple, unweighted, and undirected. Let  $G$  be a graph of order  $n = |V(G)|$ , where  $V(G)$  is the vertex set of  $G$ . The complement of  $G$  is denoted by  $\overline{G}$ . The complete graph of order  $n$  is denoted by  $K_n$ . Denote the average vertex degree of  $G$  by  $\overline{d}$ . The join of two graphs  $H_1$  and  $H_2$  is the graph  $H_1 \nabla H_2$  with the vertex set  $V(H_1) \cup V(H_2)$  and the edge set consisting of all the edges of  $H_1$  and  $H_2$  together with the edges joining each vertex of  $H_1$  with every vertex of  $H_2$ . For details on graph theory and spectral graph theory; see [1–4].

Let  $A(G)$  and  $D(G)$  be the adjacency matrix and the diagonal matrix of the vertex degrees of  $G$ , respectively. Then,  $L(G) = D(G) - A(G)$  and  $Q(G) = D(G) + A(G)$  are called the Laplacian matrix and the signless Laplacian matrix of  $G$ , respectively. In particular, the signless Laplacian spectra of join of two regular graphs are already determined [5].

The energy  $E(G)$  of  $G$  is defined as the sum of the absolute value of the eigenvalues of its adjacency matrix  $A(G)$  [6, 7]. Let  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  be the eigenvalues of  $A$ . Then,

$$E(G) = \sum_{i=1}^n |\lambda_i|. \quad (1)$$

For additional information on graph energy and its applications in chemistry, we refer to [8–10]. The eigenvalues of the Laplacian matrix  $L(G)$  of graph  $G$  are denoted by  $\xi_1 \geq \xi_2 \geq \dots \geq \xi_n = 0$ . The Laplacian energy [11] of  $G$  is defined as

$$LE(G) = \sum_{i=1}^n |\xi_i - \overline{d}|. \quad (2)$$

The eigenvalues of the signless Laplacian matrix  $Q$  of graph  $G$  are denoted by  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_{n-1} \geq \mu_n$ , which forms the signless Laplacian spectrum  $Spec_Q(G)$ . The signless Laplacian energy of  $G$  [12] is defined as  $QE(G) = \sum_{i=1}^n |\mu_i - \overline{d}|$ .

In 2015, Gong et al. [13] proposed the concept of borderenergetic graphs, namely graphs of order  $n$  satisfying  $E(G) = 2(n-1)$ . The corresponding results on borderenergetic graphs can be seen in [14–17]. For the Laplacian energy of a graph  $G$ , Tura [18] proposed the concept of L-borderenergetic graphs, that is, a graph  $G$  of order  $n$  is

*L*-borderenergetic if  $LE(G) = LE(K_n) = 2(n-1)$ . More results on *L*-borderenergetic graphs, we can refer to [18–22].

Recently, Tao and Hou [23] extended this concept to the signless Laplacian energy of a graph. If a graph has the same signless Laplacian energy as the complete graph  $K_n$ , i.e.,  $QE(G) = QE(K_n) = 2(n-1)$ , then it is called *Q*-borderenergetic. In [23, 24], several classes of *Q*-borderenergetic graphs are constructed.

Moreover, in this paper, through using the joint of two graphs, we construct a new class of *Q*-borderenergetic graphs and present a procedure to find sequences of regular *Q*-borderenergetic graphs. Especially, all regular *Q*-borderenergetic graphs of order  $7 < n \leq 10$  are presented. In addition, we obtain the signless Laplacian spectrum of the complement of any *k*-regular graph  $G$  of order  $n$ .

## 2. Construction on *Q*-Borderenergetic Graphs

At first, the signless Laplacian spectrum of the complement of any *k*-regular graph  $G$  with order  $n$  is given in Lemma 1. Denote the signless Laplacian matrix of  $\overline{G}$  by  $\overline{Q}$ .

**Lemma 1.** *Let  $G$  be a  $k$ -regular connected graph of order  $n$ . If  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$  are the eigenvalues of  $Q(G)$ , then the eigenvalues of  $Q(\overline{G})$  are as follows:*

$$\begin{aligned} 2(n-1) - \mu_1 = 2(n-1-k) &\geq n-2 - \mu_n \\ &\geq n-2 - \mu_{n-1} \geq \dots \geq n-2 - \mu_2. \end{aligned} \quad (3)$$

*Proof.* Note that the signless Laplacian matrix of the complement  $\overline{G}$  of  $G$  is written as

$$\begin{aligned} Q(\overline{G}) &= [(n-1)I - D(G)] + (J - A(G) - I) \\ &= (n-2)I + J - Q(G), \end{aligned} \quad (4)$$

where  $I$  is an identity matrix and  $J$  is the matrix with each of whose entries is equal to 1. Since  $G$  is *k*-regular, we have that  $\mu_1 = 2k$  with corresponding eigenvector  $\mathbf{e} = (1, 1, \dots, 1)^T$ . Let  $\mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n$  be the eigenvectors of  $Q(G)$  corresponding to the eigenvalues  $\mu_2, \dots, \mu_{n-1}, \mu_n$ , respectively. Thus, we have  $Q(G)\mathbf{x}_i = \mu_i\mathbf{x}_i$ ,  $i = 2, \dots, n$ . Since  $Q(G)$  is symmetric, all the eigenvectors  $\mathbf{e}, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n$  are orthogonal to each other. Thus, we obtain  $\mathbf{e}^T \mathbf{x}_i = 0$ ,  $i = 2, \dots, n$ . As  $J$  can be presented as

$$J = \begin{pmatrix} \mathbf{e}^T \\ \mathbf{e}^T \\ \mathbf{e}^T \\ \vdots \\ \mathbf{e}^T \end{pmatrix}, \quad (5)$$

it arrives at  $J\mathbf{x}_i = 0$ ,  $i = 2, \dots, n$ . Therefore,

$$\begin{aligned} Q(\overline{G})\mathbf{x}_i &= ((n-2)I + J - Q(G))\mathbf{x}_i \\ &= (n-2)\mathbf{x}_i + J\mathbf{x}_i - Q(G)\mathbf{x}_i \\ &= (n-2 - \mu_i)\mathbf{x}_i, \quad i = 2, \dots, n. \end{aligned} \quad (6)$$

Thus,  $n-2 - \mu_i$  is an eigenvalue with corresponding eigenvector  $\mathbf{x}_i$  of  $Q(\overline{G})$ , where  $i = 2, \dots, n$ . As  $\overline{G}$  is  $(n-1-k)$ -regular,  $2(n-1-k)$  is an eigenvalue with corresponding eigenvector  $\mathbf{e} = (1, \dots, 1)^T$ .

Using Lemma 1, we obtain the signless Laplacian spectrum of the join of two special graphs in the following theorem.  $\square$

**Theorem 1.** *Let  $G_1$  be a  $k$ -regular graph on  $n$  vertices and  $G_2$  be an empty graph on  $n-k$  vertices. If  $2k = \mu_1 \geq \mu_2 \geq \dots \geq \mu_n$  are the signless Laplacian eigenvalues of  $G_1$ , then the signless Laplacian eigenvalues of  $G_1 \nabla G_2$  are*

$$n-k + \mu_2, n-k + \mu_3, \dots, n-k + \mu_n, n^{(n-k-1)}, k, 2n. \quad (7)$$

*Proof.* Note that the join of  $G_1$  and  $G_2$  can also be expressed with

$$G_1 \nabla G_2 = \overline{\overline{G_1} \cup \overline{G_2}}. \quad (8)$$

Since  $2k = \mu_1 \geq \mu_2 \geq \dots \geq \mu_n$  and  $0^{(n-k)}$  are the signless Laplacian eigenvalues of  $G_1$  and  $G_2$ , respectively, by Lemma 1, we have that the signless Laplacian spectra of  $\overline{G_1}$  and  $\overline{G_2}$  are as follows:

$$\begin{aligned} \{n-2 - \mu_n, n-2 - \mu_{n-1}, \dots, n-2 - \mu_2, \\ 2(n-1) - \mu_1 = 2(n-1-k)\}, \\ \{n-k - 2^{(n-k-1)}, 2(n-k-1)\}. \end{aligned} \quad (9)$$

Thus, the set of the signless Laplacian eigenvalues of  $\overline{\overline{G_1} \cup \overline{G_2}}$  is composed of the above two sets. Using Lemma 1, we obtain the signless Laplacian eigenvalues of  $G_1 \nabla G_2$  as follows:

$$n-k + \mu_2, n-k + \mu_3, \dots, n-k + \mu_n, n^{(n-k-1)}, k, 2n. \quad (10)$$

Using Theorem 1, from any *k*-regular *Q*-borderenergetic graph, we can construct a new class of *Q*-borderenergetic graphs in the following theorem.  $\square$

**Theorem 2.** *Let  $G$  be a  $k$ -regular *Q*-borderenergetic graph with  $n$  vertices. Then  $G \nabla \overline{K}_{n-k}$  is *Q*-borderenergetic.*

*Proof.* Let  $2k = \mu_1 \geq \mu_2 \geq \dots \geq \mu_n$  be the signless Laplacian eigenvalues of  $G$ . Since  $G$  is *Q*-borderenergetic, then we have

$$\sum_{i=1}^n |\mu_i - k| = 2n - 2. \quad (11)$$

Let  $p = n - k$ . By Theorem 1, the *Q*-spectrum of  $G \nabla \overline{K}_p$  is

$$\text{Spec}_Q(G \nabla \overline{K}_p) = \left\{ p + \mu_2, p + \mu_3, \dots, p + \mu_n, \underbrace{n, \dots, n}_{n-k-1}, k, 2n \right\}. \quad (12)$$

Since  $p = n - k$ , the average degree  $\bar{d}$  of graph  $G \nabla \overline{K}_p$  is

$$\bar{d} = \frac{nk + 2np}{n+p} = \frac{nk + 2n(n-k)}{n+n-k} = n. \quad (13)$$

By the definition of signless Laplacian energy of a graph with (11), we have

$$\begin{aligned} QE(G\overline{\nabla K}_p) &= \sum_{i=2}^n |\mu_i - k| + |n - n|(n - k - 1) \\ &\quad + |k - n| + |2n - n| \\ &= \sum_{i=1}^n |\mu_i - k| - |\mu_1 - k| + n - k + n \\ &= (2n - 2) - (2k - k) + n - k + n = 2(2n - k - 1). \end{aligned} \tag{14}$$

Since  $|V(G\overline{\nabla K}_p)| = 2n - k$ , from the above result, we conclude that  $G\overline{\nabla K}_p$  is  $Q$ -borderenergetic.  $\square$

### 3. Sequences of $Q$ -Borderenergetic Graphs

In this section, by using Theorem 2 repeatedly, an infinite sequence of  $Q$ -borderenergetic graphs is constructed. Let

$$\begin{aligned} Spec_Q(G^{(s)}) &= \left( \mu_2 + s(n - k), \mu_3 + s(n - k), \dots, \mu_n + s(n - k), \underbrace{n + (s - 1)(n - k), \dots, n + (s - 1)(n - k)}_{s(n - k - 1)}, \right. \\ &\quad \left. n + (s - 2)(n - k), 2n + 2(s - 1)(n - k) \right). \end{aligned} \tag{16}$$

*Proof.* We prove this lemma by mathematical induction on  $s$ . For  $s = 1$ , by Theorem 2, (16) holds. We now assume that the result holds for  $s = t$ . Then we have

$$\begin{aligned} Spec_Q(G^{(t)}) &= \left( \mu_2 + t(n - k), \mu_3 + t(n - k), \dots, \mu_n + t(n - k), \underbrace{n + (t - 1)(n - k), \dots, n + (t - 1)(n - k)}_{t(n - k - 1)}, \right. \\ &\quad \left. n + (t - 2)(n - k), 2n + 2(t - 1)(n - k) \right). \end{aligned} \tag{17}$$

Now, we have  $G^{t+1} = G^t \overline{\nabla K}_{n-k}$ . By Theorem 1, we obtain

$$\begin{aligned} Spec_Q(G^{(t+1)}) &= \left( \mu_2 + (t + 1)(n - k), \mu_3 + (t + 1)(n - k), \dots, \mu_n + (t + 1)(n - k), \underbrace{n + t(n - k), \dots, n + t(n - k)}_{(t+1)(n - k - 1)}, \right. \\ &\quad \left. n + (t - 1)(n - k), 2n + 2t(n - k) \right). \end{aligned} \tag{18}$$

$G^{(0)}$  be any  $k$ -regular  $Q$ -borderenergetic graph with  $n$  vertices. Consider an infinite sequence  $H$  of graphs, i.e.,  $H = \{G^{(0)}, G^{(1)}, \dots, G^{(s)}, \dots\}$  such that

$$\begin{aligned} G^{(1)} &= G^{(0)} \overline{\nabla K}_{n-k}, G^{(2)} = G^{(1)} \overline{\nabla K}_{n-k}, \dots, G^{(s)} \\ &= G^{(s-1)} \overline{\nabla K}_{n-k}, \dots \end{aligned} \tag{15}$$

One can easily see that graph  $G^{(s)}$  ( $s = 1, 2, \dots$ ) is of orders  $n + s(n - k)$  and  $n + (s - 1)(n - k)$ -regular. And the signless Laplacian spectrum of  $G^{(s)}$  is given in the following lemma.

**Lemma 2.** *Let  $G^{(0)}$  be a  $k$ -regular  $Q$ -borderenergetic graph of order  $n$  with signless Laplacian eigenvalues  $2k = \mu_1 \geq \mu_2 \geq \dots \geq \mu_{n-1} \geq \mu_n$ . Then for any  $G^{(s)} \in H$  ( $s \geq 1$ ), the signless Laplacian spectrum of  $G^{(s)}$  is the following:*

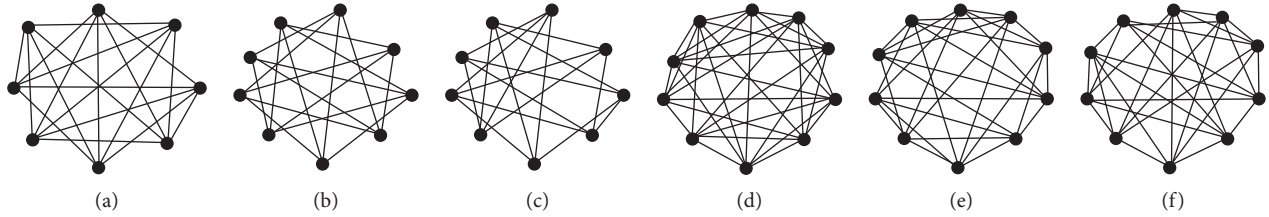


FIGURE 1: The regular  $Q$ -borderenergetic graphs (a)  $G_{8,5}^1$ , (b)  $G_{9,4}^2$ , (c)  $G_{9,4}^3$ , (d)  $G_{10,7}^4$ , (e)  $G_{10,6}^5$ , and (f)  $G_{10,6}^6$ .

Again (16) holds for  $s = t + 1$ . This completes the proof of the lemma.  $\square$

**Theorem 3.** For any  $s \geq 1$ ,  $G^{(s)} \in H$  is  $Q$ -borderenergetic.

*Proof.* Since the graph  $G^{(s)}$  is  $n + (s - 1)(n - k)$ -regular with order  $n + s(n - k)$ , by Lemma 2 and the definition of signless Laplacian energy, we have

$$\begin{aligned} QE(G^{(s)}) &= \sum_{i=2}^n |\mu_i + s(n - k) - n - (s - 1)(n - k)| + s(n - k - 1)|n + (s - 1)(n - k) - n - (s - 1)(n - k)| \\ &\quad + s|n + (s - 2)(n - k) - n - (s - 1)(n - k)| + |2(n + (s - 1)(n - k)) - n - (s - 1)(n - k)| \\ &\quad + \sum_{i=2}^n |\mu_i - k| + s(n - k) + (n + (s - 1)(n - k)) = 2n - 2 - |\mu_1 - k| + n + (2s - 1)(n - k) \\ &= 2n - 2 - k + n + 2s(n - k) - n + k = 2(n + s(n - k) - 1). \end{aligned} \quad (19)$$

Hence,  $G^{(s)}$  is  $Q$ -borderenergetic.

In fact, for a  $k$ -regular graph  $G$ , we can check that the three kinds of energies of  $G$  are equal, i.e.,  $QE(G) = LE(G) = E(G)$ .  $\square$

**Lemma 4.** [20] If  $G$  is a  $k$ -regular graph, then  $LE(G) = E(G)$ .

**Theorem 5.** If  $G$  is a  $k$ -regular graph of order  $n$ , then  $LE(G) = E(G) = QE(G)$ .

*Proof.* Obviously, the average degree of  $G$  is  $k$ . The former equality holds by Lemma 4. Moreover,

$$QE(G) = \sum_{i=1}^n |\mu_i - k| = \sum_{i=1}^n |k + \lambda_i - k| = E(G). \quad (20)$$

This completes the proof of the theorem.

For a  $k$ -regular graph of order  $n$ , if  $G$  is borderenergetic, then  $G$  is  $Q$ -borderenergetic and  $L$ -borderenergetic. In [13], Gong et al. found all the borderenergetic graphs with order  $7 \leq n \leq 9$ . Bearing in mind that there are no noncomplete borderenergetic graphs with order  $n < 7$ . Furthermore, Li et al. [17] searched for the borderenergetic graphs of order 10. Thus, we can find all the regular  $Q$  or  $L$ -borderenergetic graph of order  $n$ ,  $7 \leq n \leq 10$  (Figure 1). Denote the  $i$ -th  $k$ -regular  $Q$ -borderenergetic graph of order  $n$  by  $G_{n,k}^i$ .  $\square$

## Data Availability

The data, cited from the paper [17], used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the NSFQH (No. 2018-ZJ-925Q); NSFC (No. 11701311); and NSFQD (No. 2016A030310307).

## References

- [1] J. A. Bondy and U. S. R. Murty, *Graph Theory*, GTM 244, Springer, Berlin, Germany, 2008.
- [2] D. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Application*, Academic Press, New York, NY, USA, 1980.
- [3] J.-B. Liu, X.-F. Pan, F.-T. Hu, and F.-F. Hu, "Asymptotic Laplacian-energy-like invariant of lattices," *Applied Mathematics and Computation*, vol. 253, pp. 205–214, 2015.
- [4] J.-B. Liu, J. Zhao, and Z.-Q. Cai, "On the generalized adjacency, Laplacian and signless Laplacian spectra of the weighted edge corona networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 540, Article ID 123073, 2020.
- [5] X. Liu and P. Lu, "Signless Laplacian spectral characterization of some joins," *The Electronic Journal of Linear Algebra*, vol. 30, 2015.
- [6] I. Gutman, "Acyclic systems with extremal Huckel  $\pi$ -electron energy," *Theoretica Chimica Acta*, vol. 45, no. 2, pp. 79–87, 1977.
- [7] I. Gutman, "The energy of a graph," *Ber. Math.—Statist. Sect. Forschunz. Graz*, vol. 103, pp. 1–22, 1978.
- [8] I. Gutman, X. Li, and J. Zhang, "Graph energy," in *Analysis of Complex Networks. From Biology to Linguistics*, M. Dehmer



- and F. Emmert-Streib, Eds., Wiley-VCH, Weinheim, Germany, pp. 145–174, 2009.
- [9] X. Li, Y. Shi, and I. Gutman, *Graph Energy*, Springer, New York, NY, USA, 2012.
- [10] B. Zhou, “Energy of a graph,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 51, pp. 111–118, 2004.
- [11] I. Gutman and B. Zhou, “Laplacian energy of a graph,” *Linear Algebra and Its Applications*, vol. 414, no. 1, pp. 29–37, 2006.
- [12] D. Cvetkovic and S. Simic, “Towards a spectral theory of graphs based on the signless Laplacian, I,” *Publications de l’Institut Mathematique*, vol. 85, no. 99, pp. 19–33, 2009.
- [13] S. Gong, X. Li, G. Xu, I. Gutman, and B. Furtula, “Borderenergetic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 74, pp. 321–332, 2015.
- [14] B. Deng, X. Li, and I. Gutman, “More on borderenergetic graphs,” *Linear Algebra and Its Applications*, vol. 497, pp. 199–208, 2016.
- [15] Y. Hou and Q. Tao, “Borderenergetic threshold graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 75, pp. 253–262, 2016.
- [16] X. Li, M. Wei, and X. Zhu, “Borderenergetic graphs with small maximum or large minimum degrees,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 77, pp. 25–36, 2016.
- [17] X. Li, M. Wei, and S. Gong, “A computer search for the borderenergetic graphs of order 10,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 74, pp. 333–342, 2015.
- [18] F. Tura, “ $L$ -borderenergetic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 77, pp. 37–44, 2017.
- [19] B. Deng, X. Li, and J. Wang, “Further results on  $L$ -borderenergetic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 77, pp. 607–616, 2017.
- [20] B. Deng and X. Li, “More on  $L$ -borderenergetic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 77, pp. 115–127, 2017.
- [21] L. Lu and Q. Huang, “On the existence of non-complete  $L$ -borderenergetic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 77, pp. 625–634, 2017.
- [22] Q. Tao and Y. Hou, “A computer search for the  $L$ -borderenergetic graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 77, pp. 595–606, 2017.
- [23] Q. Tao and Y. Hou, “ $Q$ -borderenergetic graphs,” *AKCE International Journal of Graphs and Combinatorics*, 2018, In press.
- [24] B. Deng, X. Li, and Y. Li, “(Signless) Laplacian borderenergetic graphs and the join of graphs,” *MATCH Communications in Mathematical and in Computer Chemistry*, vol. 80, pp. 449–457, 2018.

## Research Article

# Research on Mathematical Model of Cost Budget in the Early Stage of Assembly Construction Project Based on Improved Neural Network Algorithm

Xin Lin <sup>1</sup> and Yinan Lu <sup>2</sup>

<sup>1</sup>School of Urban Construction Engineering, Chongqing Radio & TV University, Chongqing, China

<sup>2</sup>School of Information Engineering, Nanchang University, Nanchang, China

Correspondence should be addressed to Xin Lin; [lxqrtvu\\_edu@126.com](mailto:lxqrtvu_edu@126.com)

Received 2 May 2020; Accepted 25 June 2020; Published 15 July 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Xin Lin and Yinan Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In view of the poor performance of the original mathematical model of assembly construction project precast budget, a mathematical model of assembly construction project precast budget based on improved neural network algorithm is proposed. This paper investigates the cost content of assembly construction project and analyzes its early cost. It finds that the early cost of assembly construction project includes component production cost, transportation component cost, and installation component cost. Based on the improved neural network algorithm to build an improved neural network model, the improved neural network model to mine the cost data in the early stage of assembly construction project is used. In this paper, the earned value variable is introduced to transform the project duration and project cost in the early stage of the prefabricated construction project into quantifiable cost data, and the earned value analysis method is used to estimate the implementation cost of the prefabricated construction project. According to the result of cost estimation, the mathematical model of precast budget of prefabricated construction project is built based on the project parameters. In order to prove that the cost budget performance of the mathematical model based on the improved neural network algorithm in the early stage of assembly construction project is better, the original mathematical model is compared with the mathematical model, the experimental results show that the cost budget performance of the model is better than the original model, and the cost budget performance is improved.

## 1. Introduction

In traditional construction projects, the pouring of concrete mainly adopts manual scaffolding, formwork supporting, and binding of steel bars at site. However, such kind of cast-in-place can cause a variety of problems, including making construction site in a mess, generating lots of construction waste, and polluting the surrounding environment [1]. Meanwhile, with the progress of population aging and the increase of labor costs in China, this extensive construction model is no longer suitable for the green, energy-saving, and environmentally friendly concepts. Under this circumstance, the prefabricated construction project emerged at the right moment and eased this dilemma. As a new building construction technology, prefabricated construction project

means to process and produce the prefabricated components in factory, move them to the construction site after maintenance, and then assemble the building components accordingly through machinery equipment to achieve the building functional requirements [2]. Compared with the cast-in-place construction method, the prefabricated construction project can save about 20% of materials and 80% of water resources in manufacturing and processing building components. At the same time, in prefabricated construction project, protective nets and scaffolding are not used during construction, which reduced construction waste and lowered the pollution and harm to the environment [3].

Under the trend of urbanization, people have put forward higher requirements for environmental protection and energy-saving performance of buildings. Therefore,

prefabricated buildings are getting popular, have achieved rapid development in China, and promoted the improvement of the installation quality and component precision of prefabricated components [4]. However, based on the current construction market, the overall production scale of prefabricated buildings is restricted so that it is unable to reduce the construction cost budget through expanding the scale, which limited the development of prefabricated buildings in China and affected the industrialization process of construction field. At its initial stage of development, prefabricated buildings in China are still facing many shortcomings in technology, experience, and cost control [5]. In order to accelerate the development of prefabricated buildings, an innovative research on the mathematical model of cost budget in early stage of prefabricated construction is carried out. On this basis, a mathematical model of cost budget in the early stage of prefabricated construction project based on improved neural network algorithm is proposed [6].

## 2. Design of Mathematical Model of Cost Budget in the Early Stage of Prefabricated Construction Project Based on Improved Neural Network Algorithm

*2.1. Analysis of Cost in the Early Stage of Prefabricated Construction Project.* As found in the cost analysis in early stage of the prefabricated construction project, the preliminary cost of prefabricated construction project includes component production cost, component transportation cost, and component installation cost. Among them, statistics found that component production cost contains labor cost, material cost, mold cost, amortization expense, cost of setting the embedded parts and pipelines, management and storage cost, and water and electricity charges. The specific contents are shown in Table 1 [7].

Component transportation cost covers the cost of transporting components from the factory to the construction site, which is directly related to the size and weight of the components, as well as the distance between the factory and the construction site [8].

Component installation costs involves the cost of vertical component transportation, labor cost of component installation, machinery cost of component installation, material cost of component installation, cast-in-place cost, and amortization cost, as shown in Table 2 [9].

*2.2. Data Mining of Cost in the Early Stage of Prefabricated Construction Project.* As shown in the analysis of cost budget in early stage of prefabricated construction project, an improved neural network model is built based on the improved neural network algorithm so as to conduct data mining on the cost budget in early stage of prefabricated construction project [10]. The specific steps are as follows:

- (1) First, the improved neural network is initialized:  $(X, Y)$  is adopted to represent the input and output sequence of the improved neural network. Based on

this sequence, the specific number of nodes corresponding to the output layer, hidden layer, and input layer of the improved neural network is clarified, which is  $m$ ,  $l$ , and  $n$  nodes, respectively. Next, the threshold and the connection weight are initialized. The neuron connection weight between the hidden layer and the input layer is set to  $\omega_{ij}$ ; the neuron connection weight between the output layer and the hidden layer is set to  $\omega_{jk}$ ; the thresholds of the output layer and hidden layer is  $b$  and  $a$ , respectively. Finally, the learning rate of the improved neural network is set to  $\eta$ , and the excitation neuron function is set to  $f(x)$  [11].

- (2) The specific output of the hidden layer is computed: the specific output of the hidden layer is obtained based on the specific input variable  $X$  of the improved neural network, the threshold value  $a$  of the hidden layer, and the neuron connection weight  $\omega_{ij}$  between the hidden layer and the input layer. The details are as follows:

$$\left\{ H_j = f\left(\sum_{i=1}^n \omega_{ij}x_i - a_j\right), \quad j = 1, 2, 3, \dots, l \quad (1)$$

In formula 1,  $H_j$  represents the specific output of the hidden layer,  $x_i$  represents the  $i$ th input value,  $a_j$  represents the threshold of the  $j$ th hidden layer, and  $l$  represents the specific number of nodes of the hidden layer [12].

- (3) The specific output of the input layer is computed: the specific output of the input layer is obtained according to the specific output  $H_j$  of the hidden layer, the threshold value  $b$  of the output layer, and the neuron connection weight  $\omega_{jk}$  between the output layer and the hidden layer:

$$\left\{ O_k = \sum_{i=1}^l H_j \omega_{jk} - b_k, \quad k = 1, 2, 3, \dots, m \quad (2)$$

In formula 2,  $O_k$  represents the specific output of the input layer and  $b_k$  represents the threshold of the  $k$ th output layer [13].

- (4) Error calculation: the specific error of the improved neural network is predicted based on the specific output of the input layer and the specific expected output  $Y$  of the improved neural network.
- (5) The neuron connection weight  $\omega_{ij}$  between the hidden layer and the input layer, and the neuron connection weight  $\omega_{jk}$  between the output layer and the hidden layer are updated according to the specific predicted error of the improved neural network.
- (6) Threshold updating: the thresholds  $b$  and  $a$  of the output layer and hidden layer are updated according to the specific predicted error of the improved neural network.
- (7) The construction of improved neural network model is realized and the data mining on the cost in early

TABLE 1: Component production cost.

No.	Name of cost	Content of cost
1	Labor cost of component production	Higher salaries shall be paid to professional workers
2	Material cost of component production	Basically the same materials as required by the traditional construction method
3	Mold cost of component production	Cost of table molding, binding steel bar mold, concrete pouring mold, maintenance mold, and finished component mold
4	Amortization expense	Amortization fee based on specific types of components and specific quantities of molds
5	Cost of setting the embedded parts and pipelines	Costs incurred in arranging the embedded parts and pipelines in the installation components, mainly the pipeline costs
6	Management and storage cost	Additional management and storage costs after maintenance of component productions
7	Water and electricity charges	Electricity and water charges incurred by factory component production

TABLE 2: Component installation cost.

No.	Name of cost	Content of cost
1	Cost of vertical component transportation	Cost of vertically hoisting components
2	Labor cost of component installation	Higher labor salaries shall be paid because the vertical hoisting of components requires higher professionalism and proficiency.
3	Machinery cost of component installation	The cost generated by using machinery equipment during component installation
4	Material cost of component installation	Costs incurred by filling materials and connectors
5	Cast-in-place cost	Cast-in-place cost for assembly
6	Amortization cost	Amortization cost of tools

stage of prefabricated construction project is performed on this basis [14].

2.3. *Cost Budget in Early Stage of Prefabricated Construction Project.* Based on the data mining of cost in early stage of prefabricated construction project, a variable called earned value is introduced to convert the project duration and project cost in early stage of prefabricated construction into quantifiable cost data. Meanwhile, the cost in early stage of prefabricated construction project is estimated through the earned value analysis so as to carry out the cost budget of prefabricated construction project [15]. The calculation formula of earned value is as follows:

$$EV = A \times B. \tag{3}$$

In formula 3, *EV* represents earned value, *A* represents the project actual workload and *B* represents the cost budget of the completed project.

For the analysis by using the earned value, the difference variables of two analyzes must be obtained at first, including schedule deviation and cost deviation. Their calculation formulas are separately as follows:

$$SV = B - BSWS. \tag{4}$$

In formula 4, *SV* represents the schedule deviation and *BSWS* represents the cost budget of planned project volume.

$$CV = B - ACWP. \tag{5}$$

In formula 5, *CV* represents cost deviation and *ACWP* represents the specific cost of the completed project volume [16].

Along with two variable indexes, including performance progress index and performance cost index, the calculation formulas are as follows:

$$SPI = \frac{B}{BCWS}. \tag{6}$$

In formula 6, *SPI* represents the performance progress index:

$$CPI = \frac{B}{ACWP}. \tag{7}$$

In formula 7, *CPI* represents the performance cost index.

The data of the cost budget of completed project volume, the specific cost of completed project volume, and the cost budget of planned project volume are added separately [17] in analysis so that 3 corresponding cumulative series are obtained. By inputting data of cost budget of completed project volume, the specific cost of completed project volume and the cost budget of planned project volume into a two-dimensional coordinate axis of time and cost, and 3 analysis curves are obtained and applied to analyze the period and cost in early stage of prefabricated construction project. Among them, when the cost deviation is greater than 0, it indicates that the early stage of prefabricated construction project is in a cost-saving state; when the cost deviation is less than 0, it indicates that the early stage of prefabricated construction project is in the over-cost state; when the progress deviation is greater than 0, it indicates that the early stage of prefabricated construction project is in a state of advanced progress; when the progress deviation is less than 0, it indicates that the early stage of prefabricated construction project is in a state of delayed progress. The cost

of prefabricated buildings can be estimated by inputting the construction period and cost analysis results of early stage of the prefabricated construction project as well as the actual status of the early stage of project into the project management software.

**2.4. Mathematical Model of Cost Budget in Early Stage of Prefabricated Construction Project.** According to the estimated cost in the early stage of prefabricated construction project, the mathematical model of cost budget in early stage of prefabricated construction project is constructed based on project parameters. The items of material budget in early stage of prefabricated construction project are shown in Table 3.

The project parameters are described according to the type of project, in which the parameter of production progress in the early stage of prefabricated construction project is set to  $r_g$ , the parameter of production calculation period is set to  $T$ , the parameter of periodic component

production batch is set to  $N$ , labor demand cost is set to  $L(P)$ , machinery demand cost is set to  $L(c)$ , other expenses such as management fee are set to  $K$ , demand machinery value is set to  $F_L$ , site cost is set to  $A_C$ , equipment rental cost is set to  $F(c)$ , the upper limit of transportation cost is set to  $P(c)$ , and the assembly cost is set to  $A_{sum}$ .

The production process in early stage of prefabricated construction project can be summarized as follows: the first if component production, followed by component assembly and transportation. In this way, the cost in early stage of prefabricated construction project includes construction production cost, component assembly, and transportation fee [19]. Therefore, by setting parameters of production progress in early stage of prefabricated construction project as independent variables and the minimum budget cost as the model objective function, then the mathematical model of cost budge in the early stage of prefabricated construction projects is established as follows [20]:

$$\begin{cases} r_g = \frac{T}{N}, \\ \text{MinCost} = A(c) \int_0^T L(p)dt + L(c) \int_0^T Adt + F(L) \int_0^T F(c)dt + K \cdot P(c) \cdot A_{sum}. \end{cases} \quad (8)$$

In formula 8, MinCost represents the minimum cost budget in the early stage.

### 3. Experimental Research and Result Analysis

**3.1. Experiment Design.** The experiment of cost budget in early stage of prefabricated construction project is carried out by using the mathematical model of cost budget in early stage of prefabricated construction project designed based on improved neural network algorithm. With a total area of 920 square meters, the prefabricated construction project in this experiment contains 14 floors in prefabricated structures of assembled shear wall. The prefabricated component nodes are manufactured by secondary cast-in-place. The outer wall of the building is made of prefabricated thermal insulation Sandwich panel; the floor is made of concrete prestressed composite slab; the load-bearing wall is made of shear prefabricated wall panel; the staircase is prefabricated; the inner partition wall is made of lightweight wall panel. The prefabricated parts involved in this prefabricated construction project are listed as follows: prefabricated parts for stair, prefabricated parts for laminated panel, prefabricated parts for partition wall, shear wall, etc. Considering that it is a prefabricated construction project, all the prefabricated parts are produced at the prefabricated production base and then transported to the construction site after maintenance. Meanwhile, the transportation and production of prefabricated parts for different places and floors are arranged separately according to the specific project progress.

The specific building parameters of this prefabricated building project are shown in Table 6.

In this prefabricated construction project, all the floors are standard, and the early stage is set to a six-day construction period per floor. According to the building structure and engineering quantity, the construction sequence is arranged reasonably and the whole project is divided into three phases. The specific arrangements for the construction of each floor are as follows: hoisting 60 pieces of wall in the first construction phase, hoisting 60 pieces of wall in the second construction phase, grouting sleeve in the first construction phase, grouting sleeve in the second construction phase, hoisting 30 pieces of wall in the third construction phase, grouting sleeves in the third construction phase, binging steel bars with postcasting belts, reinforcing formwork with postcasting belts, erecting supportive frames, hoisting composite beams, hoisting stairs, hoisting of 60 pieces of composite slabs, hoisting 57 pieces of composite slabs, preburying hydropower and other pipelines, binding the upper stair reinforcement, supporting formwork joints, and pouring concrete. The cost budget in early stage of this prefabricated building construction is estimated through the mathematical model. In order to ensure the effectiveness and contrast of this experiment, the original mathematical model of cost budget in the early stage of prefabricated construction project is compared to the mathematical model of cost budget in the early stage of prefabricated construction project designed based on the improved neural network algorithm in this paper. Among



TABLE 3: Items of material budget in early stage of prefabricated construction project.

No.	Items of material budget	Unit
1	Prefabricated PC wall components	Cubic meter
2	Prefabricated PC floor (laminated) components	Cubic meter
3	Prefabricated PC stair components	Cubic meter
4	Prefabricated PC balcony components	Cubic meter
5	Prefabricated PC air conditioning panel components	Cubic meter
6	Prefabricated PC beam components	Cubic meter
7	Rebar B300HR(D < 25)	Ton
8	Screw-thread steel B225HR(D < 12)	Ton
9	Concrete C30 (premixed)	Cubic meter
10	Concrete C35 (premixed)	Cubic meter
11	Fine stone concrete C20 (premixed)	Cubic meter
12	Aerated light sand autoclaved block concrete	Cubic meter
13	Specialized component grouting	Ton
14	Remaining self-purchased materials (estimated value)	—

The items of labor cost budget in early stage of prefabricated construction project are shown in Table 4 [18].

TABLE 4: Items of labor cost budget in early stage of prefabricated construction project.

No.	Items of labor cost	Unit
1	Manual cleaning of foundation pit	Cubic meter
2	Manual lashing of cast-in-place and steel bar making	Ton
3	Manual assembly and wooden template making	Cubic meter
4	Manual maintenance of cast-in-place and concrete pouring	Cubic meter
5	Manual installation of prefabricated PC components	Cubic meter
6	Tower crane driver, surveyor, and bell man	Cubic meter
7	Manual masonry	Cubic meter
8	Manual water resistance	Cubic meter
9	Manual fitment	Cubic meter

The items of machinery budget in early stage of prefabricated construction project are shown in Table 5.

TABLE 5: Items of machinery budget in early stage of prefabricated construction project.

No.	Name	Unit	Type
1	(Tower) crane	Set	ST50/20
2	Car crane	Set	26T
3	Material hoist	Set	SES160

TABLE 6: Specific building parameters of this prefabricated building project.

No.	Name of building parameters	Specific parameters
1	Height	40 m
2	Floors	14
3	Structure type	Prefabricated structure of assembled shear wall
4	Specific floor height	3.6 M at 1 <sup>st</sup> floor 2.8 M from 2 <sup>nd</sup> to 14 <sup>th</sup> floor
5	Covered area	920 m <sup>2</sup>
6	Seismic grade	4.0 magnitude
7	Seismic intensity	Scale 6
8	Specific flame resistance	Level 2
9	Specific waterproof rating	Level 2
11	Specific type of prefabricated part	Shear wall Partition wall Composite slab Platforms, ladder beams, stairs
12	Building nature	Residence
13	Plan view size	15 × 40
14	Waterproof condition (roofing)	Composite SBS waterproof
15	Production materials such as windows and doors	High quality aluminum alloy

The specific contracting situation of this prefabricated construction project is shown in Table 7.



TABLE 7: Specific contracting situation of this prefabricated construction project.

No.	Items	Specific situation
1	Specific project name	Project of Fuligang building #2
2	Construction company	Haoqiang Real Estate co., ltd
3	Design contractor	Yicai Architectural Design co., ltd
4	Supervision company	Keli Supervision Construction co., ltd
5	Construction organization	Risheng Engineering Cconstruction co., ltd
6	Quality supervision organization	Yihe Engineering Quality Supervision co., ltd
7	Total project cost	RMB 12.6 million
8	Contracted form	Contract for labor and material
9	Contract scope	Decoration, main body, foundation
10	Planned schedule	One year
11	Overall quality goal	Good in quality

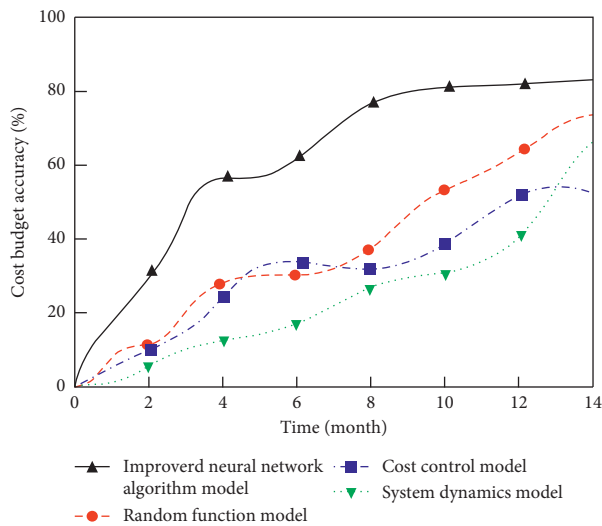


FIGURE 1: Result of comparative experiment on cost budget performance.

them, the original mathematical model of cost budget in early stage of prefabricated construction project includes the mathematical model of cost budget in early stage of prefabricated construction project based on cost control, random function, and system dynamics. As learned from comparing the performance of different mathematical models of cost budget in early stage of prefabricated construction project, that is, from analyzing the cost budget accuracy of the experimental model, the higher the cost budget accuracy, the more reasonable the cost budget result of the prefabricated construction project and the better cost budget performance [21].

**3.2. Analysis Results.** The result of comparative experiment on cost budget performance between the original mathematical model of cost budget in early stage of prefabricated construction project and the mathematical model of cost budget in early stage of prefabricated construction project designed based on the improved neural network algorithm is shown in Figure 1.

According to the result of comparative experiment on cost budget performance between two different mathematical models, the mathematical model of cost budget in

early stage of prefabricated construction project designed based on the improved neural network algorithm is better in cost budget accuracy. That means the cost budget performance of the mathematical model of cost budget in early stage of prefabricated construction project based on the improved neural network algorithm is superior to that of the original mathematical model of cost budget in early stage of prefabricated construction project.

## 4. Conclusions

The cost budget performance of the mathematical model of cost budget in the early stage of prefabricated construction project based on the improved neural network algorithm can realize the improvement of cost budget performance in the early stage of prefabricated construction project, which is of great reference significance to accurate cost budget of the overall prefabricated construction project.

## Data Availability

Simulation data and our model and related hyperparameters used are provided within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the Chongqing Municipal Education Commission Science and Technology Research Project (KJQN201904003) and research project of Chongqing Technology and Business Institute (ZD2016-04).

## References

- [1] G. E. Gurcanli, S. Bilir, and M. Sevim, "Activity based risk assessment and safety cost estimation for residential building construction projects," *Safety Science*, vol. 80, pp. 1–12, 2015.
- [2] Y. T. Chae, R. Horesh, Y. Hwang, and Y. M. Lee, "Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings," *Energy and Buildings*, vol. 111, pp. 184–194, 2016.
- [3] M.-Y. Cheng, H.-C. Tsai, and E. Sudjono, "Conceptual cost estimates using evolutionary fuzzy hybrid neural network for

- projects in construction industry,” *Expert Systems with Applications*, vol. 37, no. 6, pp. 4224–4231, 2010.
- [4] K. H. Hyari, A. Al-Daraiseh, and M. El-Mashaleh, “Conceptual cost estimation model for engineering services in public construction projects,” *Journal of Management in Engineering*, vol. 32, Article ID 04015021, 2016.
- [5] B. Pal, A. Mhashilkar, A. Pandey, B. Nagphase, and V. Chandanshive, “Cost estimation model (CEM) of buildings by ANN (artificial neural networks)—A review,” *Neural Networks*, vol. 5, pp. 1–15, 2018.
- [6] J. Liu, X. Li, D. Wu, and J. Dong, “Cost estimation of building individual cooperative housing with crowdfunding model: case of Beijing, China,” *Journal of Intelligent Manufacturing*, vol. 28, no. 3, pp. 749–757, 2017.
- [7] H. Piili, A. Happonen, T. Väistö, V. Venkataramanan, J. Partanen, and A. Salminen, “Cost estimation of laser additive manufacturing of stainless steel,” *Physics Procedia*, vol. 78, pp. 388–396, 2015.
- [8] O. Tatari and M. Kucukvar, “Cost premium prediction of certified green buildings: a neural network approach,” *Building and Environment*, vol. 46, no. 5, pp. 1081–1086, 2011.
- [9] S. C. Lhee, I. Flood, and R. R. Issa, “Development of a two-step neural network-based model to predict construction cost contingency,” *Journal of Information Technology in Construction (ITcon)*, vol. 19, pp. 399–411, 2014.
- [10] V. Chandanshive and A. R. Kambekar, “Estimation of building construction cost using artificial neural networks,” *Journal of Soft Computing in Civil Engineering*, vol. 3, pp. 91–107, 2019.
- [11] J. A. Rodger, “A fuzzy nearest neighbor neural network statistical model for predicting demand for natural gas and energy cost savings in public buildings,” *Expert Systems with Applications*, vol. 41, no. 4, pp. 1813–1829, 2014.
- [12] G. Ngowtanasuwan, “Mathematical model for optimization of construction contracting in housing development project,” *Procedia-Social and Behavioral Sciences*, vol. 105, pp. 94–105, 2013.
- [13] D.-K. Bui, T. Nguyen, J.-S. Chou, H. Nguyen-Xuan, and T. D. Ngo, “A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete,” *Construction and Building Materials*, vol. 180, pp. 320–333, 2018.
- [14] E. Asadi, M. G. d. Silva, C. H. Antunes, L. Dias, and L. Glicksman, “Multi-objective optimization for building retrofit: a model using genetic algorithm and artificial neural network and an application,” *Energy and Buildings*, vol. 81, pp. 444–456, 2014.
- [15] M. Ceylan, M. H. Arslan, R. Ceylan, M. Y. Kaltakci, and Y. Ozbay, “A new application area of ANN and ANFIS: determination of earthquake load reduction factor of pre-fabricated industrial buildings,” *Civil Engineering and Environmental Systems*, vol. 27, no. 1, pp. 53–69, 2010.
- [16] A. Nasirian, M. Arashpour, B. Abbasi, and A. Akbarnezhad, “Optimal work assignment to multiskilled resources in pre-fabricated construction,” *Journal of Construction Engineering and Management*, vol. 145, pp. 4019–4034, 2019.
- [17] H. Quan, D. Srinivasan, and A. Khosravi, “Particle swarm optimization for construction of neural network-based prediction intervals,” *Neurocomputing*, vol. 127, pp. 172–180, 2014.
- [18] H.-L. Yip, H. Fan, and Y.-H. Chiang, “Predicting the maintenance cost of construction equipment: comparison between general regression neural network and Box-Jenkins time series models,” *Automation in Construction*, vol. 38, pp. 30–38, 2014.
- [19] J. Sobhani, M. Najimi, A. R. Pourkhorshidi, and T. Parhizkar, “Prediction of the compressive strength of no-slump concrete: a comparative study of regression, neural network and ANFIS models,” *Construction and Building Materials*, vol. 24, no. 5, pp. 709–718, 2010.
- [20] R. Sonmez, “Range estimation of construction costs using neural networks with bootstrap prediction intervals,” *Expert Systems with Applications*, vol. 38, no. 8, pp. 9913–9917, 2011.
- [21] A. O. Elfaki, S. Alatawi, and E. Abushandi, “Using intelligent techniques in construction project cost estimation: 10-year survey,” *Advances in Civil Engineering*, vol. 2014, pp. 1023–1031, 2014.

## Research Article

# Multiple Positive Solutions for Fractional Three-Point Boundary Value Problem with $p$ -Laplacian Operator

Dong Li,<sup>1</sup> Yang Liu,<sup>2</sup> and Chunli Wang<sup>3</sup>

<sup>1</sup>Department of Mathematics, College of Science Jiamusi University, Jiamusi, Heilongjiang 154007, China

<sup>2</sup>School of Mathematics and Statistics, Hefei Normal University, Hefei, Anhui 230061, China

<sup>3</sup>Institute of Information Technology of GUET, Guilin, Guangxi 540004, China

Correspondence should be addressed to Chunli Wang; wangchunliwcl821222@sina.com

Received 23 May 2020; Accepted 20 June 2020; Published 9 July 2020

Academic Editor: Jia-Bao Liu

Copyright © 2020 Dong Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we investigate the existence of multiple positive solutions or at least one positive solution for fractional three-point boundary value problem with  $p$ -Laplacian operator. Our approach relies on the fixed point theorem on cones. The results obtained in this paper essentially improve and generalize some well-known results.

## 1. Introduction

Nowadays, fractional calculus has been adapted to numerous fields, such as engineering, mechanics, physics, chemistry, and biology. Many essays and monographs studying various issues in fractional calculus have been researched (see [1–6]). In particular, fractional differential equations have been found to be a powerful tool in modeling various phenomena in many areas of science and engineering such as physics, fluid mechanics, and heat conduction. More details about research achievement on fractional differential equations and their applications are shown in [7–10].

Recently, fractional differential equations have gained considerable attention (see [11–15] and the references therein). Fractional differential equations and differential equations with  $p$ -Laplacian operators have attracted much attention from many mathematicians. As a result, meaningful research results have been drawn [16–20]. Fractional-order boundary value problems involving classical, multipoint, high-order, and integral boundary conditions have extensively been studied by many researchers and a variety of results can be found in recent literature on the topic [21–25].

In [15], Chai studied the boundary value problems of fractional differential equations with  $p$ -Laplacian operator as follows:

$$\begin{cases} D_{0^+}^\beta (\phi_p(D_{0^+}^\alpha u(t))) + f(t, u(t)) = 0, & 0 < t < 1, \\ u(0) = 0, u(1) + \sigma D_{0^+}^\gamma u(1) = 0, & D_{0^+}^\alpha u(0) = 0, \end{cases} \quad (1)$$

where  $\phi_p(s) = |s|^{p-2}s$ ,  $p > 1$ ,  $f \in C((0, 1) \times R^+, R^+)$ ,  $\alpha \in (1, 2]$ ,  $\beta \in (0, 1]$ ,  $\gamma \in (0, 1]$ ,  $\alpha - \gamma \geq 1$ ,  $\sigma > 0$ , and  $D_{0^+}^\alpha$ ,  $D_{0^+}^\beta$ ,  $D_{0^+}^\gamma$  are the standard Riemann–Liouville derivatives. Some existence results of positive solutions are obtained by using the monotone iterative method.

In [16], by using Krasnosel'skii's fixed point theorem, Tian et al. obtained the existence of positive solutions for a boundary value problem of fractional differential equations with  $p$ -Laplacian operator as follows:

$$\begin{cases} D_{0^+}^\beta (\phi_p(D_{0^+}^\alpha u(t))) = f(t, u(t)), & 0 < t < 1, \\ u(0) = u'(0) = u(1) = D_{0^+}^\alpha u(0) = 0, & D_{0^+}^\alpha u(1) = \lambda D_{0^+}^\alpha u(\eta), \end{cases} \quad (2)$$

where  $\phi_p(s) = |s|^{p-2}s$ ,  $p > 1$ ,  $f \in C([0, 1] \times R^+, R^+)$ ,  $\alpha \in (2, 3]$ ,  $\beta \in (1, 2]$ ,  $\eta \in (0, 1)$ ,  $\lambda \in [0, \infty)$ , and  $D_{0^+}^\alpha$ ,  $D_{0^+}^\beta$  are the Riemann–Liouville fractional derivatives.

In [17], by using the monotone iterative method, Tian et al. obtained the existence of positive solutions for a boundary value problem of fractional differential equations with  $p$ -Laplacian operator as follows:

$$\begin{cases} D^\gamma(\phi_p(D^\alpha u(t))) = f(t, u(t)), & 0 < t < 1, \\ u(0) = D^\alpha u(0) = 0, D^\beta u(1) = aD^\beta u(\xi), \quad D^\alpha u(1) = bD^\alpha u(\eta), \end{cases} \quad (3)$$

where  $\phi_p(s) = |s|^{p-2}s$ ,  $p > 1$ ,  $\alpha \in (1, 2]$ ,  $0 < \beta \leq \alpha - 1$ ,  $\xi, \eta \in (0, 1)$ ,  $a, b \in [0, \infty)$  and  $1 - a\xi^{\alpha-\beta-1} > 0$ ,  $1 - b^{p-1}\eta^{p-1} > 0$ ,  $f \in C([0, 1] \times R^+, R^+)$ , and  $D^\alpha$  is the Riemann–Liouville fractional derivative.

In [18], by means of the  $p$ -Laplacian operator, Han et al. obtained the existence of positive solutions for the boundary value problem of fractional differential equation as follows:

$$\begin{cases} D_{0^+}^\beta(\phi(D_{0^+}^\alpha u(t))) = f(t, u(t)), & 0 < t < 1, \\ u(0) = u'(0) = 0, \quad D_{0^+}^\mu u(1) = \delta D_{0^+}^\mu u(\eta), \phi(D_{0^+}^\alpha u(0)) = (\phi(D_{0^+}^\alpha u(1)))' = 0, \end{cases} \quad (5)$$

where  $\phi(s) = |s|^{p-2}s$ ,  $p > 1$ ,  $\alpha \in (2, 3]$ ,  $\beta \in (1, 2]$ ,  $\mu \in [1, \alpha - 1]$ ,  $f \in C([0, 1] \times R^+, R^+)$ ,  $\delta \geq 0$ ,  $0 < \eta < 1$ , and  $\Delta = 1 - \delta\eta^{\alpha-\mu-1} > 0$ .

The aim is to establish some existence and multiplicity results of positive solutions for BVP (5). This paper is organized as follows. In Section 2, some properties of Green’s function will be given, which are needed later. In Section 3, the existence of multiplicity results of positive solutions of BVP will be discussed (5).

## 2. Preliminary Knowledge and Lemmas

**Lemma 1** (see [14]). *Assume that  $D_{a^+}^\alpha \in L^1(a, b)$  with a fractional derivative of order  $\alpha > 0$ . Then,*

$$I_{a^+}^\beta D_{a^+}^\beta u(t) = u(t) + c_1(t-a)^{\alpha-1} + c_2(t-a)^{\alpha-2} + \dots + c_n(t-a)^{\alpha-n}, \quad (6)$$

for some  $c_i \in R, i = 1, 2, \dots, n$ , where  $n$  is the smallest integer greater than or equal to  $\alpha$ .

**Lemma 2.** *If  $y \in C[0, 1]$ , then the fractional boundary value problem is*

$$\begin{cases} D_{0^+}^\alpha u(t) + y(t) = 0, & 0 < t < 1, \\ u(0) = u'(0) = 0, \quad D_{0^+}^\mu u(1) = \delta D_{0^+}^\mu u(\eta). \end{cases} \quad (7)$$

The unique solution is  $u(t) = \int_0^1 G(t, s)y(s)ds$ , where  $G(t, s) = G_1(t, s) + (\delta/\Delta)t^{\alpha-1}G_2(\eta, s)$ ,

$$G_1(t, s) = \frac{1}{\Gamma(\alpha)} \begin{cases} t^{\alpha-1}(1-s)^{\alpha-\mu-1} - (t-s)^{\alpha-1}, & s \leq t, \\ t^{\alpha-1}(1-s)^{\alpha-\mu-1}, & t \leq s, \end{cases}$$

$$G_2(\eta, s) = \frac{1}{\Gamma(\alpha)} \begin{cases} (1-s)^{\alpha-\mu-1}\eta^{\alpha-\mu-1} - (\eta-s)^{\alpha-\mu-1}, & s \leq \eta, \\ (1-s)^{\alpha-\mu-1}\eta^{\alpha-\mu-1}, & s \leq \eta. \end{cases} \quad (8)$$

*Proof.* The general solution to the problem (7) is

$$\begin{cases} D_{0^+}^\beta(\phi(D_{0^+}^\alpha u(t))) = \lambda f(u(t)), & 0 < t < 1, \\ u(0) = u'(0) = u'(1) = 0, \quad \phi(D_{0^+}^\alpha u(0)) = (\phi(D_{0^+}^\alpha u(1)))' = 0, \end{cases} \quad (4)$$

where  $\phi(s) = |s|^{p-2}s$ ,  $p > 1$ ,  $\alpha \in (2, 3]$ ,  $\beta \in (1, 2]$ ,  $f: (0, +\infty) \rightarrow (0, +\infty)$  is continuous, and  $D_{0^+}^\alpha, D_{0^+}^\beta$  are the Riemann–Liouville fractional derivatives.

Based on the above research, this paper analyzed the following fractional three-point boundary value problem with the  $p$ -Laplacian operator:

$$u(t) = - \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} y(s)ds + C_1 t^{\alpha-1} + C_2 t^{\alpha-2} + C_3 t^{\alpha-3}. \quad (9)$$

From the boundary value condition of (7),  $C_2 = C_3 = 0$ ,

$$C_1 = \frac{1}{\Delta} \left[ \int_0^1 \frac{(1-s)^{\alpha-\mu-1}}{\Gamma(\alpha)} y(s)ds - \delta \int_0^\eta \frac{(\eta-s)^{\alpha-\mu-1}}{\Gamma(\alpha)} y(s)ds \right]. \quad (10)$$

Therefore,

$$\begin{aligned} u(t) &= - \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} y(s)ds + \int_0^1 \frac{t^{\alpha-1}(1-s)^{\alpha-\mu-1}}{\Gamma(\alpha)} y(s)ds \\ &\quad + \frac{\delta t^{\alpha-1}}{\Delta} \int_0^1 \frac{\eta^{\alpha-\mu-1}(1-s)^{\alpha-\mu-1}}{\Gamma(\alpha)} y(s)ds \\ &\quad - \frac{\delta t^{\alpha-1}}{\Delta} \int_0^\eta \frac{(\eta-s)^{\alpha-\mu-1}}{\Gamma(\alpha)} y(s)ds \\ &= \int_0^1 G_1(t, s)y(s)ds + \frac{\delta}{\Delta} t^{\alpha-1} \int_0^1 G_2(\eta, s)y(s)ds = \int_0^1 G(t, s)y(s)ds. \end{aligned} \quad (11)$$

□

**Lemma 3.** *If  $w \in C[0, 1]$ , then the fractional boundary value problem is*

$$\begin{cases} D_{0^+}^\beta(\phi(D_{0^+}^\alpha u(t))) = w(t), & 0 < t < 1, \\ u(0) = u'(0) = 0, \quad \phi(D_{0^+}^\alpha u(0)) = (\phi(D_{0^+}^\alpha u(1)))' = 0. \end{cases} \quad (12)$$

The unique solution is  $u(t) = \int_0^1 G(t, s)\phi^{-1}(\int_0^1 H(s, \tau)w(\tau)d\tau)ds$ , where

$$H(t, s) = \frac{1}{\Gamma(\beta)} \begin{cases} t^{\beta-1}(1-s)^{\beta-2} - (t-s)^{\beta-1}, & s \leq t, \\ t^{\beta-1}(1-s)^{\beta-2}, & t \leq s, \end{cases} \quad (13)$$

*Proof.* Problem (12) is equivalent to

$$\phi(D_{0^+}^\alpha u(t)) = \int_0^t \frac{(t-\tau)^{\beta-1}}{\Gamma(\beta)} w(\tau) d\tau + C_1 t^{\beta-1} + C_2 t^{\beta-2}. \tag{14}$$

From the boundary value condition of (12),  $C_2 = 0$  and  $C_1 = -\int_0^1 ((1-\tau)^{\beta-2})/\Gamma(\beta) w(\tau) d\tau$ ; then,

$$\begin{aligned} \phi(D_{0^+}^\alpha u(t)) &= \int_0^t \frac{(t-\tau)^{\beta-1}}{\Gamma(\beta)} w(\tau) d\tau - \int_0^1 \frac{t^{\beta-1}(1-\tau)^{\beta-2}}{\Gamma(\beta)} w(\tau) d\tau \\ &= -\int_0^1 H(t, \tau) w(\tau) d\tau. \end{aligned} \tag{15}$$

Therefore,

$$D_{0^+}^\alpha u(t) + \phi^{-1}\left(\int_0^1 H(t, \tau) w(\tau) d\tau\right) = 0. \tag{16}$$

Based on Lemma 2,

$$u(t) = \int_0^1 G(t, s) \phi^{-1}\left(\int_0^1 H(s, \tau) w(\tau) d\tau\right) ds. \tag{17}$$

**Lemma 4.** The properties of  $G(t, s)$  and  $H(t, s)$  are

- (i)  $0 \leq G(t, s) \leq G(1, s)$ , for  $(t, s) \in [0, 1] \times [0, 1]$
- (ii)  $G(t, s) \geq (1/4)^{\alpha-1} G(1, s)$ , for  $(t, s) \in I \times (0, 1) = (1/4, 3/4) \times (0, 1)$
- (iii) (see [13]).  $0 \leq H(t, s) \leq H(s, s)$ , for  $(t, s) \in [0, 1] \times [0, 1]$
- (iv) (see [13]).  $H(t, s) \geq (1/4)^{\beta-1} H(1, s)$ , for  $(t, s) \in I \times (0, 1) = (1/4, 3/4) \times (0, 1)$

*Proof* (i) For  $t \leq s$ , it is easy to show that  $(\partial/\partial t)G_1(t, s) \geq 0$ ; for  $t \geq s$ ,

$$\begin{aligned} \frac{\partial}{\partial t} G_1(t, s) &= \frac{(\alpha-1)[t^{\alpha-2}(1-s)^{\alpha-\mu-1} - (t-s)^{\alpha-2}]}{\Gamma(\alpha)} \\ &\geq \frac{(\alpha-1)(t-s)^{\mu-1}[t^{\alpha-\mu-1}(1-s)^{\alpha-\mu-1} - (t-s)^{\alpha-\mu-1}]}{\Gamma(\alpha)} \geq 0. \end{aligned} \tag{18}$$

So, for  $(t, s) \in [0, 1] \times [0, 1]$ ,

$$\frac{\partial}{\partial t} G_1(t, s) = \frac{\partial}{\partial t} G_1(t, s) + (\alpha-1) \frac{\delta}{\Delta} t^{\alpha-2} G_2(\eta, s) \geq 0. \tag{19}$$

Then, for  $(t, s) \in [0, 1] \times [0, 1]$ ,  $0 \leq G(t, s) \leq G(1, s)$ .

- (ii) For  $t \leq s$ , it is easy to show that  $G_1(t, s)/G_1(1, s) = t^{\alpha-1}$ ; for  $t \geq s$ ,

$$\begin{aligned} \frac{G_1(t, s)}{G_1(1, s)} &= \frac{t^{\alpha-1}(1-s)^{\alpha-\mu-1} - (t-s)^{\alpha-1}}{(1-s)^{\alpha-\mu-1} - (1-s)^{\alpha-1}} \\ &\geq \frac{t^{\alpha-1}(1-s)^{\alpha-\mu-1} - t^{\alpha-1}(1-s)^{\alpha-1}}{(1-s)^{\alpha-\mu-1} - (1-s)^{\alpha-1}} = t^{\alpha-1}. \end{aligned} \tag{20}$$

and for  $(t, s) \in I \times (0, 1) = (1/4, 3/4) \times (0, 1)$ ,

$$\begin{aligned} G(t, s) &= G_1(t, s) + \frac{\delta}{\Delta} t^{\alpha-1} G_2(\eta, s) \\ &\geq t^{\alpha-1} G_1(t, s) + \frac{\delta}{\Delta} t^{\alpha-1} G_2(\eta, s) \\ &\geq \left(\frac{1}{4}\right)^{\alpha-1} G(1, s). \end{aligned} \tag{21}$$

□

**Lemma 5** (see [20]). Let  $E = (E, \|\cdot\|)$  be a Banach space and let  $K \in E$  be a cone in  $E$ . Assume  $\Omega_1$  and  $\Omega_2$  are open subsets of  $E$  with  $0 \in \Omega_1$  and  $\overline{\Omega_1} \subset \Omega_2$  and let  $T: K \cap (\overline{\Omega_2} \setminus \Omega_1) \rightarrow K$  be a continuous and completely continuous. In addition, suppose either

- (1)  $\|Tu\| \leq \|u\|$ ,  $u \in K \cap \partial\Omega_1$ , and  $\|Tu\| \geq \|u\|$ ,  $u \in K \cap \partial\Omega_2$ , or
- (2)  $\|Tu\| \geq \|u\|$ ,  $u \in K \cap \partial\Omega_1$ , and  $\|Tu\| \leq \|u\|$ ,  $u \in K \cap \partial\Omega_2$

**Lemma 6** (see [16]). Let  $K$  be a cone in a real Banach space  $E$ ,  $K_r = \{x \in K: \|x\| < r\}$ ,  $\psi$  be nonnegative continuous concave functional on  $K$  such that  $\psi(x) \leq \|x\|, \forall x \in \overline{K}_r$ , and

$$K(\psi, d, e) = \{x \in K: d \leq \psi(x), \|x\| \leq e\}. \tag{22}$$

Suppose  $T: \overline{K}_r \rightarrow \overline{K}_r$  is completely continuous and there exist constants  $0 < c < d < e \leq r$  such that

- (i)  $\{x \in K(\psi, d, e) \mid \psi(x) > d\} \neq \emptyset$  and  $\psi(Tx) > d$  for  $x \in K(\psi, d, e)$
- (ii)  $\|Tx\| < c$  for  $x \leq c$
- (iii)  $\psi(Tx) > d$  for  $x \in K(\psi, d, r)$  with  $\|Tx\| > e$

Then,  $T$  has at least three fixed points  $x_1, x_2$ , and  $x_3$  with  $\|x_1\| < c, d < \psi(x_2)$ , and  $c < \|x_3\|$  with  $\psi(x_3) < d$ .

### 3. Main Results

When  $E = [0, 1]$ , any  $u \in E, \|u\| = \max_{0 \leq t \leq 1} |u(t)|$ , then  $E$  is a real Banach space.  $K \in E$  is a cone, which can be defined as  $K = \{u \in E: \min_{t \in [0, 1]} u(t) \geq 0, \min_{t \in I} u(t) \geq (1/4)^{\alpha-1} \|u\|\}$ . Defining the operator  $T: E \rightarrow E$ , for any  $u \in E$ ,

$$Tu(t) = \int_0^1 G(t, s) \phi^{-1}\left(\int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau\right) ds, \tag{23}$$

and for convenience, the following notation is introduced:



$$M = \left( \int_0^1 G(1, s) \phi^{-1} \left( \int_0^1 H(\tau, \tau) d\tau \right) ds \right)^{-1}, \tag{24}$$

$$N = \left( \int_I G(1, s) \phi^{-1} \left( \int_I H(1, \tau) d\tau \right) ds \right)^{-1}.$$

**Theorem 1.** *If there are two positive numbers  $0 < r_1 < r_2$  such that the following conditions hold:*

- (B<sub>1</sub>)  $f(t, u) \geq \phi(r_1 N 4^{\alpha+\beta-2})$ , for  $(t, u) \in [0, 1] \times [0, r_1]$
- (B<sub>2</sub>)  $f(t, u) \leq \phi(r_2 M)$ , for  $(t, u) \in [0, 1] \times [0, r_2]$

*then the fractional three-point boundary value problem (5) has at least one positive solution  $u$  and  $r_1 \leq \|u\| \leq r_2$ .*

*Proof.* From the continuity of  $G, H, f$ , it can be concluded that  $T: K \rightarrow K$  is continuous. For  $(t, s) \in I \times (0, 1), u \in K$ , by Lemma 4, we have

$$\begin{aligned} \min_{t \in I} Tu(t) &= \min_{t \in I} \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \\ &\geq \left(\frac{1}{4}\right)^{\alpha-1} \int_0^1 G(1, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \\ &\geq \left(\frac{1}{4}\right)^{\alpha-1} \|Tu\|. \end{aligned} \tag{25}$$

It means that  $T(K) \subset K$ . Therefore, the Arzela–Ascoli theorem can prove that the operator  $T: K \rightarrow K$  is completely continuous.

Let  $\Omega_1 = \{u \in K: \|u\| \leq \lambda_1\}$ , for  $u \in \partial\Omega_1$ . From Lemma 4 and (B<sub>1</sub>), we can conclude that

$$\begin{aligned} Tu(t) &= \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \\ &\geq r_1 N 4^{\alpha+\beta-2} \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) d\tau \right) ds \\ &\geq r_1 N 4^{\alpha+\beta-2} \int_I \left(\frac{1}{4}\right)^{\alpha-1} G(1, s) \phi^{-1} \left( \int_I \left(\frac{1}{4}\right)^{\beta-1} H(1, \tau) d\tau \right) ds \\ &= r_1 N \int_I G(1, s) \phi^{-1} \left( \int_I H(1, \tau) d\tau \right) ds \\ &= r_1. \end{aligned} \tag{26}$$

Then, when  $u \in \partial\Omega_1, \|Tu\| \geq \|u\|$ .

Let  $\Omega_2 = \{u \in K: \|u\| \leq \lambda_2\}$ , for  $u \in \partial\Omega_2$ . Then, we also can conclude from Lemma 4 and (B<sub>2</sub>) that

$$\begin{aligned} Tu(t) &= \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \\ &\leq r_2 M \int_0^1 G(1, s) \phi^{-1} \left( \int_a^b H(\tau, \tau) d\tau \right) ds \\ &= r_2. \end{aligned} \tag{27}$$

Therefore, for  $u \in \partial\Omega_2, \|Tu\| \leq \|u\|$ . In summary, by Lemma 5, the fractional three-point boundary value problem (5) has at least one positive solution  $u$  and  $r_1 \leq \|u\| \leq r_2$ .  $\square$

**Theorem 2.** *If there exist positive real numbers  $0 < c < d < (1/4)^{\alpha-1} r$  such that the following conditions hold:*

- (B<sub>3</sub>)  $f(t, u) < \phi(rM)$ , for  $(t, u) \in [0, 1] \times [0, r]$
- (B<sub>4</sub>)  $f(t, u) > \phi(dN/(1/4)^{\alpha+\beta-2})$ , for  $(t, u) \in I \times [d, d/(1/4)^{\alpha-1}]$
- (B<sub>5</sub>)  $f(t, u) < \phi(cM)$ , for  $(t, u) \in [0, 1] \times [0, c]$

*then the fractional three-point boundary value problem (5) has at least three positive solutions  $u_1, u_2$ , and  $u_3$  with*

$$\begin{aligned} \max_{0 \leq t \leq 1} |u_1(t)| &< c, \\ d &< \min_{t \in I} |u_2(t)| < r, \\ c &< \max_{0 \leq t \leq 1} |u_1(t)| < r, \\ \min_{t \in I} |u_3(t)| &< d. \end{aligned} \tag{28}$$

*Proof.* Firstly, if  $u \in \overline{K_r}$ , then we may assert that  $T: \overline{K_r} \rightarrow \overline{K_r}$  is a completely continuous operator. To see this, suppose  $u \in \overline{K_r}$ ; then,  $\|u\| \leq r$ . It follows from Lemma 4 and (B<sub>3</sub>) that

$$\begin{aligned} \|Tu(t)\| &= \max_{0 \leq t \leq 1} |Tu(t)| \\ &= \max_{0 \leq t \leq 1} \left( \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \right) \\ &\leq r_2 M \int_0^1 G(1, s) \phi^{-1} \left( \int_a^b H(\tau, \tau) d\tau \right) ds \\ &= r_2. \end{aligned} \tag{29}$$

Therefore,  $T: \overline{K_r} \rightarrow \overline{K_r}$ . This together with Lemma 5 implies that  $T: \overline{K_r} \rightarrow \overline{K_r}$  is a completely continuous operator. In the same way, if  $u \in \overline{K_c}$ , then assumption (B<sub>5</sub>) yields  $\|Tu\| < c$ . Hence, condition (ii) of Lemma 6 is satisfied.

To check condition (i) of Lemma 6, we let  $u(t) = d/(1/4)^{\alpha-1}$  for  $t \in [0, 1]$ . It is easy to verify that  $u(t) = d/(1/4)^{\alpha-1} \in K(\psi, d, d/(1/4)^{\alpha-1})$  and  $\psi(u) = (d/(1/4)^{\alpha-1}) > d$ , and so

$$\left\{ u \in K \left( \psi, d, \frac{d}{(1/4)^{\alpha-1}} \right) \mid \psi(u) > d \right\} \neq \emptyset. \tag{30}$$

Thus, for all  $u \in K(\psi, d, d/(1/4)^{\alpha-1})$ , we have that  $d \leq u(t) \leq (d/(1/4)^{\alpha-1})$  for  $t \in I$  and  $Tu \in K$ .

From Lemma 4 and (B<sub>4</sub>), one has



$$\begin{aligned}
 \psi(Tu(t)) &= \min_{t \in I} |Tu(t)| \\
 &= \min_{t \in I} \left( \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \right) \\
 &\geq \int_I \left( \frac{1}{4} \right)^{\alpha-1} G(1, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \\
 &\geq \int_I \left( \frac{1}{4} \right)^{\alpha-1} G(t, s) \phi^{-1} \left( \int_I H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \\
 &\geq \frac{dN}{(1/4)^{\alpha+\beta-2}} \int_I \left( \frac{1}{4} \right)^{\alpha-1} G(1, s) \phi^{-1} \left( \int_I \left( \frac{1}{4} \right)^{\beta-1} H(1, \tau) d\tau \right) ds \\
 &= dN \int_I G(1, s) \phi^{-1} \left( \int_I H(1, \tau) d\tau \right) ds \\
 &= d.
 \end{aligned} \tag{31}$$

This shows that condition (i) of Lemma 6 holds.

Secondly, we verify that (iii) of Lemma 6 is satisfied. By Lemma 6, we have

$$\begin{aligned}
 \min_{t \in I} |Tu(t)| &= \min_{t \in I} \left( \int_0^1 G(t, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \right) \\
 &\geq \left( \frac{1}{4} \right)^{\alpha-1} \left( \int_0^1 G(1, s) \phi^{-1} \left( \int_0^1 H(s, \tau) f(\tau, u(\tau)) d\tau \right) ds \right) \\
 &\geq \left( \frac{1}{4} \right)^{\alpha-1} \|Tu\| > d,
 \end{aligned} \tag{32}$$

for  $u \in K(\psi, d, r)$  with  $\|Tu\| > (d/(1/4)^{\alpha-1})$ , which shows that condition (iii) of Lemma 6 holds.

To sum up, all the conditions of Lemma 6 are satisfied; from Lemma 6, it follows that there exist three positive solutions  $u_1, u_2$ , and  $u_3$  with

$$\begin{aligned}
 \max_{0 \leq t \leq 1} |u_1(t)| &< c, \\
 d < \min_{t \in I} |u_2(t)| &< r, \\
 c < \max_{0 \leq t \leq 1} |u_1(t)| &< r, \\
 \min_{t \in I} |u_3(t)| &< d.
 \end{aligned} \tag{33}$$

□

### 4. Conclusion

The existence of solutions to three-point boundary value problems of fractional differential equations with the  $p$ -Laplacian operators is discussed by using the fixed point exponential theorem and fixed point theorem of cone compression and cone tension. By extending the existence of solutions to boundary value problems, we have obtained the sufficient condition that the boundary value problem has multiple positive solutions or at least one positive solution.

### Data Availability

The data used to support the findings of the study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was supported by 2019 Project of Foundational Research Ability Enhancement for Young and Middle-Aged University Faculties of Guangxi (2019KY1046), Nature and Science Foundation of Anhui (2008085QA08), Scientific Research Projects of Institute of Information Technology of GUET (B201911), and Science and Technology Research Project of Heilongjiang Provincial Department of Education (12543079).

### References

- [1] K. B. Oldham and J. Spanier, *The Fractional Calculus (Theory and Applications of Differentiation and Integration to Arbitrary Order)*, Academic Press, San Diego, CA, USA, 1974.
- [2] S. G. Kilbas and A. A. Marichev, *Fractional Integrals and Derivatives (Theory and Applications)*, Gordon & Breach, New York, NY, USA, 1993.
- [3] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York, NY, USA, 1993.
- [4] I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, CA, USA, 1999.
- [5] R. Hilfer, *Applications of Fractional Calculus in Physics*, World Scientific, New Jersey, NJ, USA, 2000.
- [6] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam, Netherlands, 2006.
- [7] V. Lakshmikantham, S. Leela, and J. Vasundhara, *Theory of Fractional Dynamic Systems*, Cambridge Academic Publishers, Cambridge, UK, 2009.
- [8] R. Herrmann, *Fractional Calculus (An Introduction for Physicists)*, World Scientific, Singapore, 2014.
- [9] T. M. Atanacković, S. Pilipović, B. Stanković, and D. Zorica, *Fractional Calculus with Applications in Mechanics*, Wiley, New York, NY, USA, 2014.
- [10] Y. Liu, D. P. Xie, D. D. Yang, and C. Z. Bai, "Two generalized Lyapunov-type inequalities for a fractional  $p$ -Laplacian equation with fractional boundary conditions," *Journal of Inequalities and Applications*, vol. 98, no. 1, pp. 1–16, 2017.
- [11] L. S. Zhang, X. Wu, and Y. Caccetta, "Extremal solutions for  $p$ -Laplacian differential systems via iterative computation," *Applied Mathematics Letters*, vol. 26, no. 12, pp. 1151–1158, 2013.
- [12] S. N. Rao, M. Singh, and M. Z. Meetei, "Multiplicity of positive solutions for Hadamard fractional differential equations with  $p$ -Laplacian operator," *Boundary Value Problems*, vol. 43, no. 1, pp. 1–25, 2020.
- [13] Y. Wang, "Multiple positive solutions for mixed fractional differential system with  $p$ -Laplacian operators," *Boundary Value Problems*, vol. 144, no. 1, pp. 1–17, 2019.
- [14] Y. He and B. Bi, "Existence and iteration of positive solution for fractional integral boundary value problems with  $p$ -

- Laplacian operator,” *Advances in Difference Equations*, vol. 415, pp. 1–15, 2019.
- [15] G. Chai, “Positive solutions for boundary value problem of fractional differential equation with  $p$ -Laplacian operator,” *Boundary Value Problems*, vol. 18, no. 1, pp. 1–15, 2012.
- [16] Y. Tian, Y. Wei, and S. Sun, “Multiplicity for fractional differential equations with  $p$ -Laplacian,” *Boundary Value Problems*, vol. 127, no. 1, pp. 1–18, 2018.
- [17] Y. Tian, Z. Bai, and S. Sun, “Positive solutions for a boundary value problem of fractional differential equation with  $p$ -Laplacian operator,” *Advances in Difference Equations*, vol. 349, pp. 1–19, 2019.
- [18] Z. Han, H. Lu, and C. Zhang, “Positive solutions for eigenvalue problems of fractional differential equation with generalized  $p$ -Laplacian,” *Applied Mathematics and Computation*, vol. 257, no. 1, pp. 526–536, 2015.
- [19] C. Z. Bai, “Existence and uniqueness of solutions for fractional boundary value problems with  $p$ -Laplacian operator,” *Advances in Difference Equations*, vol. 4, pp. 1–12, 2018.
- [20] Y. H. Li, “Multiple positive solutions for nonlinear mixed fractional differential equation with  $p$ -Laplacian operator,” *Advances in Difference Equations*, vol. 112, no. 1, pp. 1–12, 2019.
- [21] J. B. Liu, J. Zhao, and Z. Cai, “On the generalized adjacency, Laplacian and signless Laplacian spectra of the weighted edge corona networks,” *Physica A*, vol. 540, Article ID 123073, 2020.
- [22] J. B. Liu, J. Zhao, and Z. X. Zhu, “On the number of spanning trees and normalized Laplacian of linear octagonal-quadrilateral networks,” *International Journal of Quantum Chemistry*, vol. 119, Article ID 25971, 2019.
- [23] B. B. Zhou and L. L. Zhang, “Existence of positive solutions of boundary value problems for high-order nonlinear conformable differential equations with  $p$ -Laplacian operator,” *Advances in Difference Equations*, vol. 351, no. 1, pp. 1–14, 2019.
- [24] K. S. Jong, “Existence and uniqueness of positive solutions of a kind of multi-point boundary value problems for nonlinear fractional differential equations with  $p$ -Laplacian operator,” *Mediterranean Journal of Mathematics*, vol. 15, no. 3, 2018.
- [25] H. X. Wang and W. M. Hu, “Existence of solutions for a class of fractional differential equations with  $p$ -Laplacian operators and integral boundary conditions,” *Mathematics in Practice and Theory*, vol. 46, no. 16, pp. 228–236, 2016.

## Research Article

# On Local Generalized Ulam–Hyers Stability for Nonlinear Fractional Functional Differential Equation

Dongming Nie,<sup>1</sup> Azmat Ullah Khan Niazi ,<sup>2,3</sup> and Bilal Ahmed<sup>3</sup>

<sup>1</sup>Department of Common Courses, Anhui Xinhua University, Hefei 230088, China

<sup>2</sup>Department of Mathematics, Xiangtan University, Xiangtan, Hunan 411105, China

<sup>3</sup>Department of Mathematics and Statistics, University of Lahore, Sargodha, Lahore, Pakistan

Correspondence should be addressed to Azmat Ullah Khan Niazi; azmatniazi35@gmail.com

Received 27 April 2020; Accepted 27 May 2020; Published 18 June 2020

Guest Editor: Shaohui Wang

Copyright © 2020 Dongming Nie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We discuss the existence of positive solution for a class of nonlinear fractional differential equations with delay involving Caputo derivative. Well-known Leray–Schauder theorem, Arzela–Ascoli theorem, and Banach contraction principle are used for the fixed point property and existence of a solution. We establish local generalized Ulam–Hyers stability and local generalized Ulam–Hyers–Rassias stability for the same class of nonlinear fractional neutral differential equations. The simulation of an example is also given to show the applicability of our results.

## 1. Introduction

Fractional differential equations (FDEs) have boosted considerably due to their application in various fields of sciences, such as engineering, chemistry, mechanics, and physics. Recently, fractional ordinary differential equations (ODEs) and partial differential equations (PDEs) have been developed significantly. Indeed, we can also find applications in control, electromagnetism, and electrochemistry (see [1, 2]). For more details in this area, one can also see the monographs of Kilbas et al. [3], Li et al. [4, 5], Liu et al. [6–9], Miller and Ross [10], Podulbny [11], Rehman et al. [12], and the references therein.

Positive solution of fractional ODEs have already discussed in [13–15]. However, until now research on existence of positive solution for fractional functional differential equations (FFDEs) is rare. Research in different fields including engineering, physics, and biosciences have proved that numerous system structures explain more exactly with the help of FDEs [16–21] and similarly FDEs with delay [22–27] are more accurate to illustrate the real world problems compared to FDEs without delay. So, in [28], Miller and Ross mentioned that this field has been touched by almost all fields of engineering and science.

Idea of Ulam stability started in 1940, when during the talk at Wisconsin University Ulam posed a problem that “When can we assert that an approximate solution of a functional equation can be approximated by a solution of the corresponding equation?” (for more details, see [29]). After one year, Hyers first gave the answer to the Ulam’s question [30] in case of Banach spaces. After that, stability of this type is called the Ulam–Hyers stability. Rassias [31] in 1978 provided an outstanding generalization of the Ulam–Hyers stability of mappings by considering variables.

Usually, the discussion of the existence and uniqueness of a solution for nonlinear FDEs normally fixed point theory has been used [16, 32, 33]. Motivated by these and [34–36], in this work, we have discussed existence and uniqueness of solution also after applying some sufficient conditions, obtained positive solution, and at the end established local generalized Ulam–Hyers stability and local generalized Ulam–Hyers–Rassias stability and presented stability results graphically for the class of nonlinear FDEs with delay given by

$$\begin{cases} {}^c D_0^\gamma [z(t) - g(t, z_t)] = z(t)f(t, z_t), & t \in I, \\ z(t) = \psi(t) \geq 0, & t \in [-\tau, 0], \end{cases} \quad (1)$$

where  ${}^c D_0^\gamma$  is the Caputo derivative and  $0 < \gamma < 1$ , while  $\psi \in C$  and  $I = [0, T]$  where  $C = C([- \tau, 0]; \mathbb{R}^+)$  is space of continuous functions,  $\tau$  is a nonnegative real number also  $f, g \in (I \times C, \mathbb{R}^+)$  are continuous functions, and  $z_t \in C$  is defined as  $z_t(\theta) = z(t + \theta), -\tau \leq \theta \leq 0$ .

Remaining paper is arranged as follows. Section 2 includes some basic definitions and lemmas used throughout this paper. In Section 3, we have the main results. In Section 4, we establish local generalized Ulam–Hyers stability and local generalized Ulam–Hyers–Rassias stability for problem (1), and in Section 5, simulation of an example is given to show the applicability of our results.

### 2. Preliminaries

Let  $Q$  be a cone in a real Banach space  $X$  and partial order  $\leq$  introduced by  $Q$  in  $X$  is as

$$\text{if } v - u \in Q \text{ then } u \leq v. \tag{2}$$

*Definition 1.* Let  $u, v \in X$  be the order interval  $\langle u, v \rangle$  defined as

$$\langle u, v \rangle = \{w \in X: u \leq w \leq v\}. \tag{3}$$

*Definition 2.* The functional  $h(t, z_t)$  is said to be nondecreasing w.r.t.  $\phi$  on  $I \times C$  such that

$$x(\phi) \leq y(\phi), \quad \phi \in [-\tau, 0], \tag{4}$$

implies that

$$h(t, x) \leq h(t, y), \tag{5}$$

for any  $(t, x)$  and  $(t, y) \in I \times C$ .

*Definition 3* (see [11]). The fractional integral of order  $\gamma$  for a function  $g$  with lower limit 0 can be defined as

$$I_0^\gamma g(t) = \frac{1}{\Gamma(\gamma)} \int_0^t \frac{g(s)}{(t-s)^{1-\gamma}} ds, \quad \gamma > 0, t > 0, \tag{6}$$

where  $\Gamma$  is the gamma function and right-hand side of upper equality is defined pointwise on  $\mathbb{R}^+$ .

*Definition 4* (see [11]). The left Caputo fractional derivative of order  $\gamma$  is given by

$${}^c D_t^\gamma f(t) = \frac{1}{\Gamma(n-\gamma)} \int_a^t \frac{f^{(n)}(s)}{(t-s)^{\gamma+1-n}} ds, \tag{7}$$

where  $n = [\gamma] + 1$  ( $[\gamma]$  stands for the bracket function of  $\gamma$ ).

**Lemma 1** (Leray Schauder fixed point theorem, see [11]). *Let  $S_r$  be a nonempty, closed, bounded, and convex subset of Banach space  $X$  and  $Q: S_r \rightarrow S_r$  is a compact and continuous map; then,  $Q$  has a fixed point in  $S_r$ .*

**Theorem 1.** *If  $U \subset Q$ , where  $Q$  is a cone of space of partial order  $X$ ,  $F: U \rightarrow X$  be nondecreasing. If there exists  $u_0, v_0 \in U$  such that  $v_0 \geq u_0, \langle u_0, v_0 \rangle \subset U$ , and  $u_0$  is lower solution and  $v_0$  is an upper solution of  $y - F(y) = 0$ , then*

*$y - F(y) = 0$  has minimum solution  $u^*$  and maximum solution  $v^*$  in  $\langle u_0, v_0 \rangle$  such that  $v^* \geq u^*$ , and if anyone of these conditions holds*

- (i)  $X$  is reflexive,  $F$  is weak continuous or continuous, and  $Q$  is normal
- (ii)  $F$  is completely continuous and  $Q$  is normal
- (iii)  $F$  is continuous and  $Q$  is regular

### 3. Main Results

In this section, we have discussed existence and uniqueness of solution and some conditions for positive solution of equation (1). Consider  $X = C[-\tau, T]$  is a Banach space endowed with the supremum norm and cone  $Q$  is defined as  $Q = \{z \in Q; z(t) \geq 0, -\tau \leq t \leq T\}$ . From equation (1), we have

$$I_0^\gamma ({}^c D_0^\gamma [z(t) - g(t, z_t)]) = I_0^\gamma [z(t)f(t, z_t)], \tag{8}$$

and we can easily get that

$$z(t) = \psi(0) - g(0, \psi) + g(t, z_t) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-\tau)^{\gamma-1} z(\tau) f(\tau, z_\tau) d\tau, \quad t \in [0, T]. \tag{9}$$

Let  $x(\cdot): [-\tau, T] \rightarrow \mathbb{R}^+$  be the function defined by

$$x(t) = \begin{cases} \psi(t), & t \in [-\tau, 0], \\ 0, & t \in I. \end{cases} \tag{10}$$

So,  $x_0 = \psi$  and for every  $w \in C(I, \mathbb{R}^+)$  with  $w(0) = 0$ , and we defined  $\bar{w}$  as

$$\bar{w}(t) = \begin{cases} 0, & t \in [-\tau, 0], \\ w(t), & t \in I. \end{cases} \tag{11}$$

Let  $x(\cdot)$  satisfy equation (9), and we can decompose  $z(\cdot)$  as  $z(t) = \bar{w}(t) + x(t)$ , for  $0 \leq t \leq T$  so that  $z_t = \bar{w}_t + x_t$ , for every  $0 \leq t \leq T$ , where  $\bar{w}(t)$  is such that  $|\bar{w}(t)| < 1, \forall t \in I$  and the function  $w(\cdot)$  satisfies

$$\begin{aligned} \bar{w}(t) &= \psi(0) - g(0, \psi) + g(t, \bar{w}_t + x_t) \\ &+ \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} (\bar{w}(s) + x(s)) f(s, \bar{w}_s + x_s) ds, \end{aligned} \tag{12}$$

$t \in I.$

Set  $H = \{z \in Q; z = \psi(t), t \in [-\tau, 0]\} \subset Q$ , where  $Q$  is a cone and  $Q \subset X$  and  $X$  is a Banach space with Sup norm  $\|\bar{w}(t)\| \leq \sup\{|w(t)|; -\tau \leq t \leq T\}$ . Let the operator  $P: H \rightarrow H$  be defined by

$$\begin{aligned} (P\bar{w})(t) &= \psi(0) - g(0, \psi) + g(t, \bar{w}_t + x_t) \\ &+ \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}(s) f(s, \bar{w}_s + x_s) ds. \end{aligned} \tag{13}$$

Before proving main results, we introduce the following conditions:

(C1) Let us take  $f, g: I \times C \rightarrow R^+$  with  $g(0, \psi) = 0$ , where function  $g$  is nondecreasing and  $|g(t, z_t)| \leq L_1$  and  $f$  and  $g$  satisfy Lipchitz condition that is, and there exists nonnegative constants  $c_1$  and  $c_2$  such that

$$\begin{aligned} |f(t, u) - f(t, v)| &\leq c_1 \|u - v\|, \quad \forall t \in I, \text{ and } u, v \in C, \\ |g(t, u) - g(t, v)| &\leq c_2 \|u - v\|, \quad \forall t \in I, \text{ and } u, v \in C. \end{aligned} \tag{14}$$

(C2) If  $G \subset H$  be bounded,  $f$  is nondecreasing and

$$\max_{t \in I} \{f(t, z_t), z \in G\} = L - 1, \tag{15}$$

where  $L > 0$ . Let  $\{H = \{z \in Q; z = \psi(t), t \in [-\tau, 0]\} \subset Q\}$ . Define the operator  $P$  by

$$P\bar{w}(t) = \begin{cases} \psi(0) + g(t, \bar{w}_t + x_t) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}(s) f(s, \bar{w}_s + x_s) ds, & t \in I, \\ 0, & t \in [-\tau, 0]. \end{cases} \tag{16}$$

**Theorem 2.** Assume that  $f$  and  $g$  are nonnegative and continuous; then, operator  $P: H \rightarrow H$  is completely continuous.

*Proof.* As we know that  $f$  and  $g$  are nonnegative and continuous, so operator  $P: H \rightarrow H$  is continuous. Let

$G \subset H$  is bounded, i.e., there exists a constant  $l > 0$  such that  $\|z\| \leq l, \forall z \in G$ . Here, we have to keep in mind the condition that  $x(t) = 0$  when  $t \in I$ . If  $\max\{f(t, z_t) | 0 \leq t \leq T, z \in G\} = L - 1$ , and for all  $z \in G$ , we get that

$$\begin{aligned} |(P\bar{w})(t)| &\leq \left| \psi + g(t, \bar{w}_t + x_t) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}(s) f(s, \bar{w}_s + x_s) ds \right| \\ &\leq l + L_1 + \frac{L}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} ds \\ &\leq l + L_1 + \frac{LT^\gamma}{\Gamma(\gamma+1)}. \end{aligned} \tag{17}$$

Hence,  $PG$  is bounded. Now, we show that  $PG$  is equicontinuous, and the proof is divided into three possible steps.

Step 1: for every  $\bar{w} \in G$  and  $t_1, t_2 \in I$  such that  $t_2 > t_1$ , then

$$\begin{aligned} |P\bar{w}(t_1) - P\bar{w}(t_2)| &\leq \left| g(t_1, \bar{w}_{t_1} + x_{t_1}) - g(t_2, \bar{w}_{t_2} + x_{t_2}) \right| + \frac{1}{\Gamma(\gamma)} \int_0^{t_1} (t_1-s)^{\gamma-1} |\bar{w}(s) f(s, \bar{w}_s + x_s)| ds \\ &\quad - \frac{1}{\Gamma(\gamma)} \int_0^{t_2} (t_2-s)^{\gamma-1} |\bar{w}(s) f(s, \bar{w}_s + x_s)| ds \\ &\leq c_2 \|\bar{w}_{t_1} - \bar{w}_{t_2}\| + \frac{1}{\Gamma(\gamma)} \left[ \int_0^{t_1} (t_1-s)^{\gamma-1} - \int_0^{t_2} (t_2-s)^{\gamma-1} \right] |\bar{w}(s) f(s, \bar{w}_s + x_s)| ds \\ &\leq c_2 \|\bar{w}_{t_1} - \bar{w}_{t_2}\| + \frac{1}{\Gamma(\gamma)} |\bar{w}(t)| \|f(t, \bar{w}_t + x_t)\| \left[ \int_0^{t_1} |(t_1-s)^{\gamma-1} - (t_2-s)^{\gamma-1}| ds - \int_{t_1}^{t_2} (t_2-s)^{\gamma-1} ds \right] \\ &\leq c_2 \|\bar{w}_{t_1} - \bar{w}_{t_2}\| + \frac{L}{\Gamma(\gamma+1)} (t_1^\gamma + (t_2-t_1)^\gamma - t_2^\gamma + (t_2-t_1)^\gamma) \\ &< c_2 \sup\{|w(t_1) - w(t_2)|\} + \frac{L}{\Gamma(\gamma+1)} (t_2 - t_1)^\gamma. \end{aligned} \tag{18}$$

As  $t_1 \rightarrow t_2$ , the right-hand side of above inequality tends to zero.

Step 2: for every  $\bar{w} \in G$ , as  $\psi \in C[-\tau, 0]$  and  $t_1 \in [-\tau, 0]$  while  $0 \leq t_2 \leq T$ , there exists  $\delta > 0$  such that when  $\delta_1 > t_1$ ,  $|\psi(t_1) - \psi(0)| < \epsilon/2$ .

If  $\delta = \min\{\delta_1, (\Gamma(\gamma + 1)(\epsilon - 2L_1/2L))^{1/\gamma}\}$ , then for  $t_2 - t_1 < \delta$ , we get that

$$\begin{aligned}
 |P\bar{w}(t_1) - P\bar{w}(t_2)| &\leq |\psi(t) - \psi(0)| + |g(t_2, \bar{w}_{t_2} + x_{t_2})| + \frac{1}{\Gamma(\gamma)} \int_0^{t_2} (t_2 - s)^{\gamma-1} |\bar{w}(s)f(s, \bar{w}_s + x_s)| ds \\
 &\leq \frac{\epsilon}{2} + L_1 + \frac{L}{\Gamma(\gamma + 1)} t_2^\gamma = \frac{\epsilon}{2} + L_1 + \frac{L}{\Gamma(\gamma + 1)} \delta^\gamma \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
 \end{aligned}
 \tag{19}$$

Step 3: for every  $\bar{w} \in G$ ,  $\forall \epsilon > 0$ , and  $t_1, t_2 \in [-\tau, 0]$  where  $t_2 > t_1$ . As  $\psi$  is continuous, we have

$$|P\bar{w}(t_1) - P\bar{w}(t_2)| = |\psi(t_1) - \psi(t_2)| < \epsilon, \text{ when } \delta_1 > t_2 - t_1.
 \tag{20}$$

Hence,  $PG$  is equicontinuous. So,  $\overline{PG}$  is compact by Arzela-Ascoli Theorem.  $\square$

**Theorem 3.** Suppose that C1 and C2 hold. Let  $f, g: I \times C \rightarrow \mathbb{R}^+$  and  $[c_2 + (T^\gamma/\Gamma(\gamma + 1))(c_1 + L)] < 1$ . Then, there exists a unique solution for IVP (1) on the interval  $[-\tau, T]$ .

*Proof.* Consider the operator  $P: H \rightarrow H$  defined as

$$(P\bar{w})(t) = \psi(0) + g(t, \bar{w}_t + x_t)$$

$$+ \frac{1}{\Gamma(\gamma)} \int_0^t (t - s)^{\gamma-1} \bar{w}(s)f(s, \bar{w}_s + x_s) ds.
 \tag{21}$$

We will show that  $P$  is a contraction. Let  $\bar{w}_1, \bar{w}_2 \in C(I, \mathbb{R}^+)$ , then we obtain the following sequence of inequalities:

$$\begin{aligned}
 |P\bar{w}_1(t) - P\bar{w}_2(t)| &\leq \left| g\left( \begin{matrix} t, \bar{w}_{1t} \\ +x_t \end{matrix} \right) - g(t, \bar{w}_{2t} + x_t) \right| \\
 &\quad + \left| \frac{1}{\Gamma(\gamma)} \int_0^t (t - s)^{\gamma-1} \bar{w}_1(s)f(s, \bar{w}_{1s} + x_s) ds - \frac{1}{\Gamma(\gamma)} \int_0^t (t - s)^{\gamma-1} \bar{w}_2(s)f(s, \bar{w}_{2s} + x_s) ds \right| \\
 &\leq c_2 \|\bar{w}_{1t} - \bar{w}_{2t}\| + \frac{1}{\Gamma(\gamma)} \int_0^t (t - s)^{\gamma-1} |\bar{w}_1(s)f(s, \bar{w}_{1s} + x_s) - \bar{w}_2(s)f(s, \bar{w}_{2s} + x_s)| ds \\
 &\leq c_2 \|\bar{w}_{1t} - \bar{w}_{2t}\| + \frac{1}{\Gamma(\gamma)} \left[ \int_0^t (t - s)^{\gamma-1} |\bar{w}_1(s)f(s, \bar{w}_{1s} + x_s) - \bar{w}_2(s)f(s, \bar{w}_{2s} + x_s)| \right. \\
 &\quad \left. + \bar{w}_1(s)f(s, \bar{w}_{2s} + x_s) - \bar{w}_2(s)f(s, \bar{w}_{2s} + x_s) \right] ds \\
 &\leq c_2 \|\bar{w}_{1t} - \bar{w}_{2t}\| + \frac{1}{\Gamma(\gamma)} \left[ \int_0^t (t - s)^{\gamma-1} |\bar{w}_{1s}| c_1 \|\bar{w}_{1s} - \bar{w}_{2s}\| ds + \int_0^t (t - s)^{\gamma-1} L \sup\{|w_1(t) - w_2(t)|\} ds \right] \\
 &\leq c_2 \sup\{|w_1(t) - w_2(t)|\} + \frac{c_1 T^\gamma}{\Gamma(\gamma + 1)} \sup\{|w_1(t) - w_2(t)|\} + \frac{LT^\gamma}{\Gamma(\gamma + 1)} \sup\{|w_1(t) - w_2(t)|\} \\
 &\leq \left[ c_2 + \frac{T^\gamma}{\Gamma(\gamma + 1)} (c_1 + L) \right] \|\bar{w}_1 - \bar{w}_2\|.
 \end{aligned}
 \tag{22}$$



Therefore,

$$\|P\bar{w}_1(t) - P\bar{w}_2(t)\| \leq \left[ c_2 + \frac{T^\gamma}{\Gamma(\gamma + 1)} (c_1 + L) \right] \|\bar{w}_1(t) - \bar{w}_2(t)\|. \tag{23}$$

So,  $P$  is a contraction, and hence by Banach's contraction principle  $P$  has a unique fixed point.  $\square$

*Definition 5.* The function  $m \in X$  is said to be a lower solution of equation (1) (operator  $P$ ) if

$$\begin{cases} {}^c D_0^\gamma [m(t) - g(t, m_t)] \leq m(t) f(t, m_t), & t \in I, \\ m(t) \leq \psi(t), & t \in [-\tau, 0], \end{cases} \tag{24}$$

and the function  $n \in X$  is said to be an upper solution of equation (1) if

$$\begin{cases} {}^c D_0^\gamma [n(t) - g(t, n_t)] \geq n(t) f(t, n_t), & t \in I, \\ n(t) \geq \psi(t), & t \in [-\tau, 0]. \end{cases} \tag{25}$$

If inequalities are strict, then  $m(t)$  and  $n(t)$  are strict lower and upper solutions.

**Theorem 4.** *Let*

- (H1)  $f$  and  $g$  from  $I \times C$  to  $\mathbb{R}^+$  are continuous, and  $f(t, \psi)$  and  $g(t, \psi)$  are nondecreasing in  $\psi$  for each  $t \in I$
- (H2)  $m_0$  is lower and  $n_0$  is an upper solution of equation (1), satisfying condition  $m_0(t) \leq n_0(t)$ ,  $t \in I$  and  $m_0, n_0 \in H$

*Then, equation (1) has at least a positive solution.*

*Proof.* We only take fixed point operator  $P$ . By Theorem 2,  $P$  is completely continuous and obviously by equation (16),  $u_0$  and  $v_0$  are lower and upper solutions of  $P$ , respectively. By (H1),  $\bar{w}_1, \bar{w}_2 \in H, \bar{w}_1 \leq \bar{w}_2$ , we have

$$\begin{aligned} P\bar{w}_1(t) &= \psi(0) + g(t, \bar{w}_{1t} + x_t) \\ &+ \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}_1(s) f(s, \bar{w}_{1s} + x_s) ds \\ &\leq \psi(0) + g(t, \bar{w}_{2t} + x_t) \\ &+ \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}_2(s) f(s, \bar{w}_{2s} + x_s) ds \\ &= P\bar{w}_2(t), \quad t \in I. \end{aligned} \tag{26}$$

So, the operator  $P$  is nondecreasing. Also, we have

$$\begin{aligned} Pm(t) &= \psi(0) + g(t, m_t + x_t) \\ &+ \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} m(s) f(s, m_s) ds, \end{aligned} \tag{27}$$

$$Pm(0) = \psi(0) + g(0, m_0 + x_0) \geq m(0).$$

As  $m \in X$  is a lower solution, therefore we can say that  $Pm(0) \geq m(0)$ , similarly  $Pn(0) \leq n(0)$  from the definition of upper solution of  $P$ .

Hence,  $P: \langle m_0, n_0 \rangle \rightarrow \langle m_0, n_0 \rangle$  is a completely continuous operator. Since  $Q$  is a normal cone, Theorem 1 implies that  $P$  has a fixed point  $\bar{w} \in \langle m_0, n_0 \rangle$ .

(H3) Here, we suppose that there exists  $M > 0$  such that  $0 < z f(t, z_t) \leq M, t \in I$ .  $\square$

**Theorem 5.** *Assume conditions (H1) and (H3) holds; then, equation (1) has a positive solution.*

*Proof.* Consider the equation

$$\begin{cases} {}^c D_0^\gamma [m(t) - g(t, m_t)] = 0, & t \in I, \\ m(t) = \psi(t), & t \in [-\tau, 0]. \end{cases} \tag{28}$$

Obviously, equation (28) has a solution  $m(t) = m_0 - g(t_0, \psi) + g(t, m_t)$ , which is a lower solution of equation (1). Similarly, consider the equation

$$\begin{cases} {}^c D_0^\gamma [n(t) - g(t, n_t)] = M, & t \in I, \\ n(t) = \psi(t), & t \in [-\tau, 0]. \end{cases} \tag{29}$$

We know that  $n(t) = n_0 - g(t_0, \psi) + g(t, n_t) + (Mt^\gamma/\Gamma(\gamma + 1))$ ,  $t \in I$  is an upper solution of equation (1) and  $n(t) \geq m(t)$ , so the above definition proves that equation (1) has minimum one positive solution.

For getting another result for a positive solution of equation (1), now we take the more general case of (1) and find the existence of a positive solution for the equations:

$$\begin{cases} {}^c D_0^\gamma [\nu(t) - g(t, \nu_t)] = N\nu(t), & t \in I, \\ \nu(t) = \psi(t), & t \in [-\tau, 0]. \end{cases} \tag{30}$$

**Theorem 6.** *Equation (30) has a least positive solution.*

*Proof.* Equation (30) is equivalent to

$$\nu(t) = \begin{cases} \nu(0) - g(0, \psi) + g(t, \nu(t)) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} N\nu(s) ds, & t \in I, \\ \psi(t), & t \in [-\tau, 0]. \end{cases} \tag{31}$$

Let  $P: H \rightarrow H$  be an operator defined as

$$Pv(t) = \begin{cases} v(0) + g(t, v(t)) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} Nv(s) ds, & t \in I, \\ \psi(t), & t \in [-\tau, 0]. \end{cases} \tag{32}$$

$P$  is completely continuous by Theorem 2.

Let us take  $B_R = \{v(t) \in C([-\tau, T], \mathbb{R}^+) \mid \|v(t) - v(0)\| \leq R\}$  a bounded, closed, and convex subset of the Banach space  $C[-\tau, \delta]$ , where  $\delta \in I$  and  $R > 0$ . The possible two cases are

Case 1: for all  $v \in B_R, t \in [I]$ , we obtain

$$\begin{aligned} |Pv(t) - v(0)| &\leq |g(t, v_t)| + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} |Nv(s)| ds \\ &\leq L_1 + \frac{N}{\Gamma(\gamma+1)} \|v\| t^\gamma, \quad t \in I. \end{aligned} \tag{33}$$

Since  $\|v\| \leq R + v(0)$ , hence

$$\|Pv(t) - v(0)\| \leq L_1 + \frac{N}{\Gamma(\gamma+1)} (v(0) + R) \delta^\gamma \leq R. \tag{34}$$

Here,  $\delta < ((R - L_1)\Gamma(\gamma + 1)/N(v(0) + R))^{1/\gamma}$ .

Case 2: if  $Pv(t) = \psi(t)$  and  $-\tau \leq t < 0$ , then

$$|Pv(t) - v(0)| = |\psi(t) - \psi(0)| = |v(t) - v(0)| \leq \|v(t) - v(0)\| \leq R. \tag{35}$$

Hence,  $P(B_R) \subseteq B_R$ . The Leray–Schauder fixed point theorem guarantees that operator  $P$  has at least one fixed point and so equation (30) has at least one positive solution.  $\square$

### 4. Stability

In this section, we will discuss local generalized Ulam–Hyers stability and local generalized Ulam–Hyers–Rassias stability for a class of fractional neutral differential equations. For the case  $I = [0, T]$  with  $f, g: I \times C \rightarrow \mathbb{R}^+$  are continuous functions on a closed interval or more generally compact sets, then they are bounded so we can replace the supremum by the maximum. In this case, norm is also called the maximum norm. Let  $X = \{x \in C[-\tau, 1]\}$  be a Banach space with  $\|x\| = \max_{t \in I} |x(t)|$ , where  $\tau$  is a nonnegative real number and  $z_t = z(t + \theta); -\tau \leq \theta \leq 0$ . We consider the following differential equation:

$$\begin{cases} {}^c D_0^\gamma [z(t) - g(t, z_t)] = z(t)f(t, z_t), & t \in I, \\ z(t) = \psi(t) \geq 0, & t \in [-\tau, 0], \end{cases} \tag{36}$$

where  ${}^c D_0^\gamma$  is a Caputo derivative with  $\gamma \in (0, 1)$  and  $\psi \in C([-\tau, 0], \mathbb{R}^+)$ . We focus on the following inequalities:

$$\|{}^c D_0^\gamma [u(t) - g(t, u_t)] - u(t)f(t, u_t)\| \leq \epsilon, \quad t \in I, \tag{37}$$

$$\|{}^c D_0^\gamma [u(t) - g(t, u_t)] - u(t)f(t, u_t)\| \leq \psi(t), \quad t \in I. \tag{38}$$

*Definition 6* (see [36]). Equation (36) is local generalized Ulam–Hyers stable if there exists a positive real number  $c_1$  such that, for each positive  $\epsilon$  and for every solution  $\bar{v} \in C([-\tau, T], \mathbb{R})$  of (37), there exists a solution  $\bar{w} \in C([-\tau, T], \mathbb{R})$  of (36) with  $\|\bar{v}(t) - \bar{w}(t)\| \leq c_1 \epsilon, t \in [-\tau, T]$ .

*Definition 7* (see [36]). Equation (36) is local generalized Ulam–Hyers–Rassias stable with respect to  $\psi$  if there exists  $c_{1\psi} > 0$  such that, for every solution  $\bar{v} \in C([-\tau, T], \mathbb{R})$  of (38), there exists a solution  $\bar{w} \in C([-\tau, T], \mathbb{R})$  of (36) with  $\|\bar{v}(t) - \bar{w}(t)\| \leq c_{1\psi} \psi(t), t \in [-\tau, T]$ .

*Remark 1.* A solution of differential equation is stable (asymptotically stable) if it attracts all other solutions with sufficiently close initial values. On the contrary, in Hyers–Ulam stability, we compare solution of given differential equation with the solution of differential inequality. We say solution of differential equation is stable if it stays close to solution of differential inequality. Hyers–Ulam stability may not imply the asymptotic stability.

*Remark 2* (see [37]). A function  $\bar{v} \in C([-\tau, T], \mathbb{R})$  is a solution of (37) if and only if there exists  $h \in C(I, \mathbb{R})$  such that

- (i)  $|h(t)| \leq \epsilon, t \in I$
- (ii)  ${}^c D_0^\gamma [u(t) - g(t, u_t)] = u(t)f(t, u_t) + h(t), t \in I$

Similarly, a function  $\bar{v} \in C([-\tau, T], \mathbb{R})$  is a solution of (38) if and only if there exists  $\tilde{h} \in C(I, \mathbb{R})$  such that

- (i)  $|\tilde{h}(t)| \leq \psi(t), t \in I$
- (ii)  ${}^c D_0^\gamma [u(t) - g(t, u_t)] = u(t)f(t, u_t) + \tilde{h}(t), t \in I$

*Remark 3.* Let  $0 < \gamma < 1$  and  $\bar{v} \in C([-\tau, T], \mathbb{R})$  be a solution of inequality (37); then,  $\bar{v}$  is a solution of the following inequality:

$$\left\| \bar{v}(t) - v(0) - g(t, \bar{v}_t + x_t) - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{v}(s) f(s, \bar{v}_s + x_s) ds \right\| \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)}. \tag{39}$$

From Remark 2, we have

$${}^c D_0^\gamma [u(t) - g(t, u_t)] = u(t) f(t, u_t) + h(t), \quad t \in I, \tag{40}$$

then

$$\begin{aligned} & u(t) - u(0) + g(0, \psi) - g(t, u_t) \\ &= \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} u(s) f(s, u_s) ds \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} h(s) ds. \end{aligned} \tag{41}$$

Therefore,

$$\begin{aligned} u(t) &= u(0) - g(0, \psi) + g(t, u_t) \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} u(s) f(s, u_s) ds \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} h(s) ds. \end{aligned} \tag{42}$$

By using the same process as in Section 3, i.e., let  $x(\cdot): [-\tau, T] \rightarrow \mathbb{R}^+$  be the function defined by

$$x(t) = \begin{cases} \psi(t), & t \in [-\tau, 0], \\ 0, & t \in I. \end{cases} \tag{43}$$

Therefore,  $x_0 = \psi$  and for every  $v \in C(I, \mathbb{R}^+)$  with  $v(0) = 0$ . We defined  $\bar{v}$  as

$$\bar{v}(t) = \begin{cases} 0, & t \in [-\tau, 0], \\ v(t), & t \in I. \end{cases} \tag{44}$$

Let  $x(\cdot)$  satisfies equation (42). We can decompose  $u(\cdot)$  as  $u(t) = \bar{v}(t) + x(t)$ , for  $0 \leq t \leq T$  so that  $u_t = \bar{v}_t + x_t$ , for every  $0 \leq t \leq T$ , where  $\bar{v}(t)$  is such that  $|\bar{v}(t)| < 1, \forall t \in I$  and the function  $v(\cdot)$  satisfies

$$\begin{aligned} \bar{v}(t) &= v(0) - g(0, \psi) + g(t, \bar{v}_t + x_t) \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{v}(s) f(s, \bar{v}_s + x_s) ds \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} h(s) ds. \end{aligned} \tag{45}$$

By following the conditions in Section 3, we get from (42) that

$$\begin{aligned} \bar{v}(t) &= v(0) + g(t, \bar{v}_t + x_t) \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{v}(s) f(s, \bar{v}_s + x_s) ds \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} h(s) ds. \end{aligned} \tag{46}$$

Hence,

$$\begin{aligned} & \left\| \bar{v}(t) - v(0) - g(t, \bar{v}_t + x_t) \right. \\ & \quad \left. - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{v}(s) f(s, \bar{v}_s + x_s) ds \right\| \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)}. \end{aligned} \tag{47}$$

If  $\bar{v} \in C([-\tau, T], \mathbb{R})$  is a solution of inequality (38) then  $\bar{v}$  is a solution of the following inequality:

$$\begin{aligned} & \left\| \bar{v}(t) - v(0) - g(t, \bar{v}_t + x_t) \right. \\ & \quad \left. - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{v}(s) f(s, \bar{v}_s + x_s) ds \right\| \\ & \leq \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \psi(s) ds, \\ & \leq \lambda_\psi \psi(t), \quad t \in I. \end{aligned} \tag{48}$$

Before stating stability results, let us take condition (H4) as

$$(H4) \quad c_2 + (T^\gamma/\Gamma(\gamma+1))(c_1 + L) < 1$$

**Theorem 7.** Suppose that (H4) is true and other two conditions

(a)  $f, g: I \times C \rightarrow \mathbb{R}^+, h \in C(I, \mathbb{R})$ , and  $|h(t)| \leq \epsilon, h > 0$ .

(b)  $f$  and  $g$  satisfy Lipchitz conditions:

$$\begin{aligned} & \|f(t, \bar{v}_t + x_t) - f(t, \bar{w}_t + x_t)\| \leq c_1 \|\bar{v}_t - \bar{w}_t\|, \\ & \|g(t, \bar{v}_t + x_t) - g(t, \bar{w}_t + x_t)\| \leq c_2 \|\bar{v}_t - \bar{w}_t\|, \end{aligned} \tag{49}$$

where  $c_1$  and  $c_2$  are nonnegative real numbers holds, then

(i) Equation (36) has a unique solution

(ii) Equation (36) is local generalized Ulam–Hyers stable

*Proof.* Let  $\bar{v} \in C([-\tau, T], \mathbb{R})$  be a unique solution of equation (37). Denote  $\bar{w} \in C([-\tau, T], \mathbb{R})$  as the unique solution of the equation

$$\begin{cases} {}^c D_0^\gamma [z(t) - g(t, z_t)] = z(t) f(t, z_t), & t \in I, \\ w(t) = v(t) \geq 0, & t \in [-\tau, 0]. \end{cases} \tag{50}$$

Then, we have

$$\begin{aligned} \bar{w}(t) &= w(0) + g(t, \bar{w}_t + x_t) \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}(s) f(s, \bar{w}_s + x_s) ds \\ &= v(0) + g(t, \bar{w}_t + x_t) \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}(s) f(s, \bar{w}_s + x_s) ds. \end{aligned} \tag{51}$$

We can see that  $\|\bar{v}(t) - \bar{w}(t)\| = 0$ , for  $t \in [-\tau, 0]$ . For  $t \in I$ , we obtain

$$\begin{aligned} \|\bar{v}(t) - \bar{w}(t)\| &= \left\| \bar{v}(t) - v(0) - g(t, \bar{w}_t + x_t) - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \bar{w}(s) f(s, \bar{w}_s + x_s) ds \right\| \\ &\leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \|g(t, \bar{v}_t + x_t) - g(t, \bar{w}_t + x_t)\| + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \|\bar{v}(s) f(s, \bar{v}_s + x_s) - \bar{w}(s) f(s, \bar{w}_s + x_s)\| ds \\ &\leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + c_2 \|\bar{v}_t - \bar{w}_t\| + \frac{c_1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \|\bar{v}_s - \bar{w}_s\| ds + \frac{L}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \|\bar{v}_s - \bar{w}_s\| ds. \end{aligned} \tag{52}$$

For  $t \in I$ , by using

$$\begin{aligned} \|\bar{v}_s - \bar{w}_s\| &= \max_{0 \leq s \leq t} |\bar{v}_s - \bar{w}_s| = \max_{0 \leq s \leq t} |\bar{v}(s+\theta) - \bar{w}(s+\theta)| \\ &= \max_{\theta \leq s+\theta \leq t+\theta} |\bar{v}(s+\theta) - \bar{w}(s+\theta)| \\ &\leq \max_{-\tau \leq s_1 \leq t} |\bar{v}(s_1) - \bar{w}(s_1)| \text{ where } s+\theta = s_1, \quad -\tau \leq \theta < 0 \\ &\leq \max_{-\tau \leq s_1 \leq T} |\bar{v}(s_1) - \bar{w}(s_1)|. \end{aligned} \tag{53}$$

Thus,

$$\begin{aligned} \|\bar{v}(t) - \bar{w}(t)\| &\leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \lambda_2 \max_{-\tau \leq t_1 \leq T} |\bar{v}(t_1) - \bar{w}(t_1)| + \frac{c_1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \\ &\quad \max_{-\tau \leq s_1 \leq T} |\bar{v}(s_1) - \bar{w}(s_1)| ds + \frac{L}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \max_{-\tau \leq s_1 \leq T} |\bar{v}(s_1) - \bar{w}(s_1)| ds \\ &\leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + c_2 \|\bar{v}(t_1) - \bar{w}(t_1)\| + \frac{c_1 + L}{\Gamma(\gamma+1)} \|\bar{v}(t_1) - \bar{w}(t_1)\| T^\gamma \\ &\leq \frac{T^\gamma \epsilon}{\Gamma(\gamma+1)} + \left( c_2 + \frac{T^\gamma}{\Gamma(\gamma+1)} (c_1 + L) \right) \|\bar{v}(t) - \bar{w}(t)\|. \end{aligned} \tag{54}$$

So,

$$\|\bar{v}(t) - \bar{w}(t)\| \leq c_1 \epsilon, \text{ where } c_1 = \frac{T^\gamma}{(1-c_2)\Gamma(\gamma+1) - (c_1+L)T^\gamma}. \tag{55}$$

Hence, equation (36) is local generalized Ulam–Hyers stable.  $\square$

**Theorem 8.** If (H4) is true and

- (a)  $f, g: I \times C \rightarrow \mathbb{R}^+$ ,  $\tilde{h} \in C(I, \mathbb{R})$ , and  $|\tilde{h}(t)| \leq \psi(t)$ ,  $h > 0$ .
- (b)  $f$  and  $g$  satisfy Lipschitz conditions:

$$\begin{aligned} \|f(t, \bar{v}_t + x_t) - f(t, \bar{w}_t + x_t)\| &\leq c_1 \|\bar{v}_t - \bar{w}_t\|, \\ \|g(t, \bar{v}_t + x_t) - g(t, \bar{w}_t + x_t)\| &\leq c_2 \|\bar{v}_t - \bar{w}_t\|, \end{aligned} \tag{56}$$

where  $c_1$  and  $c_2$  are nonnegative real numbers.

- (c) The function  $\psi \in C(I, \mathbb{R}^+)$  is an increasing function and there exists  $\lambda_\psi > 0$  such that

$$\frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \psi(s) ds \leq \lambda_\psi \psi(t), \quad t \in I, \tag{57}$$

holds, then equation (36) is local generalized Ulam–Hyers–Rassias stable with respect to  $\psi$ .

*Proof.* Following the same steps as in Theorem 7, we can find the result, i.e., here we will obtain

$$\begin{aligned} \|\bar{v}(t) - \bar{w}(t)\| &\leq \lambda_\psi \psi(t) + \lambda_2 \max_{-\tau \leq t_1 \leq T} |\bar{v}(t_1) - \bar{w}(t_1)| + \frac{c_1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \\ &\quad \cdot \max_{-\tau \leq s_1 \leq T} |\bar{v}(s_1) - \bar{w}(s_1)| ds + \frac{L}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \max_{-\tau \leq s_1 \leq T} |\bar{v}(s_1) - \bar{w}(s_1)| ds \\ &\leq \lambda_\psi \psi(t) + c_2 \|\bar{v}(t_1) - \bar{w}(t_1)\| + \frac{c_1 + L}{\Gamma(\gamma + 1)} \|\bar{v}(t_1) - \bar{w}(t_1)\| T^\gamma \\ &\leq \lambda_\psi \psi(t) + \left( c_2 + \frac{T^\gamma}{\Gamma(\gamma + 1)} (c_1 + L) \right) \|\bar{v}(t) - \bar{w}(t)\|. \end{aligned} \tag{58}$$

Therefore,

$$\begin{aligned} \|\bar{v}(t) - \bar{w}(t)\| &\leq c_{1\psi} \psi(t), \\ \text{where } c_{1\psi} &= \frac{\lambda_\psi \Gamma(\gamma + 1)}{(1 - c_2) \Gamma(\gamma + 1) - (c_1 + L) T^\gamma}. \end{aligned} \tag{59}$$

Hence, equation (36) is local generalized Ulam–Hyers–Rassias stable.  $\square$

*Example 1.* Consider the problem

$$\begin{cases} D_0^{1/2} [x(t) - x(t - \tau)] = tx(t) - x(t - \tau), & t \in [0, 1], \\ x(t) = 0, & t \in [-\tau, 0]. \end{cases} \tag{60}$$

One can easily prove that this problem has a unique fixed point and that it has a positive solution, i.e.,

### 5. Example

In this section, we present an example to explain the applicability of main results.

$$\begin{aligned} x(t) - x(t - \tau) &= \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{(1/2)-1} tx(t) ds - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{(1/2)-1} x(t-s) ds, \\ x(t) &= x(t - \tau) - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-(1/2)} tx(t) ds - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-(1/2)} x(t-s) ds. \end{aligned} \tag{61}$$

Define the operator  $P$  as

$$Px(t) = \begin{cases} x(t - \tau) - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-(1/2)} tx(t) ds - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-(1/2)} x(t-s) ds, & t \in [0, 1], \\ 0, & t \in [-\tau, 0], \end{cases}$$

$$\begin{aligned} |Px_1(t) - Px_2(t)| &\leq \|x_1(t) - x_2(t)\| + \frac{1}{\Gamma(1/2)} \|x_1(t) - x_2(t)\| \int_0^t (t-s)^{-(1/2)} ds + \frac{1}{\Gamma(1/2)} \|x_1(t) - x_2(t)\| \int_0^t t(t-s)^{-(1/2)} ds \\ &\leq \left( 1 + \frac{t^{3/2}}{\Gamma(3/2)} - \frac{t^{1/2}}{\Gamma(1/2)} \right) \|x_1 - x_2\|, \text{ with } \left( 1 + \frac{t^{3/2}}{\Gamma(3/2)} - \frac{t^{1/2}}{\Gamma(1/2)} \right) \leq 1, \quad \text{if } t \in [0, 1]. \end{aligned} \tag{62}$$

For stability, we take (60) as

$$\begin{cases} D_0^{1/2}[x(t) - x(t - \tau)] = tx(t) - x(t - \tau), & t \in [0, 1], \\ x(t) = z(t), & t \in [-\tau, 0], \\ x(0) = 0, \end{cases} \quad (63)$$

and the inequalities

$$\left| D_0^{1/2}[z(t) - z(t - \tau)] - tz(t) - z(t - \tau) \right| \leq \epsilon_1, \quad t \in [0, 1], \quad (64)$$

$$\left| D_0^{1/2}[z(t) - z(t - \tau)] - tz(t) - z(t - \tau) \right| \leq \psi(t), \quad t \in [0, 1]. \quad (65)$$

To prove that equation (63) is local generalized Ulam–Hyers stable, we take the conditions as in Theorem 7, i.e., a function  $z \in C([0, 1], \mathbb{R})$  is a solution of inequality (63) if and only if there exists  $h \in C([0, 1], \mathbb{R})$  such that

$$\begin{cases} |h(t)| \leq \epsilon_1, & t \in [0, 1], \\ D_0^{1/2}[z(t) - z(t - \tau)] = tz(t) - z(t - \tau) + h(t), & t \in [0, 1]. \end{cases} \quad (66)$$

Applying  $I_0^{1/2}$  on both sides, we obtain

$$\begin{aligned} & \left\| z(t) - z(t - \tau) - \frac{1}{\Gamma(3/2)} t^{3/2} z(t) \right. \\ & \left. + \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-1/2} z(t-s) ds \right\| \leq \frac{\epsilon}{\Gamma(3/2)}. \end{aligned} \quad (67)$$

Here, we take unique solution  $x(t)$  of equation (63) as

$$\begin{aligned} x(t) &= x(t - \tau) + \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-1/2} tx(s) ds \\ &\quad - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-1/2} x(t-s) ds, \end{aligned} \quad (68)$$

$$\begin{aligned} x(t) &= z(t - \tau) + \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-1/2} tx(s) ds \\ &\quad - \frac{1}{\Gamma(1/2)} \int_0^t (t-s)^{-1/2} x(t-s) ds. \end{aligned}$$

By further simplification, we obtain

$$\begin{aligned} \|z(t) - x(t)\| &\leq c_1 \epsilon; \\ c_1 &= \frac{1}{\Gamma(3/2) - t^{1/2}(t-1)}. \end{aligned} \quad (69)$$

Hence, (63) is local generalized Ulam–Hyers stable.

*Remark 4.* If we replace equation (66) by the inequality,

$$\begin{cases} |\tilde{h}(t)| \leq \psi(t), & t \in [0, 1], \\ D_0^{1/2}[z(t) - z(t - \tau)] = tz(t) - z(t - \tau) + \tilde{h}(t), & t \in [0, 1]. \end{cases} \quad (70)$$

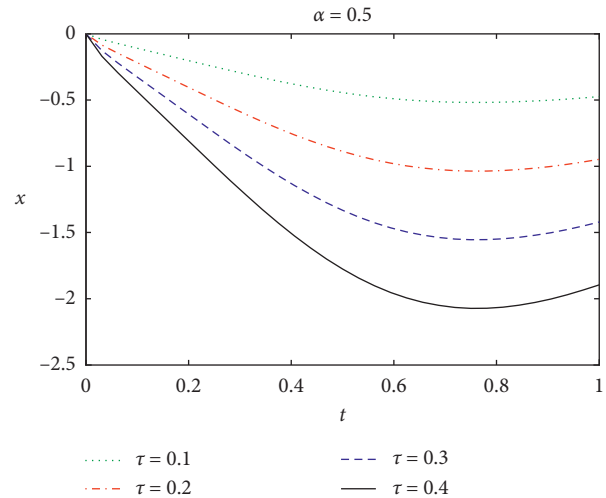


FIGURE 1: Trajectories for different values of  $\tau = 0.1, 0.2, 0.3,$  and  $0.4$  with  $\alpha = 0.5$ .

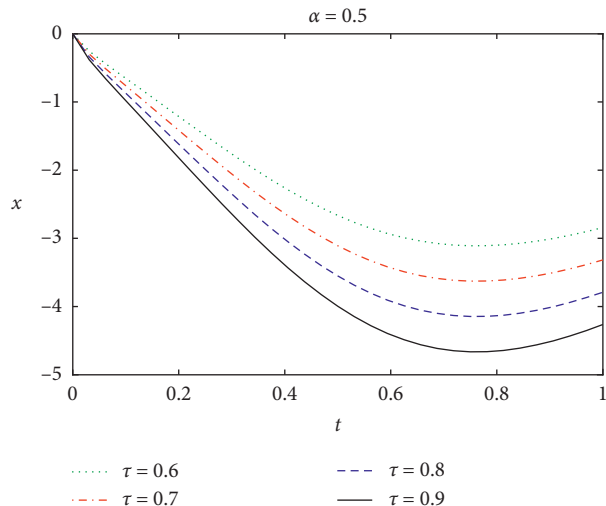


FIGURE 2: Trajectories for different values of  $\tau = 0.6, 0.7, 0.8,$  and  $0.9$  with  $\alpha = 0.5$ .

By repeating the same process as in the above example, one can easily verify the main results of Theorem 8.

### 6. Graphical Presentation

For graphical representation of the solution of the problem given in equation (60), we adopted Adams–Bashforth–Moulton scheme [38] to obtain the numerical solution for this fractional differential equation. To analyze the effect and contribution of time delayed factor, the modified predictor-corrector scheme [39] is incorporated for simulation. To check and demonstrate the stability of the consider model, graphical representation of the solution with different variations of the time delay factor along with other parameters is made. From the numerical results, we could note the UH-stability of the system with varying orders and delays. With the higher orders, the system will achieve the Ulam–Hyers stability more rapidly. It also holds



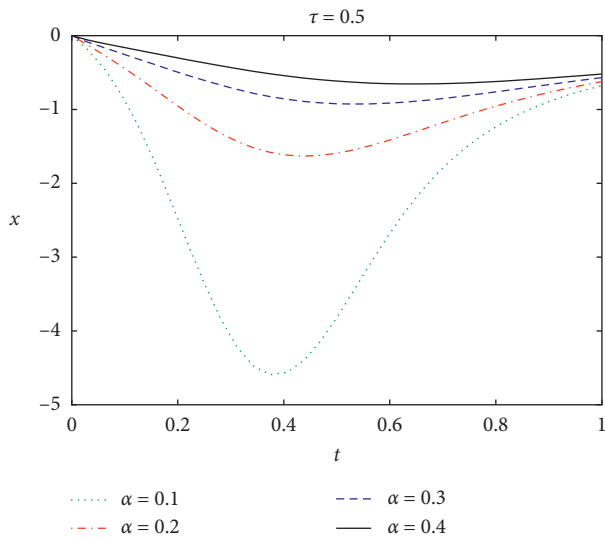


FIGURE 3: Trajectories for different values of  $\alpha = 0.1, 0.2, 0.3,$  and  $0.4$  with time delay  $\tau = 0.1$ .

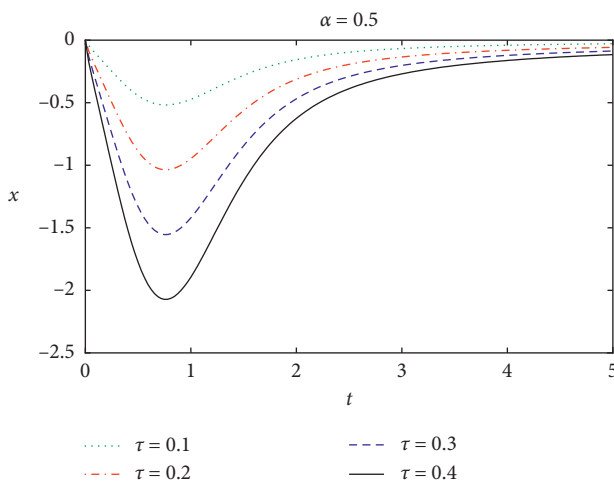


FIGURE 4: Trajectories for different values of  $\tau = 0.1, 0.2, 0.3,$  and  $0.4$  with  $\alpha = 0.5$ .

true with increasing delays. One can see the stability of system (60) in Figures (1)–(4) for various values of time delays and fractional derivatives.

### 7. Conclusions

We present some new results about the existence of positive solution for a class of nonlinear fractional differential equations with delay involving Caputo derivative. Leray–Schauder theorem, Arzela–Ascoli theorem, and Banach contraction principle are used for the fixed point property and existence of a solution. We also establish local generalized Ulam–Hyers stability and local generalized Ulam–Hyers–Rassias stability for the same class of nonlinear fractional neutral differential equations. The simulation of an example is also given to show the applicability of our results. The current concepts have significant applications since it means that if we are studying local generalized

Hyers–Ulam–Rassias stable (or Hyers–Ulam stable) system then one does not have to reach the exact solution. We just need to get a function which satisfies a suitable approximation inequality. In other words, local generalized Hyers–Ulam–Rassias stability (or Hyers–Ulam stability) guarantees that there exists a close exact solution. This is altogether useful in many applications where finding the exact solution is quite difficult such as optimization, numerical analysis, biology, and economics. It also helps, if the stochastic effects are small, to use deterministic model to approximate a stochastic one.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Authors’ Contributions

D. Nie, A. U. K. Niazi, and B. Ahmed contributed to and designed the study. A.U.K. Niazi gave the main idea of the manuscript. D. Nie wrote the first draft, and B. Ahmed performed the design of figures. All authors contributed to the manuscript revision, read, and approved the submitted version.

### Acknowledgments

This research is supported by Natural Science Research Project of Anhui Education Department (No. KJ2019A0875); School Level Scientific Research Project of Anhui Xinhua University (No. 2019zr005).

### References

- [1] W. G. Glöckle and T. F. Nonnenmacher, “A fractional calculus approach to self-similar protein dynamics,” *Biophysical Journal*, vol. 68, no. 1, pp. 46–53, 1995.
- [2] R. Hilfer, *Applications of Fractional Calculus in Physics*, World Scientific, Singapore, 2000.
- [3] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, “Theory and applications of fractional differential equations,” *North-Holland Mathematics Studies*, vol. 204, Elsevier, Amsterdam, Netherlands.
- [4] X. Li, S. Liu, and W. Jiang, “Positive solutions for boundary value problem of nonlinear fractional functional differential equations,” *Applied Mathematics and Computation*, vol. 217, no. 22, pp. 9278–9285, 2011.
- [5] S. Liu, X. Li, X. F. Zhou, and W. Jiang, “Synchronization analysis of singular dynamical networks with unbounded time-delays,” *Advances in Difference Equations*, p. 193, 2015.
- [6] J.-B. Liu, J. Zhao, and Z. X. Zhu, “On the number of spanning trees and normalized Laplacian of linear octagonal quadrilateral networks,” *International Journal of Quantum Chemistry*, vol. 119, Article ID 25971, 2019.
- [7] J.-B. Liu, J. Zhao, and Z.-Q. Cai, “On the generalized adjacency, Laplacian and signless Laplacian spectra of the

- weighted edge corona networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 540, no. 540, p. 123073, 2020.
- [8] J.-B. Liu, J. Zhao, H. He, and Z. Shao, “Valency-based topological descriptors and structural property of the generalized Sierpiński networks,” *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.
- [9] J.-B. Liu, Z.-Y. Shi, Y.-H. Pan, J. Cao, M. Abdel-Aty, and U. Al-Juboori, “Computing the Laplacian Spectrum of linear octagonal-quadrilateral networks and its applications,” *Polycyclic Aromatic Compounds*, pp. 1–12, 2020.
- [10] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley and Sons, Inc., New York, NY, USA, 1993.
- [11] I. Podlubny, *Fractional Differential Equations*, Academic Press, London, UK, 1999.
- [12] M. Rehman, R. Khan, and N. Asif, “Three point boundary value problems for nonlinear fractional differential equations,” *Acta Mathematica Scientia*, vol. 31B, no. 4, pp. 1337–1346, 2011.
- [13] A. Babakhani and V. Daftardar-Gejji, “Existence of positive solutions of nonlinear fractional differential equations,” *Journal of Mathematical Analysis and Applications*, vol. 278, no. 2, pp. 434–442, 2003.
- [14] V. Daftardar-Gejji, “Positive solutions of a system of non-autonomous fractional differential equations,” *Journal of Mathematical Analysis and Applications*, vol. 302, no. 1, pp. 56–64, 2005.
- [15] X. Su, “Positive solutions to singular boundary value problems for fractional functional differential equations with changing sign nonlinearity,” *Computers & Mathematics with Applications*, vol. 64, no. 10, pp. 3425–3435, 2012.
- [16] R. P. Agarwal, Y. Zhou, and Y. He, “Existence of fractional neutral functional differential equations,” *Computers & Mathematics with Applications*, vol. 59, no. 3, pp. 1095–1100, 2010.
- [17] M. Benchohra, J. Henderson, S. K. Ntouyas, and A. Ouahab, “Existence results for fractional order functional differential equations with infinite delay,” *Journal of Mathematical Analysis and Applications*, vol. 338, no. 2, pp. 1340–1350, 2008.
- [18] V. Lakshmikantham and A. S. Vatsala, “Basic theory of fractional differential equations,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 69, no. 8, pp. 2677–2682, 2008.
- [19] V. Lakshmikantham and A. Vatsala, “Theory of fractional differential inequalities and applications,” *Communications on Pure and Applied Analysis*, vol. 11, pp. 395–402, 2007.
- [20] V. Lakshmikantham and J. Devi, “Theory of fractional differential equations in Banach space,” *European Journal of Pure and Applied Mathematics*, vol. 1, pp. 38–45, 2008.
- [21] V. Lakshmikantham, “Theory of fractional functional differential equations,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 69, no. 10, pp. 3337–3343, 2008.
- [22] Y. Li and W. Jiang, “Fractional order nonlinear systems with delay in iterative learning control,” *Applied Mathematics and Computation*, vol. 257, pp. 546–552, 2015.
- [23] X.-F. Zhou, F. Yang, and W. Jiang, “Analytic study on linear neutral fractional differential equations,” *Applied Mathematics and Computation*, vol. 257, pp. 295–307, 2015.
- [24] Y. Li, S. Sun, D. Yang, and Z. Han, “Three-point boundary value problems of fractional functional differential equations with delay,” *Boundary Value Problems*, vol. 2013, no. 1, p. 38, 2013.
- [25] Y. Wen, X.-F. Zhou, Z. Zhang, and S. Liu, “Lyapunov method for nonlinear fractional differential systems with delay,” *Nonlinear Dynamics*, vol. 82, no. 1-2, pp. 1015–1025, 2015.
- [26] D. Guo and V. Lakshmikantham, *Nonlinear Problems in Abstract Cone*, Academic Press, Orlando, FL, USA, 1988.
- [27] M. Belmekki and K. Mekhalfi, “On fractional differential equations with State-dependent delay via Kuratowski measure of noncompactness,” *Filomat*, vol. 31, no. 2, pp. 451–460, 2017.
- [28] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York, NY, USA, 1993.
- [29] S. M. Ulam, *A Collection of Mathematical Problems*, Interscience Publishers, New York, NY, USA, 1968.
- [30] D. H. Hyers, “On the stability of the linear functional equation,” *Proceedings of the National Academy of Sciences*, vol. 27, no. 4, pp. 222–224, 1941.
- [31] T. M. Rassias, “On the stability of the linear mapping in Banach spaces,” *Proceedings of the American Mathematical Society*, vol. 72, no. 2, p. 297, 1978.
- [32] Y. Zhou, F. Jiao, and J. Li, “Existence and uniqueness for  $\alpha$ -type fractional neutral differential equations,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 7-8, pp. 2724–2733, 2009.
- [33] Y. Zhou, F. Jiao, and J. Li, “Existence and uniqueness for fractional neutral differential equations with infinite delay,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 7-8, pp. 3249–3256, 2009.
- [34] J. Wang, M. Fečkan, and Y. Zhou, “Ulam’s type stability of impulsive ordinary differential equations,” *Journal of Mathematical Analysis and Applications*, vol. 395, no. 1, pp. 258–264, 2012.
- [35] J. Wang, Y. Zhou, and M. Fečkan, “Nonlinear impulsive problems for fractional differential equations and Ulam stability,” *Computers & Mathematics with Applications*, vol. 64, no. 10, pp. 3389–3405, 2012.
- [36] Q. H. Alqifiary and S.-M. Jung, “On the Hyers-Ulam stability of differential equations of Second order,” *Abstract and Applied Analysis*, vol. 2014, pp. 1–8, 2014.
- [37] J. Wang, L. Lv, and Y. Zhou, “Ulam stability and data dependence for fractional differential equations with Caputo derivative,” *Electronic Journal of Qualitative Theory of Differential Equations*, vol. 63, no. 63, pp. 1–10, 2011.
- [38] K. Diethelm, N. J. Ford, and A. D. Freed, “A predictor-corrector approach for the numerical solution of fractional differential equations,” *Nonlinear Dynamics*, vol. 29, no. 1/4, pp. 3–22, 2002.
- [39] S. Bhalekar and V. Daftardar-Gejji, “A predictor-corrector scheme for solving nonlinear delay differential equations of fractional order,” *Fractional Calculus and Applied Analysis*, vol. 5, pp. 1–9, 2011.