

# Collective Behavior Analysis and Graph Mining in Social Networks

Lead Guest Editor: Fei Xiong

Guest Editors: Shirui Pan, Liang Wang, Huizhi Liang, and Hongshu Chen





---

# **Collective Behavior Analysis and Graph Mining in Social Networks**

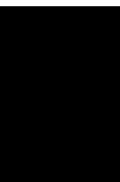
Complexity

---

## **Collective Behavior Analysis and Graph Mining in Social Networks**

Lead Guest Editor: Fei Xiong

Guest Editors: Shirui Pan, Liang Wang, Huizhi Liang, and Hongshu Chen




---

Copyright © 2021 Hindawi Limited. All rights reserved.

This is a special issue published in "Complexity." 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

Hiroki Sayama , USA

## Associate Editors

Albert Diaz-Guilera , Spain  
Carlos Gershenson , Mexico  
Sergio Gómez , Spain  
Sing Kiong Nguang , New Zealand  
Yongping Pan , Singapore  
Dimitrios Stamovlasis , Greece  
Christos Volos , Greece  
Yong Xu , China  
Xinggang Yan , United Kingdom

## Academic Editors

Andrew Adamatzky, United Kingdom  
Marcus Aguiar , Brazil  
Tarek Ahmed-Ali, France  
Maia Angelova , Australia  
David Arroyo, Spain  
Tomaso Aste , United Kingdom  
Shonak Bansal , India  
George Bassel, United Kingdom  
Mohamed Boutayeb, France  
Dirk Brockmann, Germany  
Seth Bullock, United Kingdom  
Diyi Chen , China  
Alan Dorin , Australia  
Guilherme Ferraz de Arruda , Italy  
Harish Garg , India  
Sarangapani Jagannathan , USA  
Mahdi Jalili, Australia  
Jeffrey H. Johnson, United Kingdom  
Jurgen Kurths, Germany  
C. H. Lai , Singapore  
Fredrik Liljeros, Sweden  
Naoki Masuda, USA  
Jose F. Mendes , Portugal  
Christopher P. Monterola, Philippines  
Marcin Mrugalski , Poland  
Vincenzo Nicosia, United Kingdom  
Nicola Perra , United Kingdom  
Andrea Rapisarda, Italy  
Céline Rozenblat, Switzerland  
M. San Miguel, Spain  
Enzo Pasquale Scilingo , Italy  
Ana Teixeira de Melo, Portugal

Shahadat Uddin , Australia  
Jose C. Valverde , Spain  
Massimiliano Zanin , Spain

# Contents

**Research on the Structure and Characteristics of the Overall Social Network of Professional Athletes**  
Shuqin Cui , Mingyou Gao , Yang Xun , Sai-Fu Fung , Yujiao Tan , Yu Zhang , Chenghao Wang , Huanqing Wang , and You Xiong 

Research Article (11 pages), Article ID 6484098, Volume 2021 (2021)

**Erratum to “A Hierarchical Attention Recommender System Based on Cross-Domain Social Networks”**

Rongmei Zhao , Xi Xiong , Xia Zu , Shenggen Ju, Zhongzhi Li, and Binyong Li



Erratum (1 page), Article ID 5323421, Volume 2021 (2021)

**Analysis of the Structural Characteristics of the Online Social Network of Chinese Professional Athletes**

Yue Wang , Qian Huang , Qiurong Wang , Yang Xun , Yujiao Tan , Shuqin Cui , Linxiao Ma, Penglin Huang , Meijuan Cao , and Bin Zhang 

Research Article (10 pages), Article ID 6647664, Volume 2021 (2021)

**Opinion Expression Dynamics in Social Media Chat Groups: An Integrated Quasi-Experimental and Agent-Based Model Approach**

Siyuan Ma  and Hongzhong Zhang 









Review Article (14 pages), Article ID 2304754, Volume 2021 (2021)

**Social Networks through the Prism of Cognition**

Radosław Michalski , Boleslaw K. Szymanski , Przemysław Kazienko , Christian Lebiere , Omar Lizardo , and Marcin Kulisiewicz 



Research Article (13 pages), Article ID 4963903, Volume 2021 (2021)

**A Review of the Research Progress of Social Network Structure**

Ning Li , Qian Huang , Xiaoyu Ge , Miao He , Shuqin Cui , Penglin Huang , Shuairan Li , and Sai-Fu Fung 


Review Article (14 pages), Article ID 6692210, Volume 2021 (2021)

**A Cognition Knowledge Representation Model Based on Multidimensional Heterogeneous Data**

Dong Zhong , Yi-An Zhu, Lanqing Wang , Junhua Duan, and Jiaxuan He




Research Article (17 pages), Article ID 8812459, Volume 2020 (2020)

**Research on the City Network Structure in the Yellow River Basin in China Based on Two-Way Time Distance Gravity Model and Social Network Analysis Method**

Duo Chai , Dong Zhang, Yonghao Sun, and Shan Yang


Research Article (19 pages), Article ID 6680954, Volume 2020 (2020)

**The Spread of Information in Virtual Communities**

Zhen Zhang , Jin Du, Qingchun Meng, Xiaoxia Rong , and Xiaodan Fan 

Research Article (15 pages), Article ID 6629318, Volume 2020 (2020)

### **Athlete Social Support Network Modeling Based on Modern Valence Bond Theory**

Ningshe Zhao 



Research Article (8 pages), Article ID 4392975, Volume 2020 (2020)

### **Characterization of 2-Path Signed Network**

Deepa Sinha  and Deepakshi Sharma 


Research Article (13 pages), Article ID 1028941, Volume 2020 (2020)

### **Legal Judgment Prediction Based on Multiclass Information Fusion**

Kongfan Zhu, Rundong Guo , Weifeng Hu, Zeqiang Li, and Yujun Li 


Research Article (12 pages), Article ID 3089189, Volume 2020 (2020)

### **Influence of Social Networks on Citizens' Willingness to Participate in Social Governance: Evidence from China**

Rui Nan and Fan Ouyang 


Research Article (16 pages), Article ID 3819514, Volume 2020 (2020)

### **Anatomy of Complex System Research**

Yang Wang  and Xinyue Wang



Research Article (11 pages), Article ID 5208356, Volume 2020 (2020)

### **A Tri-Attention Neural Network Model-Based Recommendation**

Nanxin Wang, Libin Yang , Yu Zheng, Xiaoyan Cai, Xin Mei, and Hang Dai



Research Article (10 pages), Article ID 3857871, Volume 2020 (2020)

### **An Adaptive Network Model to Simulate Consensus Formation Driven by Social Identity Recognition**

Kaiqi Zhang , Zinan Lv, Hai Feng Du , and Honghui Zou


Research Article (13 pages), Article ID 1742065, Volume 2020 (2020)

### **A Bichannel Transformer with Context Encoding for Document-Driven Conversation Generation in Social Media**

Yuanyuan Cai, Min Zuo , Qingchuan Zhang , Haitao Xiong, and Ke Li


Research Article (13 pages), Article ID 3710104, Volume 2020 (2020)

### **Adolescent Sports Behavior and Social Networks: The Role of Social Efficacy and Self-Presentation in Sports Behavior**

Lei Lei, Huifang Zhang , and Xin Wang

Research Article (10 pages), Article ID 4938161, Volume 2020 (2020)

### **The Collective Behaviors of Self-Excitation Information Diffusion Processes for a Large Number of Individuals**


Lifu Wang and Bo Shen 

Research Article (14 pages), Article ID 1861583, Volume 2020 (2020)

# Contents

---

## **Statistical Analysis of Dispelling Rumors on Sina Weibo**

Yue Wu, Min Deng, Xin Wen, Min Wang, and Xi Xiong 


Research Article (11 pages), Article ID 3176593, Volume 2020 (2020)

## **A Hierarchical Attention Recommender System Based on Cross-Domain Social Networks**

Rongmei Zhao , Xi Xiong , Xia Zu , Shenggen Ju, Zhongzhi Li, and Binyong Li



Research Article (13 pages), Article ID 9071624, Volume 2020 (2020)

## **Self-Presentation and Adolescent Altruistic Behaviors in Social Networks**

Yaling Zhu, Yue Shen, and Qiang Zhao 





Research Article (11 pages), Article ID 1719564, Volume 2020 (2020)

## **Attention with Long-Term Interval-Based Deep Sequential Learning for Recommendation**

Zhao Li , Long Zhang, Chenyi Lei, Xia Chen, Jianliang Gao , and Jun Gao




Research Article (13 pages), Article ID 6136095, Volume 2020 (2020)

## **Recommendation Algorithm in Double-Layer Network Based on Vector Dynamic Evolution Clustering and Attention Mechanism**

Jianrui Chen , Zhihui Wang , Tingting Zhu , and Fernando E. Rosas 

Research Article (19 pages), Article ID 5206087, Volume 2020 (2020)

## **Ranking Influential Nodes in Complex Networks with Information Entropy Method**

Nan Zhao , Jingjing Bao , and Nan Chen 

Research Article (15 pages), Article ID 5903798, Volume 2020 (2020)

## **Effective Data Transmission and Control Based on Social Communication in Social Opportunistic Complex Networks**

Weiyu Yang, Jia Wu , and Jingwen Luo

Research Article (20 pages), Article ID 3721579, Volume 2020 (2020)



## Research Article

# Research on the Structure and Characteristics of the Overall Social Network of Professional Athletes

Shuqin Cui <sup>1</sup>, Mingyou Gao <sup>2</sup>, Yang Xun <sup>3</sup>, Sai-Fu Fung <sup>4</sup>, Yujiao Tan <sup>5</sup>,  
Yu Zhang <sup>5</sup>, Chenghao Wang <sup>5</sup>, Huanqing Wang <sup>1</sup>, and You Xiong <sup>1</sup>

<sup>1</sup>Sport College, Xi'an University of Architecture and Technology, Xi'an 710311, China

<sup>2</sup>Yulin University, Yulin 719000, China

<sup>3</sup>School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710049, China

<sup>4</sup>City University of Hong Kong, Hong Kong

<sup>5</sup>Xi'an Physical Education University, Xi'an 710065, China

Correspondence should be addressed to Mingyou Gao; 350441003@qq.com

Received 26 June 2020; Revised 31 August 2020; Accepted 27 April 2021; Published 8 May 2021

Academic Editor: Fei Xiong

Copyright © 2021 Shuqin Cui 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 study chooses Chinese athletes as the research object and constructs the overall network of its social support network and discussion network. From the micro-, meso-, and macrolevels of the social network structure, the structure and characteristics of the athlete's overall social network are analyzed. Through research, we found that there is embeddedness, that is, the relevance, between society support networks, between society discussion networks, and between society support networks and society discussion networks. At the same time, in the athletes' social support network and social discussion network, some athletes have no contact with other players; they have no "power" in the group as well, so it is difficult to obtain network resources. We also found that there are small-world characteristics in the social network of Chinese professional athletes. The above findings will provide a deeper understanding of the peculiarities of athlete groups and have certain practical significance for improving athletes' daily training and life management conditions.

## 1. Research Background

For a long time, the cultivation of professional athletes in our country has mainly adopted the government-led athlete training system; that is, the government funded the establishment of sports schools at all levels and administratively intervened. Athletes competed on behalf of the local governments, and those with outstanding results were sent to provincial professional sports teams and then selected into the national team, which is a three-level training system for athletes under the guidance of the typical planned economy; that is, the national team is the first level, the provincial professional team is the second level, and the city and county amateur sports schools are the third level. This is also the difference between the current training system for professional athletes in my country and the training system for athletes based on club training abroad. When professional

athletes in our country in relatively independent training, learning, and living environment, they may have problems in terms of employment, cultural level, professional skills, and psychology [1–3]. Social network refers to the relatively stable relationship system formed by the interaction between individual members of the society, focusing on the interaction and connection between people [4–6]. Social interaction is carried out under definite circumstances. People have different ways of interaction in different situations. In different situations and different types of relationships, people's interaction process will also be different. Then, through the study of athletes' social network structure, it will further reveal the social interaction between athletes, which will have a long-term important role in improving athletes' athletic performance, mental health, and career planning. Social network is not only a research method but also a research object in itself. Social network research is mainly

carried out in two directions: Whole Network and Ego-Centric Network/Personal Network [7–9].

Social network theory emphasizes the importance of “relationship” and the interaction between individuals, breaking the conventional “attribute” research. While showing a strong empirical analysis ability, a bridge has been separately established between individual rational choice and social constraints and between microbehavior and macrophenomena. Social network theory has been widely used in many fields of social science research, such as management and sociology [10–13]. In a sense, the social performance of professional athlete groups is the effect of social networks and their sports. Social network analysis can be said to be a research paradigm, which is easier to enter the special lives of professional athletes [14].

The whole network can reveal the structure of the social network very well, so when studying the social network structure of professional athletes in our country, we use the method of the whole network structure to analyze the social network structure of professional athletes in our country.

## 2. Theoretical Analysis of the Characteristic Variables of the Whole Social Network of Professional Athletes

The whole network focuses on the relationships within the group and analyzes the interpersonal interaction and exchange patterns with the network concepts of compactness, distance, and centrality. For the whole network research of social networks, the purpose is to reveal the structural features of the social network. The whole social network structure can be divided into three levels: micro, meso, and macro [15].

**2.1. Microstructure.** Centrality is the main feature of the social network microstructure and one of the focuses of social network analysis. Centrality reflects the size of individual’s power in the network [5]. Centrality is usually measured by three indicators: degree centrality, closeness centrality, and betweenness centrality [16]. If the individuality has more degree centrality, it indicates that the individual is in a central position and thus has greater power. If the individuality has more closeness centrality, it means that he is very close to other points, so he does not depend on others. The betweenness centrality shows the extent to which an individual can connect or control others. Through the analysis of the microstructure, we can see the core leaders in the athletes group and whether there are isolated points in the group, so as to grasp the whole network of athletes.

**2.2. Mesostructure.** For human society, individuals combine into small groups through interpersonal interactions and use social networks to obtain various social resources embedded in them [17]. Relevant research shows that there is a small group phenomenon in the athlete group, but the research is mainly qualitative analysis and lacks quantitative research. This paper will analyze the structure of the small group of

athletes from the modular index of the community structure [18, 19].

**2.2.1. Newman Community Structure Definition.** Let the network node set  $V_p, V_q$  be the true subset of  $V$ ; that is,  $V_p \neq \emptyset, V_q \neq \emptyset$  and  $V_p \subset V, V_q \subset V$ ; if  $V_p \cap V_q = \emptyset$ , then  $A_{pq} = \{a_{ij}\}, i \in V_p, j \in V_q$ , and  $A_{pq} \subset A$ . Note that  $\|A\| = (1/2)\sum_i \sum_j a_{ij}$  is the number of node relationships. So when  $p \neq q$ ,  $\|A_{pq}\| = \sum_{i \in V_p} \sum_{j \in V_q} a_{ij}$ ,  $a_{ij} \in A$  is the number of relationships between subset  $V_p$  and subset  $V_q$ ; when  $p = q$ ,  $\|A_{pp}\| = (1/2)\sum_{i \in V_p} \sum_{j \in V_p} a_{ij}$ ,  $a_{ij} \in A$  is the number of internal relationships in the subset. Actually, Newman’s modularity index that measures the effectiveness and rationality of community structure division is

$$Q = \sum_{p=1}^m \left[ e_{pp} - \left( \sum_{q=1}^m e_{pq} \right)^2 \right]. \quad (1)$$

**2.3. Macrostructure.** The network indicators of all actors in a group reflect the macrostructure of the network, that is, the whole structure. Such network indicators mainly include density, out-degree centralization, in-degree centralization, and betweenness centralization [20, 21]. Density is the most commonly used measure in network analysis. It shows the difference of edges’ distributions between the measured graph and the completed graph. Degree centralization is different from centrality. Centrality is to describe the “power” of a single actor from the perspective of an individual, while centralization reveals the whole centrality of the network graph. In fact, the centralization measures the centrality difference of all actors. Out-degree centralization, in-degree centralization, and betweenness centralization measure in the network the difference in all actors’ ability to expand relationships, reputation, and the ability to act as a bridge or control resource between any other two actors [22, 23].

Centralization measures the whole centrality of the network. It is a kind of “variance” of the centrality of each individual in the overall network. There are out-degree centralization, in-degree centralization, and betweenness centralization. The calculation formula of the centralization [18, 24] is

$$C = \frac{\sum_{i=1}^n (C_{\max} - C_i)}{\max[\sum_{i=1}^n (C_{\max} - C_i)]}. \quad (2)$$

The members of the professional athlete’s social support network and social discussion network overlap to a great extent and whether the two types of networks are embedded in each other can be analyzed through traditional social networks. After as a whole revealing the embeddability of the two types of networks, the macrostructure of the social support network and the social discussion network is analyzed separately through indicators such as network density. Complex network analysis is mainly to reveal the features of the small world. The small-world features (macrostructure) are based on the mesostructure and

calculate two feature indicators, the clustering coefficient and the path length, and reveal the features of the relationship transmission of the athlete's social network by comparing with the index of the random network of the same scale and the same relationship [24–27].

The clustering coefficient reflects the degree of aggregation of network nodes, referring to the probability that two other nodes connected to the same node are also connected [28, 29]. The greater the clustering coefficient, the more stable the relationship between network members. The calculation formula of the average clustering coefficient is

$$C = \frac{1}{n} \sum_{i=1}^n \sum_{j,l \in \Gamma_i} \frac{X(j,l)}{\#\Gamma_i(\#\Gamma_i - 1)/2} \quad (3)$$

Average path length refers to how many steps one node in the overall network averages to reach another node [30, 31]. If we define the shortest distance between node  $i$  and node  $j$  as  $d(i, j)$ , the average path length is calculated as

$$L = \frac{1}{n} \sum_{i=1}^n \sum_{j \neq i} \frac{d(i, j)}{n-1} \quad (4)$$

### 3. Research Design

**3.1. Data Sources and Research Methods.** The data used by research come from a sampling survey of professional athletes in Zhangba Training Base affiliated to Shaanxi Provincial Sports Bureau. This training base uses a three-level training system for athletes under the guidance of a typical planned economy. It is not only a large training base in Shaanxi Province but also a relatively large and concentrated training items in the country. Zhangba Training Base is located in relatively independent training, learning, and living environment, including Shaanxi Gymnastics Management Center (competitive team and competitive gymnastics team), Shaanxi Martial Arts Center (Martial arts Taolu and Sanda team), Shaanxi Heavy Competitive Center (wrestling, weightlifting, and judo team), Shaanxi football Center (men's and women's football team), and Shaanxi Swimming Center (swimming team and diving team). Nearly 1000 professional athletes are training here. Therefore, it has certain representation to select Zhangba Training Base of Shaanxi Province as the investigation place of Chinese professional athletes and to take samples from the athletes in the base [32].

We sent out 144 questionnaires, all of which recover accurately and effectively. Among 144 professional athletes, the proportion of male athletes is more, accounting for 68.8%. The athletes were relatively young, the maximum age was 30 years old, and the minimum age was 12 years old, with an average age of 18.8. In the geographical area, we were mainly divided into outside Shaanxi Province, Xi'an city, and other cities in Shaanxi Province, of which the proportion of people outside Shaanxi Province was the largest, accounting for about 50%. In sports programs, we chose five sports centers to accommodate athletes engaged in different types of training as much as possible, so that the survey data

are more objective and reasonable. In terms of education, most of them have received high school education or above, and nearly one-third of them have received college or undergraduate education, which shows that the athletes' education level is relatively high. At the same time, we also investigated the reasons why they chose to be an athlete at the beginning. For most of the athletes, their reasons are "their own interests" and "suggestions of family and friends." The overall network data are divided into two aspects. One is the overall support network (A–E), which includes A-actual support, B-emotional support, C-society communication support, D-employment support, and E-achievement support. Second is the whole discussion network (F–H), F-income discussion network, G-professional discussion network, and H-marriage discussion network. Investigation process: first, encode all the members in the network to form the overall network boundary, then investigate each individual in the network to determine the network members who have communication with themselves in the network boundary, and ask the respondents to answer whether the members of the network give their actual support, emotional support, and society intercourse support [33].

**3.2. The Construction of the Overall Network Topology Diagrams of Professional Athletes.** Based on the principles of complex network analysis and system engineering, we have constructed 8 whole social networks of professional athletes from two aspects: overall support network and whole discussion network. Among them, A–E are the overall support networks, which include A-actual support, B-emotional support, C-society communication support, D-employment support, and E-achievement support. F–H are whole discussion networks, which includes F-income discussion network, G-professional discussion network, and H-marriage discussion network. The topology diagrams built with UCINET are shown in Figure 1.

## 4. Analysis of the Overall Social Network Characteristics of Professional Athletes

### 4.1. Microstructure Analysis

**4.1.1. Degree Centrality.** Table 1 provides the degree centrality of the social support network and social discussion network of professional athletes. From the perspective of minimum value, it can be seen that each professional athlete's network has isolated points and they do not have support behavior with anyone. From the view of maximum value, the out-degree is generally greater than the in-degree, indicating that the professional athletes are more likely to ask for help than to accept it passively. Among the five kinds of support networks, the mean value of degree centrality of actual support and society intercourse support is larger which means that athletes support each other more in daily life and society intercourse, while the mean value of degree centrality of employment support network is the smallest, then, in turn, emotional support and achievement support.

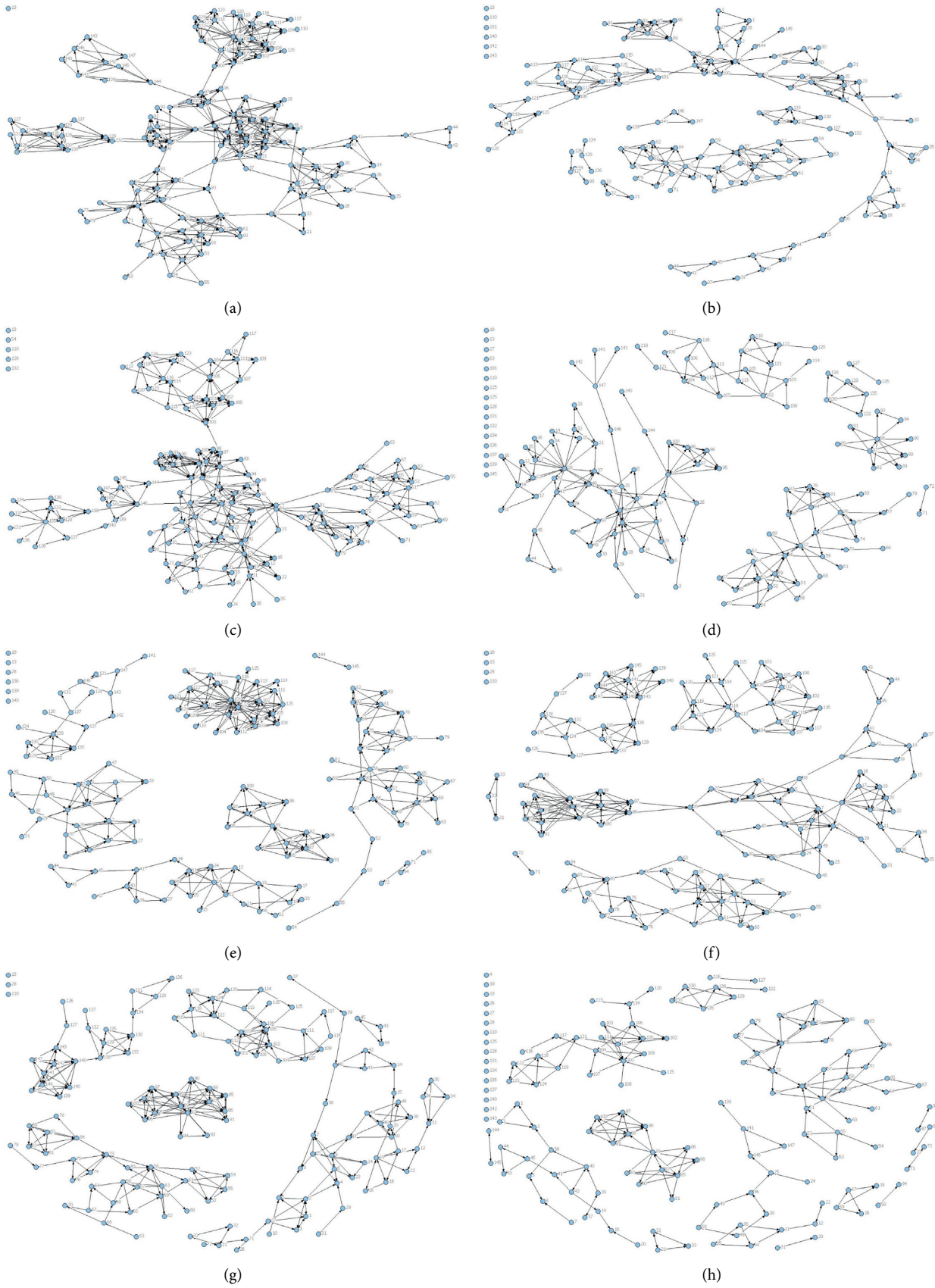


FIGURE 1: Topology diagrams of the overall social network of professional athletes.

TABLE 1: Social network degree centrality of professional athletes.

		Out-degree				In-degree			
		Mean	Standard deviation	Minimum value	Maximum value	Mean	Standard deviation	Minimum value	Maximum value
A	C <sub>AD</sub>	3.471	3.349	0	24	3.471	2.268	0	10
	C <sub>RD</sub>	2.054	1.981	0	14.201	2.054	1.342	0	5.917
B	C <sub>AD</sub>	2.012	1.919	0	10	2.012	1.659	0	6
	C <sub>RD</sub>	1.19	1.135	0	5.917	1.19	0.982	0	3.55
C	C <sub>AD</sub>	2.853	3.048	0	14	2.853	2.108	0	8
	C <sub>RD</sub>	1.688	1.804	0	8.284	1.688	1.247	0	4.734
D	C <sub>AD</sub>	1.606	2.146	0	15	1.606	1.527	0	7
	C <sub>RD</sub>	0.95	1.27	0	8.876	0.95	0.904	0	4.142
E	C <sub>AD</sub>	2.347	3.211	0	24	2.347	2.004	0	9
	C <sub>RD</sub>	1.389	1.9	0	14.201	1.389	1.186	0	5.325
F	C <sub>AD</sub>	2.488	2.59	0	15	2.488	1.974	0	8
	C <sub>RD</sub>	1.472	1.533	0	8.876	1.472	1.168	0	4.734
G	C <sub>AD</sub>	2.347	2.407	0	14	2.347	1.819	0	7
	C <sub>RD</sub>	1.389	1.424	0	8.284	1.389	1.076	0	4.142
H	C <sub>AD</sub>	1.541	1.719	0	9	1.541	1.511	0	7
	C <sub>RD</sub>	0.912	1.017	0	5.325	0.912	0.894	0	4.142

Note. CAD is an absolute value, CRD is a relative value, and relative value is a percentage.

Among the three discussion networks, there are athletes who do not participate in any discussion network. Similarly, from the view of maximum value, the out-degree is generally greater than the in-degree, indicating that professional athletes actively take part in discussions more than passively. Among them, there are more discussions on income and professional sports team management, but less on marriage, which has something to do with the younger age of professional athletes and the sensitive topics of marriage involved.

From the perspective of the characteristics of degree centrality, several adjustment ideas are proposed for the practice of training guidance and life guidance in the future: first, according to the athlete's strong initiative, the coaching staff can set up a special person to communicate with the athletes for help information and provide as many assistance items as possible according to the athlete's type of help. Secondly, due to the strong initiative of the athletes to participate in the discussion, the coaching staff needs to conduct psychological counseling and dredging frequently. Finally, the coaching staff should pay attention to some athletes who tend to self-isolate and guide them to establish social support and discussion links with other teammates.

**4.1.2. Betweenness Centrality.** Table 2 provides the betweenness centrality of the social support network and social discussion network of professional athletes. It can be seen that, in the five social support networks, the mean value of the actual support network is obviously higher than that of other networks. It can be understood that more athletes play the role of bridge in the actual support network, followed by society intercourse, emotional support, employment support, and achievement support. In three society discussion networks, we can see that the minimum value is zero and we can assume that, in every topic, there are athletes who cannot fit in other two discussion relationships; that is, they do not

have the ability to control the discussion relationship. Athletic groups have different centrality in social support and social discussion which means that they have different "power" structure in the group. Whether social support or social discussion, there are a small number of athletes in isolation. They do not have any "power." Overall speaking, the centrality of actual support is high, while that of emotional support is poor. For social discussion, the more common the topic is, the higher the centrality of the corresponding network is. On the contrary, the more private the topic is, the lower the centrality of the corresponding network is [33, 34].

According to the analysis of the centrality characteristics, the coaching staff should encourage athletes to increase the necessary contact between teammates, especially to increase emotional and spiritual communication, which will help improve the network "power" structure. At the same time, we must focus on the more important nodes in the "bridge relationship," which will have a multiplier effect on team management.

**4.1.3. Closeness Centrality.** Table 3 provides the closeness centrality of the society support network and society discussion network of professional athletes. The closeness centrality requires that the graph must be a connected graph; that is, any two points in the network can be linked together through a certain path. According to the data, we can find that the professional athlete network is relatively sparse and there are isolated points and disconnected subnetworks [33, 34].

According to the characteristics of the sparse network, the information circulation space of the athlete group is worthy of further exploration. It is recommended to guide the players to strengthen mutual learning and cooperation among different objects in the process of training and life guidance.

TABLE 2: Social network betweenness centrality of professional athletes.

		Mean value	Standard deviation	Total	Minimum value	Maximum value
A	C <sub>AD</sub>	249.69	444.21	36704	0	2062.55
	C <sub>RD</sub>	1.18	2.10	173.38	0	9.74
B	C <sub>AD</sub>	55.42	110.33	8147	0	610.83
	C <sub>RD</sub>	0.26	0.52	38.48	0	2.89
C	C <sub>AD</sub>	125.62	276.75	18466	0	1954.97
	C <sub>RD</sub>	0.59	1.31	87.23	0	9.24
D	C <sub>AD</sub>	18.79	40.76	2762	0	181.00
	C <sub>RD</sub>	0.09	0.19	13.05	0	0.86
E	C <sub>AD</sub>	16.73	39.81	2459	0	297.50
	C <sub>RD</sub>	0.08	0.19	11.62	0	1.41
F	C <sub>AD</sub>	27.57	51.61	4052	0	408.05
	C <sub>RD</sub>	0.13	0.24	19.14	0	1.93
G	C <sub>AD</sub>	20.25	38.54	2976	0	289.17
	C <sub>RD</sub>	0.10	0.18	14.06	0	1.37
H	C <sub>AD</sub>	13.71	42.36	2016	0	306.83
	C <sub>RD</sub>	0.07	0.20	9.52	0	1.45

Note. C<sub>AD</sub> is an absolute value, C<sub>RD</sub> is a relative value, and relative value is a percentage.

TABLE 3: The closeness centrality of the society of professional athletes.

	Out-degree				In-degree			
	Mean value	Standard deviation	Minimum value	Maximum value	Mean value	Standard deviation	Minimum value	Maximum value
A	13893.24	5115.757	8082	21462	13893.24	5305.243	1397	21462
	1.214	0.444	0.68	1.806	1.534	1.838	0.68	10.451
B	18842.74	2227.149	14785	21462	18842.74	2613.77	13037	21462
	0.786	0.095	0.68	0.987	0.793	0.129	0.68	1.12
C	16592.83	2736.615	12063	21462	16592.83	5873.589	3684	21462
	0.904	0.147	0.68	1.21	1.164	0.878	0.68	3.963
D	20197.58	1458.593	16003	21462	20197.58	1453.342	16701	21462
	0.727	0.058	0.68	0.912	0.727	0.056	0.68	0.874
E	20107.12	1148.178	17448	21462	20107.12	1303.288	17545	21462
	0.729	0.043	0.68	0.837	0.729	0.05	0.68	0.832
F	19534.95	1688.046	14983	21462	19534.95	1398.181	17469	21462
	0.754	0.071	0.68	0.974	0.751	0.054	0.68	0.836
G	19935.84	1286.519	16737	21462	19935.84	1056.923	18150	21462
	0.736	0.05	0.68	0.872	0.734	0.039	0.68	0.804
H	20457.69	1192.434	17448	21462	20457.69	1220.173	17456	21462
	0.716	0.045	0.68	0.837	0.716	0.046	0.68	0.836

4.2. *Mesostructure Analysis.* According to the view of Newman, the larger the  $Q$  value is, the better the detection effect of community structure and the clearer the community structure are. Generally,  $Q$  value is between 0.3 and 0.7. However, the modular  $Q$  value is problematic and the three networks in Figure 2 are used as an example to illustrate.

Network 1 and network 2 are very similar with a clear community structure and the same community; that is, each subgroup has five interconnected nodes. The only difference is that network 1 has two communities, while network 2 has four communities. Their  $Q$  values are, respectively, 0.4036 and 0.6591. For network 3, the same as network 1, there are two communities, but the number of nodes in the community is different, one is 5, the other is 10, and the  $Q$  value is 0.2688. The above example is quite different from the view of

Newman. Therefore, the  $Q$  value of the community structure can be affected by the network structure [34].

Table 4 is the optimal modularity index of social network community structure detection for athletes. It can be seen that the modularity value of all other networks except social communication support and employment support network is greater than 0.8, indicating that community structure widely exists in professional athlete network.

The modular value of social discussion network is generally greater than that of the social support network, which shows that once personal privacy issues are involved, athletes are often only discussing with people in small communities and are not willing to communicate with other community members [35]. In a word, the differences in personality, growth background, education level, and other aspects will naturally lead to individual differences of

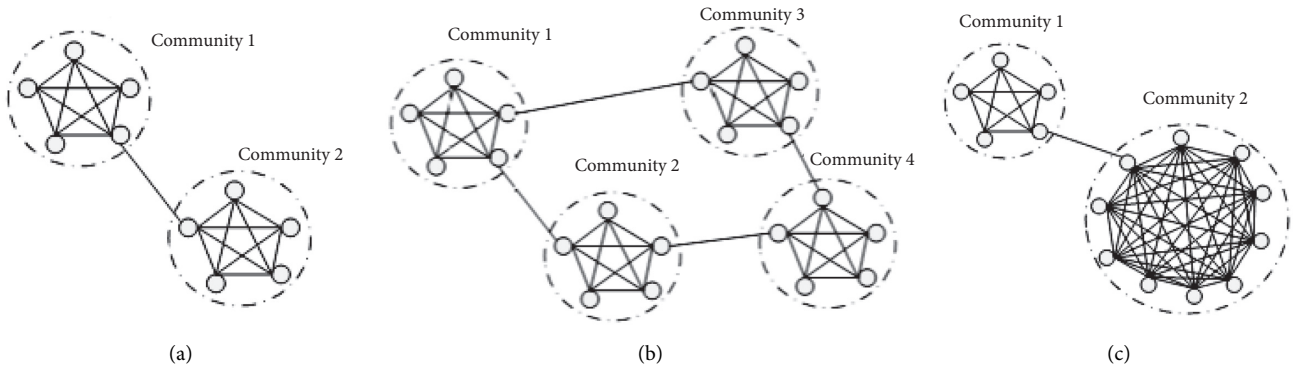


FIGURE 2: Schematic diagram of community structure. (a) Network1. (b) Network2. (c) Network3.

TABLE 4: Modular indicators of professional athletes.

	A	B	C	D	E	F	G	H
Modulari	0.8011	0.8392	0.7899	0.7991	0.8147	0.8406	0.8577	0.8367

athletes, which determines that the social network structure of the athletes group cannot be exactly the same [36].

Based on the above analysis results, it is recommended to strengthen the improvement of athletes' individual attributes, such as the improvement of education level and the psychological guidance of lonely personality, which will contribute to the beingness of the network structure of the entire athlete team.

### 4.3. Macrostructural Analysis

**4.3.1. Degree Centralization.** Table 5 provides the degree centralization of the professional athletes' social support network and social discussion network. The in-degree centralization is a measure of the difference between the number of other individuals in the network and the individual as a whole. When the difference is greater in degree, it shows that the difference in the popularity of individual Internet users is greater. The out-degree centralization measures the size of the difference in the number of individuals in contact with other individuals in the network as a whole. When the difference is greater in degree, it shows that the difference between the network members' ability to contact others is greater and the centralization is greater. We can see that the centralization of achievement support networks and actual support networks is relatively greater, in the top five social support networks. It indicates that these two social networks are relatively concentrated, then followed by employment support, society intercourse support, and emotional support. Among the three social discussion networks, the centralization is in sequence income discussion network, profession discussion network, and marriage discussion network [36].

Based on the above analysis, it is recommended to strengthen the support of athletes' social support appeals as a

whole, encourage more communication between players, and further reduce the differences in passive communication.

**4.3.2. Betweenness Centralization.** Table 6 provides the betweenness centralization of professional athletes' social support network and social discussion network, which is a measure of the difference in the number of paths between each individual in the overall network. The betweenness centralization is to measure the size of the difference between the individual in the overall network as the communication bridge between the other two individuals. When the difference is greater in degree, it shows that the minority have the greater ability to control the individual communication in the network [37]. It can be seen from the table that, in the actual support and society intercourse support network, the betweenness centralization is the greater, indicating that some of the athletes have greater individual control capabilities, while there is not much difference in the discussion network [38].

According to the analysis of the betweenness centralization, the research believes that the management of actual support relationships and social interaction support relationships can focus on the important bridge node sets, because they will affect some of the surrounding nodes. For other relationships, such as employment support relationships, this is not necessary.

**4.3.3. Embeddability.** The method of embeddability analysis is to check the correlation of the two relationship matrices. The specific method is the QAP method, that is, the Quadratic Assignment Procedure. This method based on resampling has been widely used in social networks. In the specific calculation, QAP compares the similarity of each element in the two square matrices, gives the correlation

TABLE 5: Degree centralization of professional athletes' social network.

	A	B	C	D	E	F	G	H
Out	13.78%	5.29%	7.38%	9.06%	14.68%	8.36%	7.78%	4.98%
In	4.13%	2.53%	3.24%	3.55%	4.34%	3.53%	2.96%	3.60%

TABLE 6: Betweenness centralization of professional athletes' social network.

	A	B	C	D	E	F	G	H
Betweenness centralization	8.62%	2.64%	8.70%	0.77%	1.34%	1.81%	1.28%	1.39%

coefficient between the two matrices and performs a non-parametric test on the coefficient, and calculates that the correlation coefficient after random replacement is greater than the probability of the actual correlation coefficient. The principle is that all the numerical values in each matrix are treated as a long vector, and each vector contains  $n(n-1)$  numbers (diagonal numbers are not considered) and then calculates the correlation coefficient. If the probability is less than 0.05, it means that there is a correlation between the two matrices. The calculation of the correlation coefficient is done using the software UCINET [34, 39].

It can be seen from Table 7 that there is a significant correlation between professional athletes' social networks, indicating that there are overlapping members in eight different types of networks, and there is also a strong correlation between athletes' society support network and society discussion network. It shows that a considerable part of the support network members is also members of its discussion network.

According to the analysis of embeddability, the research believes that, in order to reduce research and management costs, the structure of the social discussion network can be inferred through the athlete's social support network in the case of nonstrict requirements.

**4.3.4. Density Structure Analysis.** Presentation of the overall network structure features is mainly density, and the density in a directed network is  $d = (\sum_{i=1}^n \sum_{j=1}^n X_{ij}) / n(n-1)$ ,  $n$  represents the size of the network.

It can be seen from Table 8 that the relationship density of the social support of professional athletes is different. Among the five social networks, the actual support network and the social interaction support network are relatively dense, followed by achievement support network, emotional support network, and employment support network. For the three discussion networks, the discussion density on income topics is higher, followed by professional discussions, and the discussion density on marriage is lower. In general, no matter the social support network or discussion network, professional athletes are lacking emotional communication [40].

According to the above analysis, we should focus on the actual support relationship and social interaction support relationship between players in the work, because athletes use these two relationships more, and other relationships take second place.

TABLE 7: Social network relevance of professional athletes.

	AA	BB	CC	DD	EE	FF	GG	HH
AA	1	0.641	0.670	0.515	0.557	0.622	0.633	0.568
BB	0.641	1	0.667	0.543	0.591	0.638	0.646	0.670
CC	0.670	0.667	1	0.551	0.559	0.635	0.622	0.591
DD	0.515	0.543	0.551	1	0.520	0.507	0.511	0.517
EE	0.557	0.591	0.559	0.520	1	0.588	0.574	0.563
FF	0.622	0.638	0.635	0.507	0.588	1	0.750	0.647
GG	0.633	0.646	0.622	0.511	0.574	0.750	1	0.641
HH	0.568	0.670	0.591	0.517	0.563	0.647	0.641	1

Note. The number of random transformations in the calculation process is 5000, and all  $p$  values are 0.000. The output  $\text{Prop} \geq 0$  (these randomly calculated correlation coefficients are greater than or equal to the actual correlation coefficient) is 0.000, and  $\text{Prop} \leq 0$ , the probability is 1.000.

**4.3.5. Theoretical Analysis of Structural Characteristics of Small World.** The small-world phenomenon reveals the most effective way of information transfer between real-world actors. At present, there is not any precise definition of the small-world network. It is generally believed that if the average distance  $L$  between two nodes in the network increases logarithmically with the number of network nodes  $n$ ,  $L \propto \ln n$ , the network is called a small-world network [41, 42].

The clustering coefficient and the average path length are two important characteristic indicators for investigating small-world networks. The clustering coefficient reflects the degree of aggregation of the network nodes and refers to the possibility that two other nodes connected to the same node are also connected. The larger the clustering coefficient, the more stable the relationship between network members.

The average path length refers to the average number of steps, which an individual in the overall network gets to another individual.

It can be seen from Tables 9 and 10 that the society discussion network of athletes has obvious small-world characteristics, while the society support network has high clustering, but the average path length is long. This characteristic shows that income discussion, professional discussion, and marriage discussion spread quickly in the social network because of the short path and the small spread range of the athletes' social network, with a shorter path, but a smaller range of relationship transmission. In contrast, social support relationships have a longer propagation path and wider relationships. It explains that the most private discussion relationship has the shortest propagation path, so



TABLE 8: Social network density of professional athletes.

	A	B	C	D	E	F	G	H
Density	0.0275	0.0159	0.0226	0.0127	0.0186	0.0197	0.0186	0.0122

TABLE 9: Average clustering coefficient of professional athletes' social network.

	A	B	C	D	E	F	G	H
Local clustering	0.405	0.352	0.382	0.370	0.427	0.423	0.423	0.387
Transitive clustering	0.298	0.320	0.331	0.221	0.313	0.375	0.411	0.347

TABLE 10: Average path length of professional athletes' social network.

	A	B	C	D	E	F	G	H
Path length	5.663	4.025	4.672	3.138	2.781	3.059	2.911	2.967

the propagation speed may be faster. Among the five social support networks, the average path of achievement support is the shortest, indicating that the support for professional technology among athletes is better, and everyone communicates more. This is consistent with the shorter average path of the professional discussion network in the discussion network.

The small-world characteristics in the athlete network can be effectively used in athlete management. In the "small world," it is conducive to the rapid establishment of contacts and the effective dissemination of management information. It is also conducive to the role of group wisdom.

## 5. Conclusion

This study uses complex network analysis methods to analyze the micro-, meso-, and macrostructural characteristics of the athlete's overall social network (a total of eight overall networks including society support network and society discussion network). Using complex society network analysis methods, this study analyzes the complex overall network characteristics of professional athletes' social network including three aspects: the first is scale-free feature and connection tendency analysis at microlevel; second, it improves the community structure detection algorithm based on complex network analysis and analyzes the community structure of athletes' social network; third, it analyzes the small-world characteristics of athletes' social network. Specific analysis methods include structural hole analysis, cohesive subgroup analysis, core-periphery structure analysis, small-world feature analysis, community structure analysis, and scale-free feature analysis.

The study found the following:

- (1) The microstructure of social networks of professional athletes shows that no matter whether it is a society support network or a society discussion network, some athletes have no contact with other team members. They have no "power" in the group and it is difficult to obtain network resources.

Because access to resources is more urgent than the exchange of ideas between athletes, the centrality of athletes' society support networks is higher than that of society discussion networks as a whole. For the social support network, the actual support and society support are more central than emotional support, because the return of the energy invested in the former is relatively greater, and insufficient emotional support is the main problem of the athletic team. In terms of collective discussions, income and professionalism are the main topics of discussion. Compared with the topic of marriage, the topic of marriage is relatively few discussions, which are related to the age of athletes and the applicable regulations of sports teams. Some athletes act as "information bridge" in group communication and have a dominant role in the acquisition of network resources or the spread of ideas.

- (2) The mesostructure of professional athletes' social network shows that the structure of athletes' small groups is sparse, and social discussions mainly take place among small groups with communal support relationships. There are more factions in the social support network than in the communal discussion network; correspondingly, there are more cliques overlapping members in the social support network. Overlapping cliques become more active than other members. No matter what social support or social discussion, some athletes are relatively isolated. They neither communicate with other athletes in the same structural position nor actively communicate with other groups. They only accept passively and are in a passively dominant position. Society discussion networks are easier to form small groups than society support networks.
- (3) The macrostructure of professional athletes' society networks shows that there is embeddedness, that is, the relevance, between society support networks, between society discussion networks, and between

society support networks and society discussion networks. On the whole, athletes are more practicing and active when seeking social support or social discussion and more passive when receiving social support or social discussion. At the same time, the social network of professional athletes has the characteristics of a small world, and society discussion network is more obvious than the society support network. The characteristics of the small world reflect the rich local connection and few random long-distance connections in the society network of athletes, indicating that there are short paths in society relations. Particularly for the societal discussion network, this short path is conducive to the rapid establishment of society support relationships between individual athletes.

- (4) The research on the complex society network of professional athletes shows that professional athletes' society intercourse has self-organizing features. A few athletes are in an advantageous position in interpersonal communication and have more society relationships. In the process of communication, diverse small groups will be formed. Most of them are small-scale groups, and the small-world phenomenon still exists.

## Data Availability

The original data used in this study are the questionnaire data obtained from the survey. The original 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 work was supported in part by the 2020 Science and Technology Project in Yulin National High Tech Industrial Development Zone under Grant CXY-2020-17.




## References

- [1] Q. Li, B. Liu, and Y. Zhao, "Construction of statistical metadata in the big data environment," *Journal of Statistics and Information*, vol. 35, no. 3, pp. 14–20, 2020.
- [2] R. Zhang, L. Liu, X. Tang, and B. Zhang, "Research on commodity retail price index prediction based on web search data in the context of big data," *Journal of Statistics and Information*, vol. 35, no. 11, pp. 49–56, 2020.
- [3] Z. Pang and M. Zeng, "Has internet use affected youth subjective well-being? An analysis from CGSS 2010-2015 data," *Journal of Xi'an University of Finance and Economics*, vol. 33, no. 3, pp. 71–77, 2020.
- [4] J. Scott and P. J. Carrington, *The SAGE Handbook of Social Network Analysis*, SAGE, Thousand Oaks, CA, USA, 2011.
- [5] A. Biswas and B. Biswas, "Investigating community structure in perspective of ego network," *Expert Systems with Applications*, vol. 42, no. 20, pp. 6913–6934, 2015.
- [6] J. O. Kereri and C. M. Harper, "Social networks and construction teams: literature review," *Journal of Construction Engineering and Management*, vol. 145, no. 4, Article ID 03119001, 2019.
- [7] H. Yang, Z. Hao, and K. Wang, "Social network and HIV transmission China AIDS," *STD*, vol. 1, pp. 47–50, 2003.
- [8] J. Liu, *Introduction to Social Network Analysis*, China Social Science Publishing, Beijing, China, 2004.
- [9] R. Leonard, D. Horsfall, and K. Noonan, "Identifying changes in the support networks of end-of-life carers using social network analysis," *BMJ Supportive & Palliative Care*, vol. 5, no. 2, pp. 153–159, 2015.
- [10] T. W. Valente and S. R. Pitts, "An appraisal of social network theory and analysis as applied to public health: challenges and opportunities," *Annual Review of Public Health*, vol. 38, no. 1, pp. 103–118, 2017.
- [11] D. Hussein, S. Park, S. N. Han, and N. Crespi, "Dynamic social structure of things: a contextual approach in CPSS," *IEEE Internet Computing*, vol. 19, no. 3, pp. 12–20, 2015.
- [12] B. S. Hanson, B. Liedberg, and B. Owall, "Social network, social support and dental status in elderly Swedish men," *Community Dentistry & Oral Epidemiology*, vol. 22, no. 5, pp. 331–337, 2015.
- [13] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 3804–3816, 2020.
- [14] F. W. Frey, E. Abrutyn, and D. S. Metzger, "Focal networks and HIV risk among African-American male intravenous drug users/needle RH," *Social Networks, Drug Abuse, and HIV Transmission*, S. G. Genser and R. T. Trotter, Eds., Springer, Berlin, Germany, 1995.
- [15] G. C. Loury, "A dynamic theory of racial income differences," in *Women, Minorities, and Employment Discrimination*, P. A. Wallace and A. M. La Mond, Eds., D.C. Heath and Company, Lexington, MA, USA, 1977.
- [16] D. Mehrle, A. Strosser, and A. Harkin, "Walk-modularity and community structure in networks," *Network Science*, vol. 3, no. 3, pp. 348–360, 2015.
- [17] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, Article ID 026113, 2004.
- [18] L. Nan, *Social Capital-Theory on Social Structure and Action*, Shanghai People's Publishing House, Shanghai, China, 2005.
- [19] J. Luo, *Lecture Notes on Social Network Analysis*, Social Sciences Literature Press, Beijing, China, 2005.
- [20] B. Bollobás, *Random Graphs*, Academic Press Inc., London, UK, 1985.
- [21] A.-L. Barabási, R. Albert, and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.
- [22] N. Gupta, V. Ho, J. M. Pollack, and L. Lai, "A multilevel perspective of interpersonal trust: individual, dyadic, and cross-level predictors of performance," *Journal of Organizational Behavior*, vol. 37, no. 8, pp. 1271–1292, 2016.
- [23] M. A. J. van Duijn, T. A. B. Snijders, and B. J. H. Zijlstra, "p2: a random effects model with covariates for directed graphs," *Statistica Neerlandica*, vol. 58, no. 2, pp. 234–254, 2004.
- [24] C. J. Mitchell, "The concept and use of social networks," in *Social Network in Urban Situations*, C. J. Mitchell, Ed., Manchester University Press, Manchester, UK, 1969.
- [25] P. Carrington, "Network as personal communities," in *Social Structure: A Network Approach*, B. Wellman, Ed., Cambridge University Press, Cambridge, UK, 1988.

- [26] Welman and W. Zhang, "Network analysis: from method and metaphor to theory and essence," *Foreign Sociology*, vol. 17, no. 2, p. 8, 1994.
- [27] H. Qian and X. Zhang, "Theoretical analysis of Chinese athletes' social network and social support," *Journal of Xi'an Institute of Physical Education*, vol. 29, no. 1, pp. 1–5, 2012.
- [28] N. R. Taylor, "Small world network strategies for studying protein structures and binding," *Computational & Structural Biotechnology Journal*, vol. 5, no. 6, pp. 1–7, 2013.
- [29] A. Roxin, H. Riecke, and S. A. Solla, "Self-sustained activity in a small-world network of excitable neurons," *Physical Review Letters*, vol. 92, no. 19, Article ID 198101, 2004.
- [30] Y. B. Xie, T. Zhou, and B. H. Wang, "Scale-free networks without growth," *Physica A: Statistical Mechanics and Its Applications*, vol. 387, no. 7, pp. 1683–1688, 2017.
- [31] S. C. Ponten, F. Bartolomei, and C. J. Stam, "Small-world networks and epilepsy: graph theoretical analysis of intracerebrally recorded mesial temporal lobe seizures," *Clinical Neurophysiology*, vol. 118, no. 4, pp. 918–927, 2007.
- [32] Z. He, "A review of foreign social support network research," *Foreign Social Sciences*, vol. 24, no. 1, pp. 76–81, 2001.
- [33] J. Liu, *Fa Village Social Support Network: A Perspective of Holistic Research*, Social Science Literature Press, Beijing, China, 2006.
- [34] Y. Bian, "The source and function of urban residents' social capital: network views and survey findings," *Chinese Social Sciences*, vol. 3, pp. 136–146, 2004.
- [35] Y. Hu, F. Xiong, S. Pan, X. Xiong, L. Wang, and H. Chen, "Bayesian personalized ranking based on multiple-layer neighborhoods," *Information Sciences*, vol. 542, pp. 156–176, 2021.
- [36] H. Qian and X. Zhang, "The impact of personal characteristics of professional athletes on the structure of social support network in my country," *Journal of Shanghai Institute of Physical Education*, vol. 38, no. 6, pp. 59–64, 2014.
- [37] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4186–4199, 2020.
- [38] W. Wang, "Social capital and personal capital of Chinese urban residents," *Sociological Research*, vol. 21, no. 3, pp. 151–166, 2006.
- [39] M. Tortoriello, "The social underpinnings of absorptive capacity: the moderating effects of structural holes on innovation generation based on external knowledge," *Strategic Management Journal*, vol. 36, no. 4, pp. 586–597, 2015.
- [40] H. Qian and X. Zhang, "Research on the structural characteristics of individual center network of Chinese professional athletes," *Journal of Xi'an Institute of Physical Education*, vol. 31, no. 1, pp. 1–6, 2014.
- [41] L. I. Bo, W. Lei, and School of P.E., "Feasibility analysis of passing performance in football match by social network analysis," *Journal of Beijing Sport University*, vol. 40, no. 8, pp. 112–119, 2017.
- [42] Z. Jianyu, X. Xi, and S. Yi, "A simulation research on small world effect in knowledge flow's network evolution," *Management Review*, vol. 27, no. 5, pp. 70–81, 2015.

## Erratum

# Erratum to “A Hierarchical Attention Recommender System Based on Cross-Domain Social Networks”

**Rongmei Zhao** <sup>1</sup>, **Xi Xiong** <sup>1,2</sup>, **Xia Zu** <sup>3</sup>, **Shenggen Ju**,<sup>4</sup> **Zhongzhi Li**,<sup>1</sup> and **Binyong Li**<sup>1</sup>

<sup>1</sup>School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China

<sup>2</sup>School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China

<sup>3</sup>School of Management, Chengdu University of Information Technology, Chengdu 610103, China

<sup>4</sup>College of Computer Science, Sichuan University, Chengdu 610065, China

Correspondence should be addressed to Xia Zu; [sunnyzu@163.com](mailto:sunnyzu@163.com)

Received 2 February 2021; Accepted 2 February 2021; Published 19 February 2021

Copyright © 2021 Rongmei Zhao 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 the article titled “A Hierarchical Attention Recommender System Based on Cross-Domain Social Networks” [1], authors Xia Zu and Shenggen Ju both had incorrect affiliations.

The correct affiliations are as follows:

Xia Zu: School of Management, Chengdu University of Information Technology, Chengdu 610103, China.

Shenggen Ju: College of Computer Science, Sichuan University, Chengdu 610065, China.

The error was introduced during the production process of the article, and Hindawi apologizes for causing this error in the article.

## References

- [1] R. Zhao, X. Xiong, X. Zu, S. Ju, Z. Li, and B. Li, “A hierarchical attention recommender system based on cross-domain social networks,” *Complexity*, vol. 2020, Article ID 9071624, 13 pages, 2020.

## Research Article

# Analysis of the Structural Characteristics of the Online Social Network of Chinese Professional Athletes

Yue Wang <sup>1</sup>, Qian Huang <sup>1</sup>, Qiurong Wang <sup>1</sup>, Yang Xun <sup>2</sup>, Yujiao Tan <sup>1</sup>,  
Shuqin Cui <sup>3</sup>, Linxiao Ma,<sup>3</sup> Penglin Huang <sup>4</sup>, Meijuan Cao <sup>1</sup> and Bin Zhang <sup>1</sup>

<sup>1</sup>*Xi'an Physical Education University, Xi'an 710065, China*

<sup>2</sup>*School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710049, China*

<sup>3</sup>*Sport College, Xi'an University of Architecture and Technology, Xi'an 710311, China*

<sup>4</sup>*Shandong Sport University, Jinan 250063, China*

Correspondence should be addressed to Qian Huang; [huangqian168@126.com](mailto:huangqian168@126.com)

Received 9 October 2020; Revised 16 January 2021; Accepted 8 February 2021; Published 19 February 2021

Academic Editor: Hongshu Chen

Copyright © 2021 Yue 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.

In response to the lack of research on the online social network structure of athletes, elements of research on the online social network structure of athletes were constructed based on the whole network perspective and through the study of the characteristics of the whole online social network structure of athletes, in order to provide reference for the physical and mental health development of athletes from a new perspective. Data were collected through questionnaires, and several software programs were used to preprocess and analyse the collected data. Through the analysis of the online whole network structure, it was found that the network density of the online support network was generally greater than that of the online discussion network, and athletes still showed stronger practical support demands and behaved more rationally in the process of training and learning life, while from the perspective of the relationship structure, the athletes' family and classmates' online support is weaker than that in previous studies; in terms of the whole network, strong relationships still dominate in this population, while attention should be paid to the impact of weak relationships.

## 1. Introduction

China's competitive sports implement the "one piece of mind," "one organization," and "consistent training" training system, and athletes are in a relatively independent and closed environment for a long time. The training, learning, and living environment of athletes is relatively independent and closed for a long time, and the breadth, depth, and integration of real social interaction are subject to certain limitations [1–5]. With the rapid development of the Internet, online social network platforms such as WeChat, Weibo, QQ, and social networking sites have been integrated into people's daily lives, and the huge Internet user base has formed an intricate virtual society-Online Social Network

(OSN) [6–14], which provides athletes with a more modern and diverse way to socialize [15–17].

Athletes, in particular Chinese competitive athletes, have very specific social relationships. The offline social relationships of athletes described in this article are relatively closed, i.e., closed for a certain amount of time and closed for a certain amount of space. The reality is that athletes training in preparation for an event, for example, will remain in one or several relatively closed environments for a period of time (e.g., weeks or even months), and connections to the outside world are negligible compared to the large number of internal connections. Due to the influence of offline social interaction, this group is also relatively closed online compared to the general public.

Online social networks are having an increasing impact on athletes' social support and social inclusion, as athletes learn about society, access information, express their aspirations, and have greater emotional resonance through the Internet [18].

Due to the special nature of athletes' social relationships, online social networking has become an important way for athletes to connect with others. The survey data shows that athletes' social networks are limited in time and space compared to those of the general public, in that their online time is relatively limited and fixed, and their online activities are relatively concentrated. The challenge in studying this online social network of this population is that online data is not easy to collect and there is a lack of references, so we continue to use traditional questionnaires. The innovation of this paper is to conduct an online social support network study based on the previous offline social support network, which is an exploratory approach to fill the gap in this field in China, and is groundbreaking in terms of both the athlete population and the research methodology.

Current research on athletes' social networks focuses on the attribute variables and structural analysis of real social networks, and their online social data and research are lacking. Athletes' online social network is an open and complex system, a large-scale collection consisting of connected relationships between athletes and other network citizens, including teammates, coaches, family members, hometown, friends, and even strangers [19, 20]. Thinking about and understanding athletes' online social structure from different dimensions and studying the inner laws of professional athletes' network behaviour at multiple levels (micro-, meso-, and macrolevel) can help grasp the mechanisms of athletes' behavioural patterns and provide references for improving their online social support networks, promoting social integration, and further improving the sports system.

Online social networks are an extension of social networks [21, 22], a form of social networks existing in another space, and their structural characteristics can be analysed by applying the methods of social network analysis. According to the different boundaries of social networks, the study of social network structure is divided into two directions: individual networks and whole networks [23, 24]. Individual networks locate the network with individuals as the centre and the individual characteristics and behavioural concepts of concern, while whole networks consist of a set of specific individuals and the interrelationships between them, and the network members have relatively obvious boundaries [25–27]. The network as a whole exhibits a certain structure and has an impact on the actors within it [28]. In other words, the “structuralist” perspective of “social structure” influences the “dynamic role.” This study examines the structural characteristics of athletes' whole online social network by collecting online data from athletes in order to provide a new perspective on athletes' physical and mental health development.

## 2. Elements of Athletes' Online Social Structure

The term “online social network” refers to the network of social users in the Internet [29]. According to the EU study on social computing, online social networks can be divided into four categories: instant messaging (WeChat, QQ, etc.), online social networking (Facebook, Renren, etc.), micro-blogging (Twitter, Weibo, etc.), and shared space (forums, blogs, etc.). The latest data released by CNNIC shows that applications such as WeChat, short video, and live streaming are increasingly integrated with the lives of Chinese Internet users. In the preliminary stage of this study, a sample of 451 athletes were surveyed on online social network applications, and the most frequently used application by respondents was instant messaging, accounting for 93.1%, of which WeChat had the highest usage rate.

As shown in Figure 1, online social network for athletes consists of two categories: social support network and social discussion network [23, 30]. The online social support network is based on the Van der Poole social support classification method, and it is divided into social online support, emotional online support, achievement online support, and practical online support. Based on the Institute of Population and Development of Xi'an Jiaotong University's social network of migrant workers and the social network of athletes under Huang Qian's complex network, the online social discussion network is analysed [23, 31] from four aspects: employment online discussion, revenue online discussion, professional online discussion, and love and marriage online discussion. The questionnaire data are collected by positioning method.

## 3. Data Collection

Whole social networks require closed groups, and such sociologically significant networks are generally not large in size. This study sampled a closed group formed by 147 professional taekwondo athletes from the Shaanxi Boxing and Taekwondo Sports Management Centre. The questionnaire was divided into two parts, the first part investigated the basic information of the athletes, including gender, date of birth, and total number of WeChat friends, and the second part designed a list of 147 athletes' names and code names according to the roster method [32, 33].

The data collection was conducted from January 10 to January 13, 2020. To ensure accurate and valid data, the author interpreted the questions on site, and in the second part of the questionnaire filling, athletes first filled in their own code names according to the code name list and then selected other athletes who had correspondence with them in the test questions item by item and filled in their corresponding code names. Through a survey of coaches and relevant researchers and drawing on Huang Qian et al.'s analysis of athletes' whole social networks [3, 4], the following measurement questions about athletes' online social networks were selected: (1) Who do you interact with on

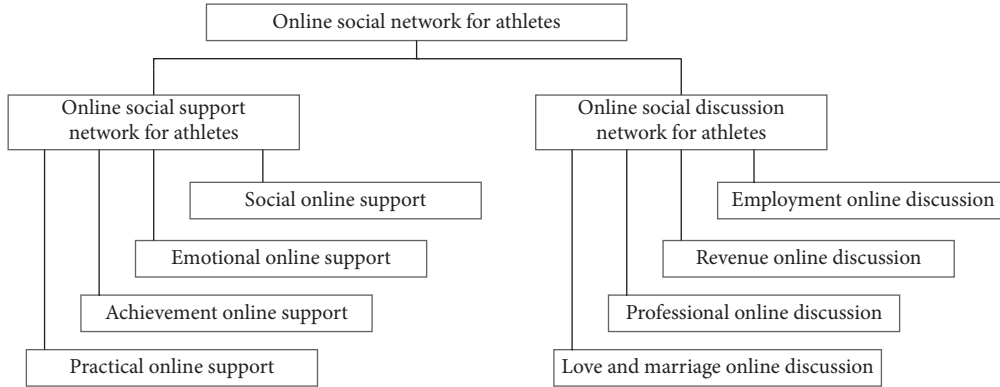


FIGURE 1: Organization structure of social network of athletes.

online platforms (conversations, holiday greetings, file transfers, etc.)? (2) In online chats, to whom do you confide your feelings? (3) Who do you turn to for help with small daily tasks on online platforms (including online advice, online payments)? (4) Who are the people who push information online (private messages, group notifications, or sharing with friends) that would be helpful to your training? (5) Who would you like to talk to online about issues related to training or sports team management, etc.? (6) Who would you like to discuss marriage topics with online? (7) Who would you like to discuss employment issues with online? (8) With whom would you like to discuss issues such as income online? Respondents ticked each item according to the list, and the relationship was recorded as “1” if selected and “0” otherwise. Eight relationship matrices consisting of “0” and “1” are formed, for example, in Table 1. The questionnaires were collected on site and four athletes were away for training during the collection period, so the author instructed them to complete the questionnaires online via the Internet. After the questionnaires were collected, the “informants” were used to understand the general situation of the training team, and the questionable data was returned to the author. 147 valid questionnaires were collected. In this paper, the data cleaning and two-party relationship data statistics are done by self-programmed Excel VBA, while other characteristic index calculation and topology drawing are processed by Ucinet and Gephi software.

To illustrate the relationship matrix for athletes using an online practical support network as an example, it is sufficient to intercept parts of the matrix that illustrate the problem, as it is large. In Table 1,  $N$  represents the total number of nodes and  $a_i$  ( $i = 1, 2, \dots, 20$ ) denote athlete nodes, and due to space limitations, only  $a_1$  to  $a_{20}$  are taken. Using formula (1) to calculate its density which is only 0.0273, we can find that the matrix is a sparse matrix ( $L$  represents the actual number of connections;  $N = 147$  is the number of nodes; generally less than 0.05 is considered as a sparse matrix), in line with the basic characteristics of social networks; secondly, the distribution of connections has no

obvious regularity (i.e., nondiagonal array, triangular array, unit array, etc.), which is a random network. Other thematic networks  $B \sim H$  are similar.

$$d(G) = \frac{L}{N(N-1)}. \quad (1)$$

#### 4. Analysis of the Characteristics of Whole Online Social Network

**4.1. Topology of Athletes Network.** In order to visualise the relationship structure of the eight online social networks for athletes, the online social network topology diagram [34–36] was drawn using Gephi as shown in Figure 2, with each of the eight subgraphs labelled with the letters  $A \sim H$ . In each directed graph, the arrow indicates the direction of the relationship; e.g.,  $a_{23}$  points to  $a_8$  in subgraph B, which means that  $x_{m23}$  is willing to confide in  $x_{m8}$  about what is on his mind. These eight subnetworks are derived from the eight relationship matrices  $A$  to  $H$  collated through the questionnaire as isomorphic to Table 1.

As can be seen, the online practical support network is the most densely connected and all nodes are connected to each other, indicating that athletes are most socially active and connected online. The practical online support network is also denser, with no isolated points. The emotional online support, employment online discussion, achievement online support, revenue online discussion, professional online discussion, and love and marriage online discussion networks all have isolated points, especially the love and marriage online discussion networks, which have the sparsest relationships, the most isolated points, and multiple subnetworks that are not connected to each other.

**4.2. Network Viscosity.** Network viscosity reflects the degree of overall network node association and overall cohesion, and is measured by network density and network distance.

TABLE 1: Athlete online relationship matrix.

$N$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$a_{15}$	$a_{16}$	$a_{17}$	$a_{18}$	$a_{19}$	$a_{20}$	...
$a_1$	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_2$	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_3$	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_4$	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	...
$a_5$	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_6$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_7$	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_8$	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_9$	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_{10}$	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_{11}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_{12}$	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	...
$a_{13}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
$a_{14}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	...
$a_{15}$	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	...
$a_{16}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	...
$a_{17}$	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	0	1	...
$a_{18}$	0	0	0	0	1	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	...
$a_{19}$	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	1	...
$a_{20}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

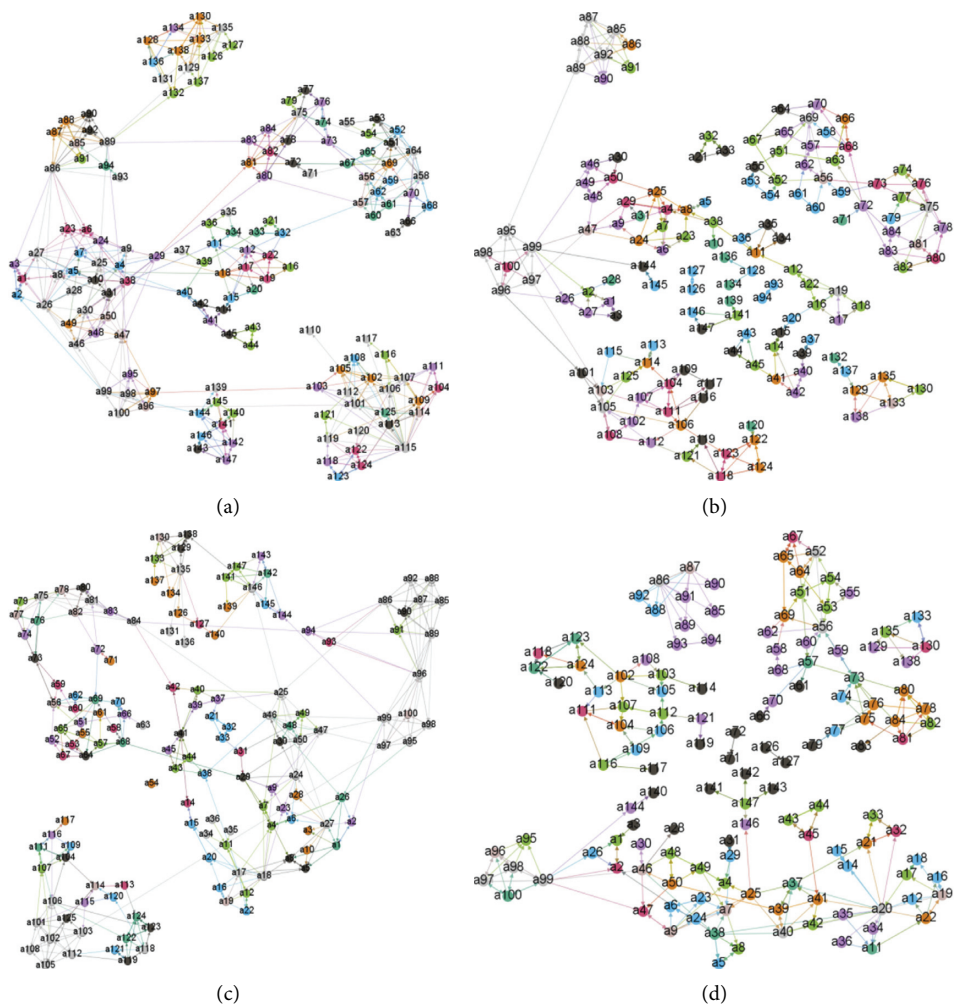


FIGURE 2: Continued.



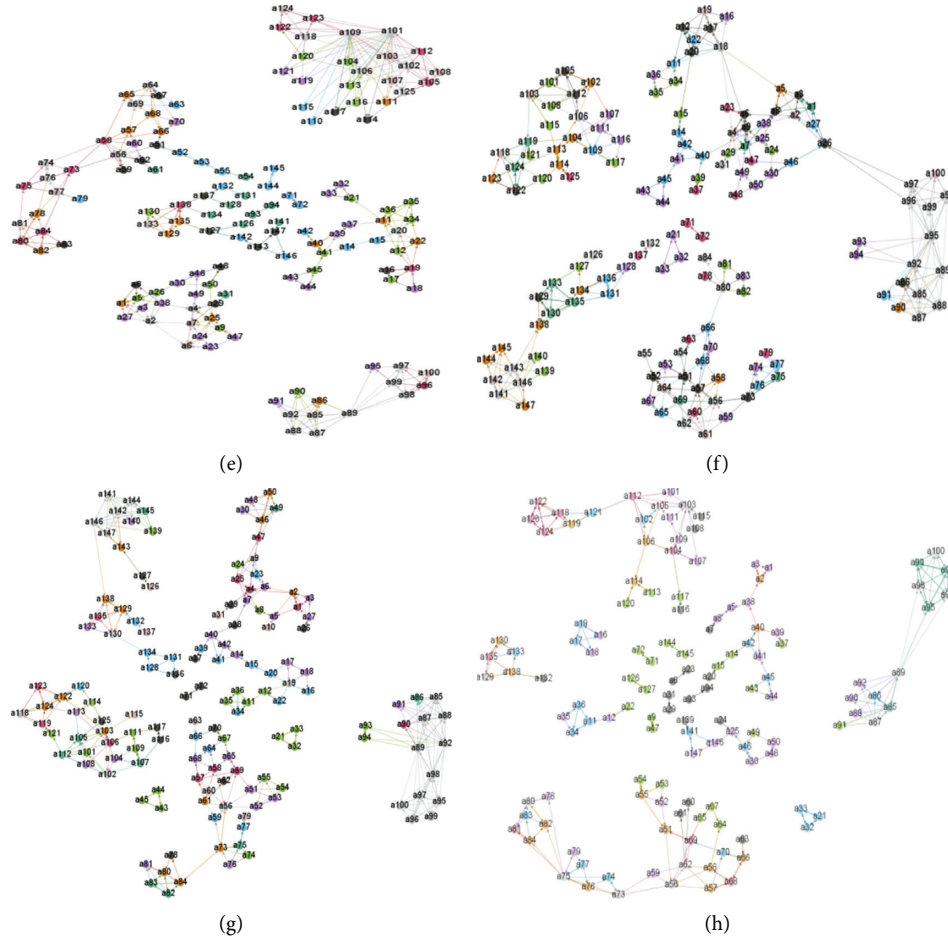


FIGURE 2: Topology of the online social network. (a) Practical online support network. (b) Emotional online support network. (c) Social online support network. (d) Employment online discussion network. (e) Achievement online support network. (f) Revenue online discussion network. (g) Professional online discussion network. (h) Love and marriage online discussion network.

Network density is an important metric in the structural form of a network and is defined as the ratio of the number of edges actually present in the network to the maximum number of edges that can be accommodated. Network density is commonly used in online social networks to measure the density of social relationships, with higher density indicating more connections and the greater effectiveness of the network's influence on nodes. The network distance is used to describe the length of the shortest path between nodes, and the clustering coefficient reflects the degree of network aggregation. The combination of network distance and clustering coefficient can demonstrate the "small world" effect, showing how nodes are embedded in their surrounding nodes. By referring to relevant research results [37–39], an indicator is proposed here to combine the two, as shown in (2), and the last column of  $Cd$  is obtained as in Table 2; the higher the  $Cd$ , the more obvious the "small world" effect.

$$Cd_i = \frac{(ci/\sum_i ci + di/\sum_i di)}{2}. \quad (2)$$

Statistics on the whole network viscosity characteristics of the athletes' online social networks are shown in Table 2. The network density of each of the athletes' online social networks was in the interval [0.016, 0.028], indicating that the network had essentially equal influence on the athletes. Among the online social support networks, emotional online support was the most lacking, and their network density was only 0.016; among the online social discussion networks, the athletes were more concerned with income issues. Comparing the average distance of the networks revealed that athletes had the longest chain of online practical support relationships and were prone to extend online practical support relationships. The average clustering coefficient for love and marriage online discussions was only 0.3, with the

TABLE 2: The viscosity characteristics of online social network.

Name of network	Density	Mean distance	Average clustering coefficient
Practical online support network	0.028	5.663	0.401
Emotional online support network	0.017	4.025	0.310
Social online support network	0.024	4.672	0.361
Employment online discussion network	0.016	3.138	0.314
Achievement online support network	0.02	2.781	0.379
Revenue online discussion network	0.021	3.059	0.388
Professional online discussion network	0.019	2.911	0.388
Love and marriage online discussion network	0.016	2.967	0.3

sparsest network relationships. The “small world” effect of practical online support network is most evident in the  $Cd$  measure.

**4.3. Centrality.** Centrality measures the structural properties of social networks at a microlevel and is usually measured by three indicators: degree centrality, closeness centrality, and betweenness centrality.

As shown in (3),  $g$  represents the size of the network and degree centrality  $C_D$  is the most direct indicator of the centrality of a network node; it measures the total number of direct links between a node and other nodes. The higher the degree centrality is, the more important the node is in the network; in the case of directed graphs, the two parameters out-degree and in-degree need to be considered.

$$\begin{cases} C_D(N_i) = \sum_{j=1}^g x_{ij}, \\ C'_D(N_i) = \frac{C_D(N_i)}{g-1}, \quad i \neq j. \end{cases} \quad (3)$$

Closeness centrality  $C(x)$  examines the extent to which a node propagates information without relying on other nodes; if a node is at a short distance from all other nodes, then that point is at the overall centre. As shown in (4),  $d(x, y)$  indicates the distance between node  $x$  and node  $y$ :

$$C(x) = \frac{1}{\sum_y d(y, x)}. \quad (4)$$

As shown in (4), betweenness centrality characterises the extent to which a node is known as a bridge to the various others or the extent to which it can control others.  $\sigma_{st}(v)$  denotes the number of shortest paths of two nodes through point  $v$ , and  $\sigma_{st}$  denotes the total number of shortest paths of two nodes.

$$CB(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}. \quad (5)$$

The data was processed using the formulae above. The characteristics of the whole online social network structure of the athletes in-degree and out-degree are shown in Table 3. The maximum values show that the out-degree is

generally greater than the in-degree, indicating that athletes actively seek online support and online discussion more than they passively receive it. The minimum value of in-degree and out-degree for all networks is 0, indicating that there are athletes who do not engage in online support and online discussion behaviours with others in their sport team. In addition, there is more variation between individual athletes, such as the practical online support network and achievement online support networks, which have a maximum value of 25, while both have a minimum value of 0.

Table 4 shows the parameter values for the whole online social network centrality characteristics of the athletes. Comparing the indicators of eccentricity, closeness centrality, harmonic centrality, and betweenness centrality reveals that athletes have a preference for seeking achievement support over other online support. Online social discussion and online social support in general show a more active and active search for support and discussion, while acceptance is more passive. In terms of the propagation of online social support and online social discussion relationships, athletes were more likely to extend their daily help support relationships and were susceptible to the influence of intermediaries. Overall, the higher closeness centrality of love and marriage online discussion network indicates a higher chance of connection between nodes. Harmonic centrality is usually for nonconnected networks and can be ignored here. Network with higher betweenness centrality indicates that more nodes in the network play a role in connecting these groups together, such as practical online support network.

**4.4. Bipartite Relationship, Tripartite Relationship.** Bipartite relationship is the basic unit of statistical analysis of social networks. For directed networks, the bipartite relationship between two nodes in the network includes reciprocal relationships, one-way relationships, and nulls.

Tripartite relationships, i.e., relationships between three nodes, are formed by combining three pairs of bipartite relationships. The tripartite structure is the basis of social structure and is described by scholars such as Holland as 16 isomorphic categories; each represented by three or four symbols: the first number represents the number of reciprocal tripartite relations; the second number represents the number of asymmetrical tripartite relations; the third number represents the number of nulls in tripartite

TABLE 3: Characteristics of the whole online social network structure of athletes.

Name	Out-degree				In-degree			
	Mean	Std.	Max	Min	Mean	Std.	Max	Min
Practical online support network	4.068	3.362	25.000	0.000	4.068	1.925	10.000	0.000
Emotional online support network	2.433	1.880	10.000	0.000	2.433	1.537	6.000	0.000
Social online support network	3.418	3.055	14.000	0.000	3.411	1.863	8.000	0.000
Employment online discussion network	2.092	2.238	15.000	0.000	2.092	1.444	7.000	0.000
Achievement online support network	2.837	3.384	25.000	0.000	2.837	1.869	9.000	0.000
Revenue online discussion network	2.936	2.550	15.000	0.000	2.957	1.808	8.000	0.000
Professional online discussion network	2.716	2.358	14.000	0.000	2.752	1.657	7.000	0.000
Love and marriage online discussion network	2.015	1.728	9.000	0.000	2.015	1.468	8.000	0.000

TABLE 4: Central characteristics of the whole online social network structure of athletes.

Name	Eccentricity	Closeness centrality	Harmonic centrality	Betweenness centrality
Practical online support network	8.247	0.294	0.368	251.397
Emotional online support network	4.376	0.435	0.502	57.780
Social online support network	5.266	0.366	0.440	129.133
Employment online discussion network	3.092	0.436	0.491	21.084
Achievement online support network	2.823	0.497	0.558	17.440
Revenue online discussion network	3.832	0.450	0.528	28.336
Professional online discussion network	3.694	0.453	0.531	20.667
Love and marriage online discussion network	2.718	0.550	0.607	15.389

relations; and if there is a final letter, it is used to distinguish between two similar tripartite relations ( $T$  for transmission,  $C$  for circular,  $U$  for up, and  $D$  for down).

The overall number of relationships between the 147 nodes in this study totalled 3,668, and Tables 5 and 6 show the proportion of each relationship to the overall relationship in the bipartite and tripartite relationships, respectively. It can be seen that the athlete group has an absolute majority proportion of the nulls in all 8 networks. The proportion of one-way relationships is very small, while the proportion of reciprocal relationships is even smaller. This suggests that online one-way communication between groups of athletes is rare and two-way interaction is more difficult. Examining the actual situation, based on the training patterns of the athletes, this is not unrelated to the regulation of communication tools. A comparison of the bipartite relationship data reveals that athletes interact with each other mainly through social online support network to enhance communication with each other, but less in terms of marriage and emotions, which may be that athletes are close to each other before they engage in emotional confessions and discussions on topics such as marriage.

The very large percentage of 003 in the athletes' tripartite relationships suggests that in most cases there are no online support and online discussion relationships between two out of three people in the athletes' group. 012 had the second highest proportion, indicating that a larger proportion of the online relationships among the three athletes, of which only two had relationships with each other, were difficult to transmit among the three and that it was difficult to communicate and collide their perceptions effectively. 300 and 030C had the smallest proportion of both, indicating that a

small proportion of the three athletes had one-way transmission relationships and two-way transmission relationships; i.e., the whole online social network of athletes has a poor sense of identity and belonging in the group. Bipartite relationships are the simplest relationships to consider and tripartite relationships have complex social relationship attributes and so are generally of greater concern.

**4.5. Community Structure.** Realistic social networks are often not evenly distributed, but consist of many subnetworks with high similarity of nodes within the subnetwork and low similarity of nodes outside the subnetwork, resulting in a community structure. The modularity  $Q$  is used to quantify or judge the merit of community division in a network, as shown in (6). It is defined as the ratio of the total number of edges within the community to the total number of edges in the network minus an expectation value that is the magnitude of the ratio of the total number of edges within the community to the total number of edges in the network that would result from the same community assignment if the network was set as a random network. If  $Q$  is greater than 0.3, then the network has a significant community structure [40].

$$Q = \sum_i (e_{ii} - a_i^2) = \sum_i e_{ii} - \sum_i a_i^2 = \text{Tre} - \|e^2\|. \quad (6)$$

Figure 3 shows the modularity of the online social support network and the online social discussion network. It can be seen that the modularity  $Q$  of both is even more than 0.8. It indicates that community structure exists widely in

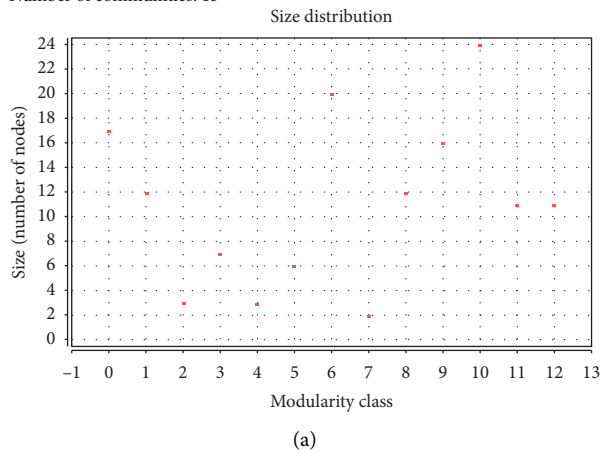
TABLE 5: Distribution of bipartite relationship.

Name	Reciprocal relationships (%)	One-way relationship (%)	Nothingness (%)
Practical online support network	3.05	6.69	90.26
Emotional online support network	0.49	2.49	97.02
Social online support network	1.42	3.93	94.65
Employment online discussion network	0.24	3.61	96.14
Achievement online support network	0.59	4.08	95.34
Revenue online discussion network	0.22	1.39	98.39
Professional online discussion network	0.51	4.06	95.43
Love and marriage online discussion network	0.33	3.31	96.36

TABLE 6: Distribution of tripartite relationship.

Type	Practical online support network	Emotional online support network	Social online support network	Employment online discussion network	Achievement online support network	Revenue online discussion network	Professional online discussion network	Love and marriage online discussion network
003	73.5	91.32	84.74	89.04	86.79	95.24	87.84	92.62
012	16.47	7.05	10.66	9.72	10.94	4.06	9.67	6.02
102	7.72	1.39	3.89	0.67	1.58	0.63	1.87	1.12
021D	0.41	0.05	0.14	0.20	0.23	0.02	0.17	0.08
021U	0.27	0.06	0.10	0.17	0.11	0.01	0.19	0.06
021C	0.36	0.07	0.14	0.10	0.15	0.02	0.10	0.05
111D	0.32	0.02	0.13	0.02	0.03	0.01	0.03	0.04
111U	0.40	0.01	0.09	0.03	0.08	0.01	0.05	0.00
030T	0.10	0.01	0.03	0.04	0.04	0.00	0.03	0.00
030C	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01
201	0.12	0.01	0.03	0.00	0.01	0.00	0.02	0.00
120D	0.08	0.00	0.02	0.00	0.01	0.00	0.01	0.00
120U	0.06	0.00	0.01	0.00	0.02	0.00	0.01	0.00
120C	0.05	0.00	0.01	0.00	0.00	0.00	0.00	0.00
210	0.10	0.00	0.01	0.00	0.01	0.00	0.01	0.00
300	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00

Modularity: 0.854  
 Modularity with resolution: 0.854  
 Number of communities: 13



Modularity: 0.867  
 Modularity with resolution: 0.867  
 Number of communities: 19

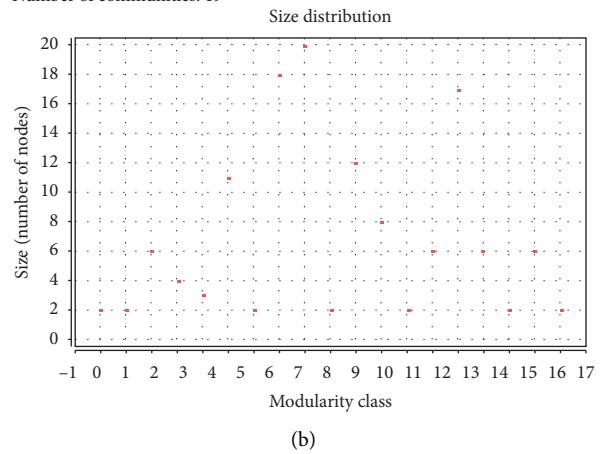


FIGURE 3: Modularization of community structure. (a) Modularity for online social support network. (b) Modularity for online social discussion network.

athletes' online social networks. A relatively stable community structure is very important for the physical and mental development of athletes.

## 5. Conclusion

In summary, the network density of the athletes' online social support network is greater than that of the online social discussion network, and specifically among the sub-networks, the practical online support network is the most dense and the love and marriage online discussion network is the least dense. This indicates that in recent years professional athletes have been gradually increasing their online practical support claims in the course of their daily lives and training. Athletes also show higher social online support, indicating that, in the new era and environment, new athletes still attach more importance to social interaction, and they are actively expanding their life circle.

The proportion of the relational composition of athletes' online social support networks has also changed in a more interesting way. While the proportion of teammates and family members remains high overall, some studies have shown that the proportion of social support provided by family members has declined significantly, while the social support provided by classmates or friends even surpasses that of family members [27], a feature that is particularly evident when it comes to social online support. Of course, it is still important not to overlook the strong social support power of teammates, which is unmatched by any other social role.

In terms of strong and weak relationships, athletes' online social support networks still show a strong tendency towards "strong relationships," but compared to online social discussion networks, weak relationships are slowly gaining ground, especially in social online support, where the proportion of weak relationships reaches almost 40%. This is mainly due to the fact that classmates or friends play a greater role in this area, while team managers have the lowest percentage, which is also related to the authority of the managers and the fact that athletes are more distant from coaches and instructors and have to turn to their teammates or classmates for social support.

## Data Availability

The original data used in this study are the questionnaire data obtained from the survey. The original 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 work was supported in part by the National Natural Science Foundation of China under Grant nos. 61440036 and 61040029 and the National Social Science Foundation of China under Grant no. 16ATY002.

## References

- [1] J. Ding and H. Qian, "A study on the multi-level relationship between the overall social network of Chinese professional athletes," in *Proceedings of the 2015 Abstract of papers of the Tenth National Sports Science Congress*, pp. 1217-1218, China Sports Science Society, Hangzhou, Jiangsu, January 2015.
- [2] H. Qian and X. Zhang, "Effects of personal characteristics of Chinese professional athletes on the structure of social support network," *Journal of Shanghai Institute of Physical Education*, vol. 38, no. 6, pp. 59-63, 2014.
- [3] H. Qian, X. Yang, J. Ding et al., "Analysis of social network characteristics of Chinese professional athletes," *Journal of Wuhan Institute of Physical Education*, vol. 50, no. 7, pp. 77-83, 2016.
- [4] H. Qian and J. Ding, "Analysis of the overall social network characteristics of Chinese professional athletes," in *Proceedings of the Compendium of Abstracts of the 10th National Sports Science Conference 2015*, pp. 811-813, Chinese Society of Sports Science, Hangzhou, Jiangsu, September 2015.
- [5] X. Zhang and H. Qian, "Research on structural characteristics of job search network of Chinese professional athletes—based on sample survey of professional athletes in Shaanxi province," *Journal of Wuhan Institute of Physical Education*, vol. 52, no. 3, pp. 17-23, 2018.
- [6] Z. Xiong, *A Study on Community Discovery Technology and its Application in Online Social Network*, Central South University, Changsha, China, 2012.
- [7] C. Xiao, *Analysis and Prediction of User's Behavior in Online Social Network*, University of Electronic Science and Technology of China, Chengdu, China, 2013.
- [8] Z. Li, *A Study on the Measurement and Law of Information Dissemination of Node Information in Social Network*, Harbin Institute of Technology, Harbin, China, 2015.
- [9] Y. Su, *A Study on the Communication Model and Intervention of Public Opinion in Online Social Network*, Yanshan University, Qinhuangdao, China, 2018.
- [10] L. Wang and X. Cheng, "Dynamic community discovery and evolution of online social network," *Journal of Computer Science*, vol. 38, no. 2, pp. 219-237, 2014.
- [11] Y. Chen, Z. Li, X. Liang et al., "A review of online social network rumor detection," *Chinese Journal of Computers*, vol. 41, no. 7, pp. 1648-1677, 2017.
- [12] Li Qian, B. Liu, and Y. Zhao, "Construction of statistical metadata in the big data environment," *Statistics and Information Forum*, vol. 35, no. 03, pp. 14-20, 2020.
- [13] R. Zhang, L.-X. Liu, X.-B. Tang, and B.-R. Zhang, "Research on the prediction of commodity retail price index based on web search data in the context of big data," *Statistics and Information Forum*, vol. 35, no. 11, pp. 49-56, 2020.
- [14] Z. Pang and M. Zeng, "Has internet use affected youth subjective well-being? -an analysis of CGSS data from 2010-2015," *Journal of Xi'an University of Finance and Economics*, vol. 33, no. 03, pp. 71-77, 2020.
- [15] H. Li and Y. Chen, "Research on the generation and dissemination mechanism of network public opinion—based on the perspective of big data social network analysis," *Age of Media Contemporary Communications*, vol. 186, no. 1, pp. 24-25, 2016.
- [16] X. Wang, "Data science and social networks: big data, small world," *Science and Society*, vol. 1, no. 1, pp. 27-35, 2014.
- [17] M. Jia, H. Xu, J. Wang et al., *Handling Big Data of Online Social Networks on a Small Machine*, Springer International Publishing, New York, NY, USA, 2015.

- [18] Y. Tian, Y. Lin, and L. Gong, "Research on the influence of Internet culture on campus culture in the Internet era--analysis based on questionnaire survey of teachers and students in some universities in Shaanxi," *Statistics and Information Forum*, vol. 35, no. 07, pp. 122–128, 2020.
- [19] J. Ding, L. Xu, H. Qian et al., "A study on the multi-level relationship of the overall social network of Chinese professional athletes," *Journal of Xi'an Institute of Physical Education Nol*, vol. 4, pp. 405–409, 2016.
- [20] J. Ding and H. Qian, "A central analysis of the overall social network of professional athletes in China," *A Journal of Xi'an University of Technology*, vol. 36, no. 5, pp. 408–413, 2016.
- [21] Y. Qian, *Online Social Networking Analysis*, Electronic Industry Press, Beijing, China, 2014.
- [22] Y. Zhao and J. Luo, "How to measure social capital: a review of empirical research," *Social Sciences Abroad*, no. 2, pp. 18–24, 2005.
- [23] X. Jin, Y. Ren, and H. du, *Change of Social Network and Conceptual Behavior of Migrant Workers*, Social Sciences Press, Beijing, China, 2014.
- [24] J. S. A. W. Carolyn, "AP primer: logit models for social networks," *Social Networks*, vol. 21, no. 1, pp. 37–66, 1999.
- [25] Z. Li and J. Luo, "A view on local management theory from the perspective of social network," *Journal of Management*, vol. 8, no. 12, pp. 1737–1747, 2011.
- [26] Z. Li and J. Luo, "The social behavior and relational network characteristics of Chinese—viewpoint of a social network," *Social Science Front*, vol. 199, no. 1, pp. 159–164, 2012.
- [27] J. Luo and F. Zeng, "A study on organization based on complex system perspective," *Foreign Economics and Management*, vol. 41, no. 12, pp. 112–134, 2019.
- [28] J. Liu, *Practical Guide to Integrated Network Analysis UCINET Software*, GE Zhi Publishing House, Shanghai, China, 2014.
- [29] K. Xu, S. Zhang, H. Chen et al., "Measurement and analysis of online social network," *Chinese Journal of Computer*, vol. 37, no. 1, pp. 165–188, 2014.
- [30] Y. Ren, S. Li, H. Du et al., "An analysis of the social network structure of migrant workers," *A Journal of Xi'an Jiaotong University (Social Sciences Edition)*, vol. 91, no. 5, pp. 44–51, 2008.
- [31] Q. Huang, S. F. Fung, B. Liu et al., "Modeling for professional athletes' social networks based on statistical machine learning," *Advanced Data Mining Methods for Social Computing*, vol. 8, 2019.
- [32] J. Liu, "A study on social network model," *Sociological Studies*, vol. 1, pp. 1–12, 2004.
- [33] J. Liu, *A Study on Social Network Extraction Technology for Text-Oriented Information and its Application*, University of Defence Science and Technology, Changsha, China, 2009.
- [34] Y. Wang, Field, T. Li et al., "Visual analysis of shipping recruitment information based on Gephi," *Big Data*, vol. 4, no. 3, pp. 81–91, 2018.
- [35] S. Hussain, L. Muhammad, and Y. Atomsa, "Mining social media and DBpedia data using Gephi and R," *Journal of Applied Computer Science & Mathematics*, vol. 12, no. 25, pp. 14–20, 2018.
- [36] Y. Guan, Y. Xiang, and C. Kang, "Research and application of visual analysis method based on Gephi," *Telecommunications Science*, vol. 29, no. S1, pp. 112–119, 2013.
- [37] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 3804–3816, 2020.
- [38] Y. Hu, F. Xiong, S. Pan, X. Xiong, L. Wang, and H. Chen, "Bayesian personalized ranking based on multiple-layer neighborhoods," *Information Sciences*, vol. 542, pp. 156–176, 2021.
- [39] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4186–4199, 2020.
- [40] S. L. Deng, B. Wang, B. Wu et al., "Modeling and validation of complex network association partitioning based on information entropy," *Computer Research and Development*, vol. 49, no. 4, pp. 725–734, 2012.

## Review Article

# Opinion Expression Dynamics in Social Media Chat Groups: An Integrated Quasi-Experimental and Agent-Based Model Approach

Siyuan Ma <sup>1</sup> and Hongzhong Zhang <sup>2</sup>

<sup>1</sup>Department of Communication, Michigan State University, East Lansing, MI, USA

<sup>2</sup>School of Journalism & Communication, Beijing Normal University, Beijing, China

Correspondence should be addressed to Hongzhong Zhang; zhanghz9@126.com

Received 25 July 2020; Revised 17 September 2020; Accepted 28 December 2020; Published 9 January 2021

Academic Editor: Hongshu Chen

Copyright © 2021 Siyuan Ma and Hongzhong Zhang. 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.

Social media chat groups, such as WeChat and WhatsApp groups, are widely applied in online communication. This research has conducted two studies to examine the individual level and collective level's opinion dynamics in those groups. The opinion dynamic is driven by two variables, people's perceived peer support and willingness of opinion expression. The perceived peer support influences the willingness of opinion expression, and the willingness influences the dynamics of real opinion-expression. First, the quasi-experimental study recruited twenty-five participants as the observation group and found that decreasing perceived peer support would significantly increase individuals' expression willingness to protect his/her opinion. To generalize the individual level findings to a collective level, the second study treated the social media chat groups as an undirected fully-connected social network and simulated people's opinion expression dynamics with an agent-based model. The simulation indicated that (1) with the help of increased willingness of opinion expression, the minority opinion supporters as a collective did not fall silent but continue to express themselves and (2) increasing willingness of opinion expression would maintain the existence of minority opinion but could not help the minority reverse to the majority.

## 1. Introduction

Opinion discussions on social media have become increasingly diversified when accompanied by the progress of media technology, diversification of personal interests, and requirements of gratification [1]. This diversified feature boosts the formation of *social media chat groups* (short for social media groups), such as WhatsApp groups, WeChat groups, and Interest groups in online forums [2]. Each social media group has regular members and focuses on some specific topics [3]. By analyzing the number of active users monthly, social media groups have become increasingly popular and hold a broader influence on people's daily life [4]. Unlike open social media platforms that produce outbursts of information, social media groups provide an environment that allows a more in-depth discussion of opinions, which has a long-term impact on their users [5].

Therefore, the dynamic of opinion expression in social media groups is an essential topic and worth further study.

An ideal media environment should not hinder the expression of opinions, especially minority opinions. According to Habermas' public sphere theory [6], preserving space for the exchange of opinions helps maintain political legitimacy, social stability, and protect the welfare of citizens. Unfortunately, it is hard for the mass media to provide an ideal environment for opinions. For instance, opinions in mass media tended to grow more extreme [7, 8], and the expression of minority opinions tended to fall silent [9–11]. However, some recent studies proposed different findings that social media groups could provide a space for the expression of opinion, especially minority opinions (e.g., [12, 13]). Thus, researchers design two studies with the following questions:

- (1) Quasi-experiment: is individuals' willingness of opinion expression protected in social media groups?
- (2) Agent-based model simulation: how do individuals' willingness of opinion expression trigger the collective behavior, like the whole group's dynamics of opinion expression?

## 2. Related Work

*2.1. Conflict Presumptions of Individual Opinion Expression Willingness: Spiral of Silence versus Corrective Action.* Willingness of opinion expression refers to individuals' eagerness to share or post their opinions in public [14, 15]. This term is widely used as a dependent variable in communication and political science studies. Previous studies have long been concerned that perceived peer support affects individuals' willingness of opinion expression. However, their primary two presumptions of that "effect" are in conflict with each other.

First, the spiral of silence theory suggests that the decrease of perceived peer support could lessen individuals' willingness of opinion expression, or even silence it because of people's fear of isolation [9, 16, 17]. The spiral of silence theory proposes that because individuals are afraid of being isolated, thus more perceived support will increase people's willingness of opinion expression, while less support will lessen the willingness [18, 19]. This theory was proposed in the traditional mass media environment. Thus, the perceived support referred to support from mass media. Researchers later extended this support to peer relations and social media. Research by Oshagan [20] introduces the concept of the *reference group*, which is used as a synonym of *peer*, to describe family members, intimate partners, colleagues, and neighbors. Studies use this concept for reference and propose that reducing perceived peer support is even more effective than reducing mass media support when it tries to lessen individuals' willingness of opinion expression [16, 17, 21].

On the contrary, the corrective action hypothesis proposes that controversial issues will inspire people to speak out their thoughts to change others' perceptions of public opinion [22]. When people perceive news or comments on social media going biased against them, they are concerned that the public may be swayed in that direction. Because of this fear, people are motivated to take action to correct public opinion by speaking out even when their opinion is in the minority [23]. Similar findings are found both in the news report and social media discussion situation; people who have received hostile news or received negative social media comments on their opinion would like to speak out to protect their opinion [24, 25].

Evidence suggests that whether audiences are motivated to keep silent or speak out depends on how they perceive others' opinions [26]. To deduct from the theory, when people feel the fear of isolation, the decrease of perceived peer support will lessen the willingness of opinion expression [27]. However, if people do not feel such fear, the perceived peer support will not diminish or even increase the willingness of opinion expression. Nowadays, people have

more peers on social media (although they may never meet offline) comparing with the traditional mass media age; and they are more likely to find online peers who share the same opinions [12, 28, 29]. Because people are able to get psychological support conveniently from their online peers, they will not easily feel the fear of isolation [13]. Therefore, even though individuals perceive, they are the minority rather than the majority, their willingness of opinion expression may not decline but even increase. For instance, some studies found people who held the minority opinion did not have much change in their willingness of opinion expression while comparing with majority opinion [15, 30]. Besides, some studies proposed that minority perception increased individuals' willingness of opinion expression [23–25, 29].

To sum up the findings above, there are conflicting conclusions between perceived peer support and willingness of opinion expression. Traditional mass media studies find that a decrease of support will lessen individuals' willingness of opinion expression [9, 16, 17, 20]. However, recent social media studies have found that the perceived peer support will not decrease or even encourage this willingness (e.g., [29, 30]). In the current study, social media groups created an unbiased fully-connected social network to help members get in touch with each other more conveniently, which suggests that people may not feel much fear of isolation. Therefore, this research proposes the first research question as follows:

RQ1. Will the decrease of perceived peer support enhance individuals' willingness of opinion expression in social media groups?

The current research has conducted a quasi-experiment (study 1) to answer this question. Details are presented later in this article.

Moreover, some other variables also influence the willingness of opinion expression. First, the spiral of silence theory [9] proposes *hardcore* to describe the people who are willing to express their opinions consistently regardless of social pressure. Studies have shown that if the hardcore is removed, minority opinion supporters' willingness to express will significantly decrease [31–33]. Furthermore, fear of authority, lack of self-efficiency, feeling the importance of the issue, accessibility of different opinions, and technical conditions of communication could all influence personal opinion expression [13, 33]. In order to keep these variables consistent, the quasi-experiment applies repeated measurements on each individual.

*2.2. Simulating the Collective Opinion Expression Dynamics by the Agent-Based Model.* According to Sawyer's [34] mechanism of emergence, collective phenomena are collaboratively created by individuals yet may not reducible to individual action. Therefore, to generalize individual-level findings to the collective behavior of opinion expression in a social media chat group, we need an agent-based model (ABM) to simulate the dynamics. The ABM consists of agents, system space, and external environment. Agents have



various attributes and are heterogeneous. ABM's rules act on the agents rather than the process of action [35, 36]. Each agent is autonomous and decides his/her behavior by learning memories, interacting with the neighbors, or the external environment with the rules of behavior [37]. To apply ABM in the current research, it will take into account people's initial opinion distribution in the social media group, the rules of opinion expression, and the evolution process. After setting up rules, new phenomena will emerge from ABM's simulation. Previous ABMs of opinion dynamics were originated from the Ising model. The two spin states of the Ising model corresponded to any relative set of concepts in social science, such as the concepts of right or wrong, and speaking out or silence. In principle, this model describes all multibody systems with two possible states [38]. Several representative evolutions of the Ising model are applied to the research of public opinion, such as the Voter model, Sznajd model, Majority-rule model, Krause–Hegselmann model, and Deffuant model [39, 40].

Although there were conflicting presumptions of the opinion expression willingness from individual-level studies, previous simulation studies mainly bought the idea of the spiral of silence, which proposed that less peer support would reduce the willingness and then also reduce or even silence the opinion expression. For instance, the study assumed people who received external opinion pressure larger than peer support would be less likely to speak out to protect their opinion. As time progressed, the minority opinion would fall into silence in any kind of social network, especially when it was the Barabasi–Albert scale-free network [41]. Social bot studies also proposed that the spiral of silence theory would lead agents to consensus opinion in a social network, with some agents being silenced. However, applying a small number of bots would easily shape the norms adopted by social media users and tip over the opinion climate [42].

The current research has found that those proposed models can be improved from two perspectives. First, we should analyze changes in opinion expression instead of changes of opinions. Some studies did not pay much attention to the difference between opinion expression and the opinion itself. From their assumption, every change of an agent would be assumed as the change of opinion itself, which referred an individual would change his or her opinion after receiving some influence from an opposite position (e.g., [35, 42, 43]). However, this assumption is not entirely consistent with the evidence. According to the spiral of silence theory and echo-chamber effect, the change of opinion does not happen in such a straightforward way. On the contrary, the opinion expression is relatively easier to change [44, 45]. Besides, a silenced position does not necessarily mean the person has changed his/her idea. Secondly, we should use behavioral data to support the design of ABM, while previous studies' simulation did not have much support from real behavioral data. For instance, Sohn and Geidner [46] did distinguish the opinion expression and the opinion itself, interpreting the opinion dynamics more accurately. However, lacking real data intended that their model needed more parameters to adjust and thus reduced

the validity [47]. Besides, previous models did not pay much attention to the possibility that the spiral of silence may not exist in the social media environment. Thus, behavioral data can help revise the rules of ABM.

In order to improve the two perspectives of model simulation mentioned in the previous paragraph, this study has planned to develop an agent-based model in study 2 that is consistent with quasi-experimental findings. Thus, the current research proposes as follows:

RQ2. How do individuals' change of opinion expression willingness influence the whole group's dynamics of real opinion-expression?

In order to clarify the relationship between study 1 and study 2, we provide a workflow. See Figure 1 for more details.

### 3. Study 1: Quasi-Experiment in Social Media Group

**3.1. Methods.** This study has designed a three-condition quasi-experiment to explore the relationship between perceived peer support and willingness of opinion expression from an individual level.

**3.1.1. Participants.** The quasi-experiment recruited 25 students from a North China university to form the observation group. Those 25 participants all had the same opinions towards three similar questions. The quasi-experiment consisted of three conditions that the observation group had (1) more members, (2) evenly-matched members, or (3) fewer members when compared with the opposite opinion group. When the observation group members were more than the opposite opinion group, we named the observation group as a *majority opinion condition*; when the members were balanced, we named it as a *balanced opinion condition*; when the members were fewer than the opposite opinion group, we named it as a *minority opinion condition*. For each condition, we asked those people to discuss one of the three questions. The opposite opinion group contained different numbers of people in different conditions, and the current study recruited volunteers to form the opposite opinion group. Besides, those volunteering opposite group members were not the object of analysis.

**3.1.2. Stimuli.** The stimuli in this quasi-experiment were the proportion of the observation group in the total participants. When the proportion was high (which referred to majority opinion condition), we observed high perceived peer support; when the proportion was balanced (balanced opinion condition), we observed medium perceived peer support; when the proportion was low (minority opinion condition), we observed low perceived peer support. The manipulation check results are shown in the result part of study 1.

The quasi-experimental conditions and participants' arrangements are shown in Table 1.

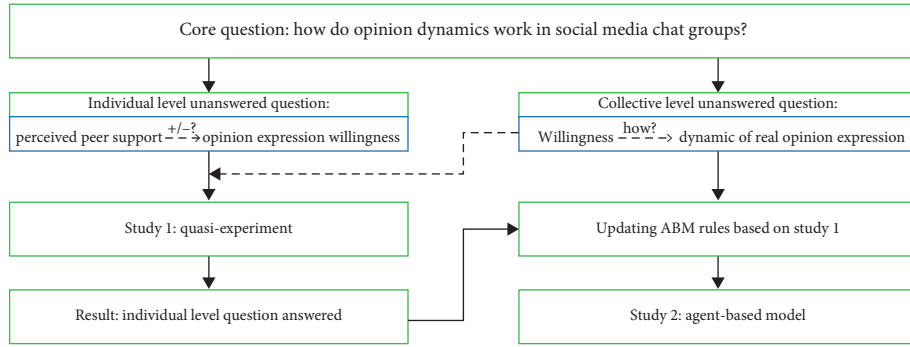


FIGURE 1: The workflow between study 1 and study 2.

TABLE 1: Quasi-experimental participants arrangement.

Condition	Observation group members	Opposite opinion members	Expected level of perceived peer support
Majority opinion (more observation members)	64% ( $N = 25$ )	36% ( $N = 14$ )	High
Balanced opinion (evenly-matched observation members)	53% ( $N = 25$ )	47% ( $N = 22$ )	Medium
Minority opinion (fewer observation members)	39% ( $N = 25$ )	61% ( $N = 40$ )	Low

Note. The majority opinion situation refers that observation group members are more than the opposite opinion group; the balanced opinion situation refers that observation group members are evenly-matched with the opposite opinion group; the minority opinion situation refers that observation group members are fewer than the opposite opinion group.

**3.1.3. Procedure.** First, the researchers designed a survey with 15 questions, and each of them had two opposite opinions. All the questions were about the campus life in the college to make sure the participants had enough experience and background knowledge to discuss them. A total of 127 students from a North China University participated in the survey.

Second, 25 students who had the same opinion on three of those questions were recruited to form the observation group. The researcher announced that all participants who showed up in the experiment would be awarded the same research credits in the course, the credits were only related to whether they showed up in the experiment or not, not related to how active they behaved. Thus, even though the researcher could know how many comments each person had made in the discussion, the participants would not feel too much pressure of expressing themselves. Not being active does not impair their profits in this course.

Third, the researchers recruited volunteers who participated in the survey. They were assigned opinions that were opposed to the observation group. The researchers also provided credible materials so that the opposite opinion group members had enough knowledge to discuss with the observation group.

Fourth, the quasi-experiment manipulated the level of perceived peer support by changing the different conditions mentioned as follows: majority opinion, balanced opinion, and minority opinion condition. In each condition, we organized a group discussion which lasted approximately 30 minutes. Discussion topics were the three questions mentioned in the second step.

Finally, researchers controlled experimental conditions. In order to simulate the anonymous social media

circumstance, each participant had a pseudonym. The pseudonym changed after each discussion. In addition, offline communication was not allowed. Therefore, an individual of the observation group only knew that some people supported his or her opinion, but could not match the people in the social media group to the offline. This design also helped to reduce the pressure of “you must speak out” from acquaintance friends.

Because of participants’ personal reasons, three observation group members did not participate in the first condition, six members missed the second condition, and five members missed the third condition. The current study used an average interpolation method to handle the missing values.

**3.1.4. Measurement.** Perceived peer support and willingness of opinion expression all need operational definitions. In this quasi-experimental design, the two concepts are measured as follows:

- (1) *Perceived Peer Support.* In a social media group, an individual’s perceived peer support includes all comments from the same-opinion people. In traditional mass media, perceived peer support was hard to quantify. Researchers relied on self-reporting, asking how much the participant thought that other commenters could support his or her opinion with a Likert scale [14]. In this quasi-experiment, participants are not able to report their perceptions while participating in the discussion. However, all comments in the social media group are recorded so that we can use the real “received peer support” to represent the perceived support. Some evidence in

previous studies supported this operationalization. For instance, youngsters who got less support from their partners would report less perceived peer support and exhibited more psychosocial adjustment problems [48]; and another psychological study about new mothers indicated that receiving more

peer support would increase the perceptions of being trusted, accepted in the relationship [49]. Those studies support that received peer support can be a good predictor of perceived support. Therefore, perceived peer support is measured as a ratio as follows:

$$\text{perceived peer support} = \frac{\text{the number of comments that an individual received from same opinion people}}{\text{the number of all comments}} \quad (1)$$

Note: “the number of all comments” should exclude comments made by this individual him/herself.

- (2) *Willingness of Opinion Expression*. Similar to the perceived peer support, all comments in the group discussion are recorded. Therefore, the willingness of opinion expression is measured as the number of comments that an individual has posted during the discussion.

The current study created a workflow of doing study 1. See Figure 2 for more details.

3.2. *Results*. The current study used a quasi-experiment to measure two variables and their relationships: perceived peer support and willingness of opinion expression. The research used individual-level repeated measurement.

3.2.1. *Manipulation Check*. The current study applied repeated measurement *F*-test to examine whether changing the situation of the majority opinion to a balanced opinion and a minority opinion would manipulate the perceived peer support or not.

The result in Table 2 depicted that all different types of *F*-test concluded that the variation of the majority opinion, balanced opinion, and minority opinion changed perceived peer support significantly.

The current study made pairwise comparisons to examine whether the direction of manipulation was correct or not. During the comparisons, 1 represented the majority opinion condition; 2 represented the balanced opinion condition; 3 represented the minority opinion condition.

This pairwise comparison in Table 3 concluded that, during the process from the majority condition to a balanced condition, the perceived peer support decreased nonsignificantly; during the process of the balanced to minority condition and from majority to minority, the perceived peer support decreased significantly. In general, it meant the perceived peer support was successfully decreased by changing an opinion group from the majority condition to a balanced condition and then to a minority condition. Therefore, the manipulation was acceptable.

3.2.2. *Correlation Analysis*. Repeated measurement correlation analysis (Rmcorr) is a method specifically designed to analyze the correlation between two variables under the

repeated measurement condition [50]. This method is designed and efficiently reported within-group variance [50]. Because the current study measured each individual three times, the Rmcorr would be an appropriate tool to analyze the data. A Rmcorr *R*-package was prepared to measure the correlations between perceived peer support and willingness of opinion expression.

The results in Table 4 showed that when the observation group increased from the minority situation to a balanced situation, then to a majority situation, the correlation between perceived support and willingness was negative and significant ( $r = -0.367$ ,  $p = 0.008$ ). According to the  $r$  and  $p$  value, statistics indicated that the decrease of perceived peer support would significantly increase the willingness of opinion expression.

Moreover, the current study examined the relationship in two more concrete scenarios as follows: (1) when the balanced opinion situation increased to the majority opinion situation; (2) when the minority opinion situation increased to the balanced opinion situation.

Results in Table 5 showed that when the observation group increased from the balanced situation to a majority situation, the correlation between perceived support and willingness was negative and nonsignificant ( $r = -0.200$ ,  $p = 0.327$ ). When the observation group increased from the minority situation to a balanced situation, the correlation between perceived support and willingness was negative and significant ( $r = -0.604$ ,  $p = 0.001$ ). According to the  $r$  and  $p$  value, statistics indicated that people in a minority opinion situation was more sensitive to the willingness of opinion expression rather than people in a majority opinion situation.

Therefore, the quasi-experiment concluded as (1) the decrease of perceived peer support would increase the willingness of opinion expression in a social media group and (2) members in a minority opinion were more sensitive to the change of peer support. When perceived less peer support, they were more willing to speak out rather than members in a majority opinion, or vice versa.

With the significant negative influence of perceived peer support towards the willingness of opinion expression, the research moved to a further step. Would individuals' change of opinion expression willingness influence a whole group's opinion expression dynamics? An agent-based model was designed to answer the question in the next study of the current research.

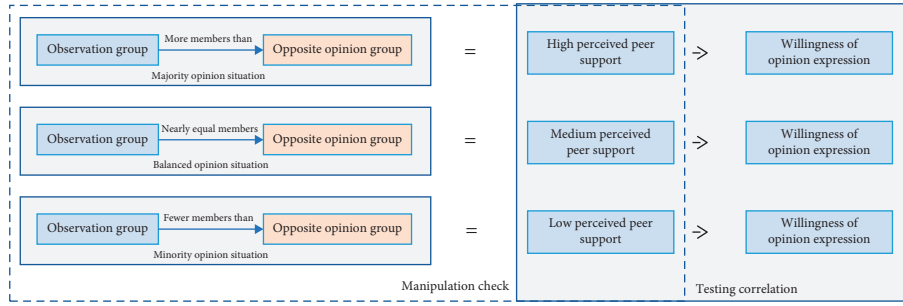


FIGURE 2: The workflow of study 1.

TABLE 2: The significant test of manipulating the level of perceived peer support.

Source	Measurement	Mean square	<i>F</i>	Sig.
Perceived peer support	Spherical distribution hypothesis	0.124	10.489	<b>0.000</b>
	Greenhouse–Geisser	0.171	10.489	<b>0.001</b>
	Huynh–Feldt	0.162	10.489	<b>0.001</b>
	Lower limit	0.248	10.489	<b>0.004</b>

Note. Spherical distribution hypothesis, Greenhouse–Geisser, Huynh–Feldt, and lower limit refer to different measurement methods.

TABLE 3: Pairwise comparison of manipulating the level of perceived peer support.

Measurement	Opinion proportion	Mean difference	S.E.	Sig. b
Perceived peer support	1 → 2	−0.055	0.042	0.621
	1 → 3	<b>−0.152 ***</b>	0.031	0.000
	2 → 3	<b>−0.097 **</b>	0.025	0.003

Note. The significance level of the mean difference is 0.05; the adjusted multiple comparison method is Bonferroni \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

## 4. Study 2: Opinion Expression Dynamics with Agent-Based Model Simulation

**4.1. Modeling Opinion Expression Dynamics of a Social Media Group.** This study has designed an ABM to simulate the opinion expression process and observe the collective level's emergent phenomena of two conflict opinion clusters' gambling. Specific operations and conclusions are detailed in the following sections.

**4.1.1. Basic Rules of the ABM.** Combined quasi-experiment conclusions with previous theories (e.g., corrective action) which mentioned opinion expression, and the study has summarized three basic rules to design the ABM.

- (1) Once the opinions are determined, it is difficult to reverse them. The only factor that can change is the willingness of opinion expression, not the opinion itself.
- (2) When people perceived their peer support is decreasing, they are more willing to speak out. The perceived peer support of one's opinion is measured

TABLE 4: The repeated measurement correlation between perceived peer support and willingness of opinion expression.

Item	Result
<i>r</i> (correlation)	−0.367 ***
degrees of freedom	49
<i>p</i> value	0.008

\*\*\*  $p < 0.001$ .

TABLE 5: The repeated measurement correlation under balanced → majority condition and minority → balanced condition.

	Balanced → majority	Minority → balanced
Item	Result	Result
<i>r</i> (correlation)	−0.200	−0.604 ***
degrees of freedom	24	24
<i>p</i> value	0.327	0.001

\*\*\*  $p < 0.001$ .

by the same ratio in the quasi-experiment that equals to

$$P(\text{support}) = \frac{\text{the number of comments that an individual received from the same opinion cluster}}{\text{the number of all comments}}. \quad (2)$$

- (3) Minority opinion cluster members are more sensitive to the change of perceived peer support while comparing with the majority.

*4.1.2. Agents and Initial Assigned Parameters in the ABM.* The study defined that all the agents were contained in the same chat group. Each agent represented a person in the group, and those agents could freely communicate with each other. In other words, this social media chat group was defined as an undirected and fully-connected social network. This network was separated into two opinion clusters, named *A* and *B*. The two clusters had conflict opinions (assigned as opinion = 1 or -1) towards each other. For each agent in one cluster, it also had an initial percentage of opinion expression willingness, which was randomly assigned by numbers ranging from 0% to 100% which is similar to Gaussian distribution.

$$P(\text{willingness}) = \text{random number}[0, 1]. \quad (3)$$

0% represented that the agent did not have any willingness of opinion expression, while 100% represented that the agent must speak out when it was given a chance. The parameter cannot exceed 1 or below 0 at any time during the simulation. When it exceeds 1 or below 0, it will be counted just as 1 or 0. To be specific, the code set random number of  $P(\text{willingness})$  based on Gaussian distribution with  $\mu = 0.5$  and  $\sigma = 1$ , if the random number of an agent exceeded 1 or below 0, the code would randomly assign a new number to this agent, until it fell into the range of [0, 1]. Thus, the final distribution of  $P(\text{willingness})$  is an approximately Gaussian distribution but with higher kurtosis. This is an acceptable assumption because social science discipline usually assumes

$$P(\text{initial support}) = \frac{\text{the number of statements in the first 15 rounds from one opinion cluster (e.g., opinion A)}}{\text{the number of all statements in the first 15 rounds}}. \quad (4)$$

The second one was  $p(\text{later support})$ , which represented the perceived peer support in the later one round of speaking. The  $p(\text{later support})$  was operationalized as

$$P(\text{later support}) = \frac{\text{the number of statements in one round from one opinion cluster (e.g., opinion A)}}{\text{the number of all statements in one round}}. \quad (5)$$

When the agents in one opinion cluster found their  $P(\text{later support})$  was less than  $P(\text{initial support})$ , that meant they perceived less peer support or vice versa.

*4.1.5. Modeling Willingness of Opinion Expression.* According to the basic rules, when agents perceive less peer support (than before), they will increase the willingness of opinion expression. On the contrary, when they perceive more peer support, they will decrease the willingness of opinion expression. Besides, minority opinion agents change

human behavior concepts following Gaussian distribution. For instance, Chun and Lee's paper [14] used structural equation modeling (SEM) to analyze the relationship of perceived peer support and willingness to speak out, while the basic presumption of SEM is the two variables follow Gaussian distribution.

To sum up, the social media group was defined as a fully connected network and had two competing opinion clusters *A* and *B*. For each agent in any cluster, it was assigned as an opinion and a percentage of opinion expression willingness. The assigned opinion could not change in the ABM, but the percentage of willingness could change over time.

*4.1.3. Round of Speaking.* The discussion in a real social media group could be conducted alternately by two conflict opinion clusters and formed a flowing stream of statements. In the ABM, the current study simulated the dynamics by defining round of speaking. The whole discussion process was segmented by several rounds of speaking. In each round, we randomly picked up several agents and asked if they chose to speak. The percentage of opinion expression willingness (which was defined in 4.1.2) was used here to determine whether the agent chose to speak or not. After one round of speaking, we moved to the next round.

*4.1.4. Modeling Perceived Peer Support.* The current study explicitly defined two types of perceived peer support. The first one was  $p(\text{initial support})$ , which represented the initial perceived peer support. The  $p(\text{initial support})$  was operationalized as

their willingness broader than majority opinion agents. The current study defined the changing percentage of willingness as  $p\%$  for minority agents, and the changing percentage of willingness as  $q\%$  for majority agents.  $p\%$  was larger than  $q\%$  if there were one majority and one minority opinion clusters. On a special occasion,  $p\%$  was equaled to  $q\%$  if the two opinion clusters had the same number of agents.

*4.1.6. The Procedure of ABM Simulation.* Using the basic rules and parameters, let the agents work as follows:

*Step 1.* Assumed that there was a social media group of 100 agents, some belong to opinion cluster *A* and some belong to *B*.

*Step 2.* Randomly assigned initial  $P(\text{willingness})$  to each agent with a similar Gaussian distribution. For instance, suppose *A* was the majority opinion and *B* was the minority opinion,  $A_i(1, 30\%)$  meant an agent in cluster *A* was assigned to opinion = 1 and  $P(\text{willingness}) = 30\%$  possibility to speak out when given a chance.  $B_j(-1, 50\%)$  meant an agent in cluster *B* was assigned to an opposite opinion = -1 and  $P(\text{willingness}) = 50\%$  possibility to speak out when given a chance.

*Step 3.* Set 100 rounds of speaking, and in each round, we randomly selected  $N = 10$  agents. Whether the ten agents chose to speak or not were depended on their percentage of opinion expression willingness. For instance, if  $A_i(1, 30\%)$  was one of the 10 agents selected in one round, it would have a 30% possibility to speak out in that round. Here, the 100 rounds of speaking are used to simulate a real discussion time of a topic. Even though it is possible that, after several thousands of discussion rounds, later the majority and minority situations are switched, it is less possible to have such a long-lasting discussion in the real world.

*Step 4.* The simulation ran 15 rounds first, then calculated  $p(\text{initial support})$  of opinion *A* and opinion *B*.

*Step 5.* Starting from round 16, for each round, if one opinion cluster's perceived peer support changed comparing with the total first 15 rounds, we had four situations:

*A.* The opinion cluster was the minority and the peer support decreased. Those agents would increase their opinion expression willingness by  $p\%$ . For instance, if cluster *B* found

$$\begin{aligned} P(\text{later support}) &< P(\text{initial support}), \\ \text{then } B_j(t+1) &= (-1, 50^* (1 + p)\%). \end{aligned} \quad (6)$$

*B.* The opinion cluster was the majority and the peer support decreased. Those agents would increase their opinion expression willingness by  $q\%$ . For instance, if cluster *A* found

$$\begin{aligned} P(\text{later support}) &< P(\text{initial support}), \\ \text{then } A_i(t+1) &= (1, 30^* (1 + q)\%). \end{aligned} \quad (7)$$

*C.* The opinion cluster was the minority and the peer support increased. Those agents would decrease their opinion expression willingness by  $p\%$ . For instance, if cluster *B* found

$$\begin{aligned} P(\text{later support}) &> P(\text{initial support}), \\ \text{then } B_j(t+1) &= (-1, 50^* (1 - p)\%). \end{aligned} \quad (8)$$

*D.* The opinion cluster was the majority and the peer support increased. Those agents would decrease their opinion expression willingness by  $q\%$ . For instance, if cluster *A* found

$$\begin{aligned} P(\text{later support}) &> P(\text{initial support}), \\ \text{then } A_i(t+1) &= (1, 30^* (1 - q)\%). \end{aligned} \quad (9)$$

Once determined an opinion, people in the group would not change their views. The most extreme possibility was the opinion remained unchanged, but the percentage of opinion expression willingness was 0% or 100%.

This study used a C++ program to run the ABM.

*4.2. Results.* Three key parameters were needed to adjust in the ABM as follows: the percentage of opinion expression willingness  $p\%$  and  $q\%$ , and the number of agents in two opinion clusters:  $a$  and  $100 - a$ . This simulation set several levels of each parameter: When  $a = 10, 20, 30,$  and  $40,$   $(p\%, q\%) = (5\%, 2.5\%), (10\%, 5\%), (15\%, 7.5\%),$  and  $(20\%, 10\%).$  When  $a = 50,$   $(p\%, q\%) = (5\%, 5\%), (10\%, 10\%), (15\%, 15\%),$  and  $(20\%, 20\%).$  Therefore, this simulation had  $5 * 4 = 20$  in different situations. The ABM in each situation ran 100 times and took the average to construct the curve of two clusters' opinion expression.

*4.2.1. Model Analysis When Two Opinions Were Balanced.* To simulate the situation that two opinions were balanced, this research assigned two opposite opinions, *A* and *B*, and each of them had 50 agents. Their initial willingness of opinion expression was a randomly assigned normal distribution before the simulation. Following the rules mentioned above, the ABM drew the *opinion expression percentage* of opinion *A*. The opinion expression percentage was operationalized as

$$P(\text{expression}) = \frac{\text{the number of statements in one round from one opinion cluster (e.g., opinion A)}}{\text{the number of all statements in one round}} \quad (10)$$

Note that the opinion expression percentage is counted by statements made by agents, which is different from the willingness of opinion expression in people's minds. If at a specific time point, the opinion expression percentage of

cluster *A* equaled to 50%, that meant opinion cluster *A* agents made 50% of the statements in the discussion.

From Figure 3, the current study concluded that the two equally balanced opinions antagonized one another and

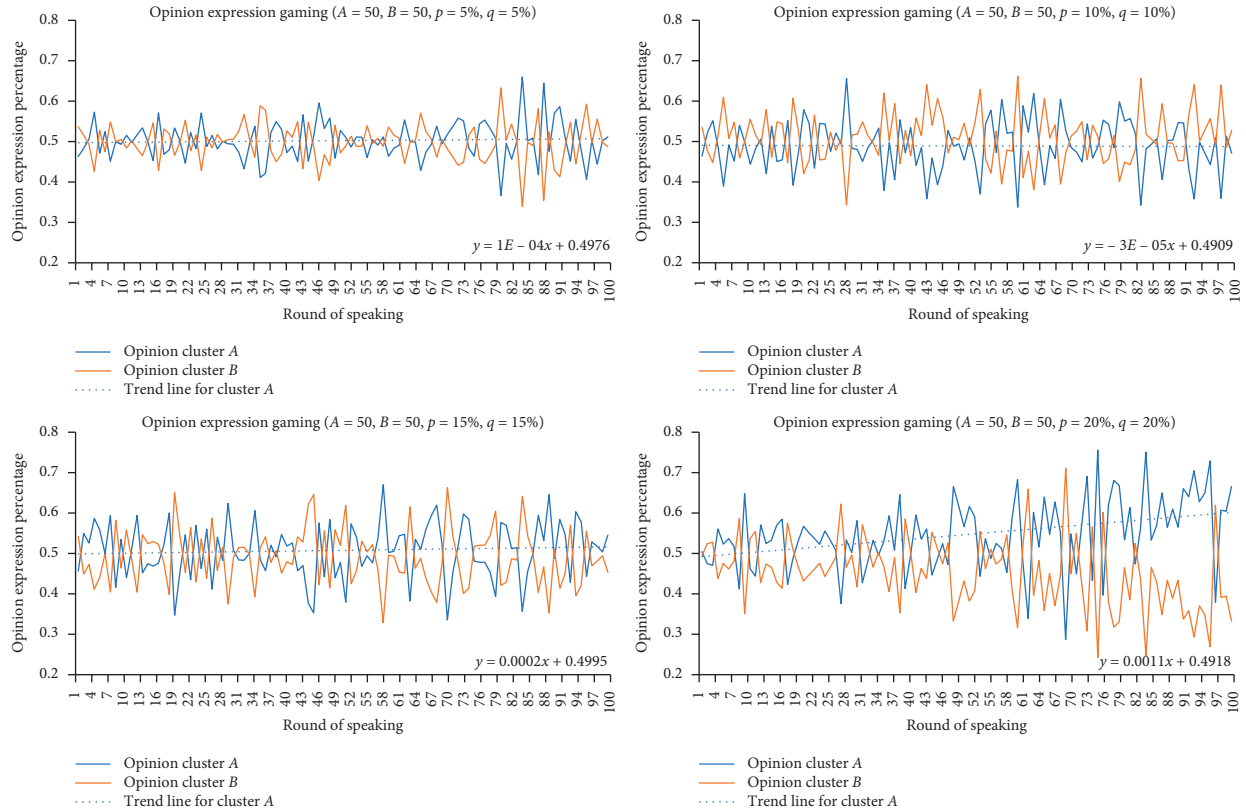


FIGURE 3: Balanced opinion situation ( $A = 50$  and  $B = 50$ ). *Note.* Opinion expression percentage, with  $A$  (minority cluster) = blue thick line and  $B$  (majority cluster) = red thin line. Round of speaking represents time because each discussion allows 100 rounds of speaking. All the results were shown in the average score of 100 times' simulation.

were always fluctuating up and down around 50%. Different values of the parameter  $p\%$  (5, 10, and 15) shared similar trends. However, when the parameter  $p\%$  gets larger to 20%, the opinion expression percentage of opinion  $A$  could fluctuate to 60% at the end of 100 rounds of speaking, which is more unstable than small  $p\%$  values.

**4.2.2. Model Analysis When the Social Media Group Was Divided into Two Opinions: One Majority and One Minority.** To simulate the situation that two opinions had one minority and another majority, this research assigned two opposite opinions,  $A$  and  $B$ , with four situations:  $A = 10, B = 90$ ;  $A = 20, B = 80$ ;  $A = 30, B = 70$ ;  $A = 40, B = 60$ . Each agent's initial willingness of opinion expression was a proportion that was randomly assigned with normal distribution before the simulation. Following the rules mentioned above, the simulation ran 100 times in each situation and then drew the opinion expression percentage on the minority/majority opinion discussion in Figure 4.

In Figure 4, the blue thick line referred to the minority opinion; the red thin line referred to the majority opinion. From the figure, the current study indicated that the two opinions antagonized one another, and the minority opinion spoke out more than expected. For instance, the minority opinion members constituted 10% of the population, while

constituted 15–25% in opinion expression at the end of 100 rounds of speaking. Different values of parameter  $p\%$  shared the same trends. Besides, the increasing range of minority opinion was enhanced by parameter  $p\%$ .

However, with the increasing members of the minority opinion, the opinion expression percentage became closer to the percentage of the population. For instance, in Figure 5, when the minority population increased from 10 to 40, the expression percentage did not increase continuously above 40% but fluctuated around 40% instead. Different values of parameter  $p\%$  shared the same trends.

Figures 6 and 7 show the simulation results when  $A = 20$  and 30. Their opinion dynamics fluctuated less than the position in Figure 4 but more than the position in Figure 5.

To sum up, according to the rules of ABM, the minority opinion's expression percentage behaved better than the percentage of the total population when the opinion had very few supporters. For instance, when  $A = 10$  and  $B = 90$ , the opinion expression percentage could be around 15%–25% for cluster  $A$ , but only 75%–85% for cluster  $B$ . The larger the number of parameters  $p$  and  $q$ , the more unstable the trend would be. However, we could not find enough support that the minority opinion could be reversed to be the majority opinion. As the number of minority opinion members increased, the superiority of opinion expression gradually declined.

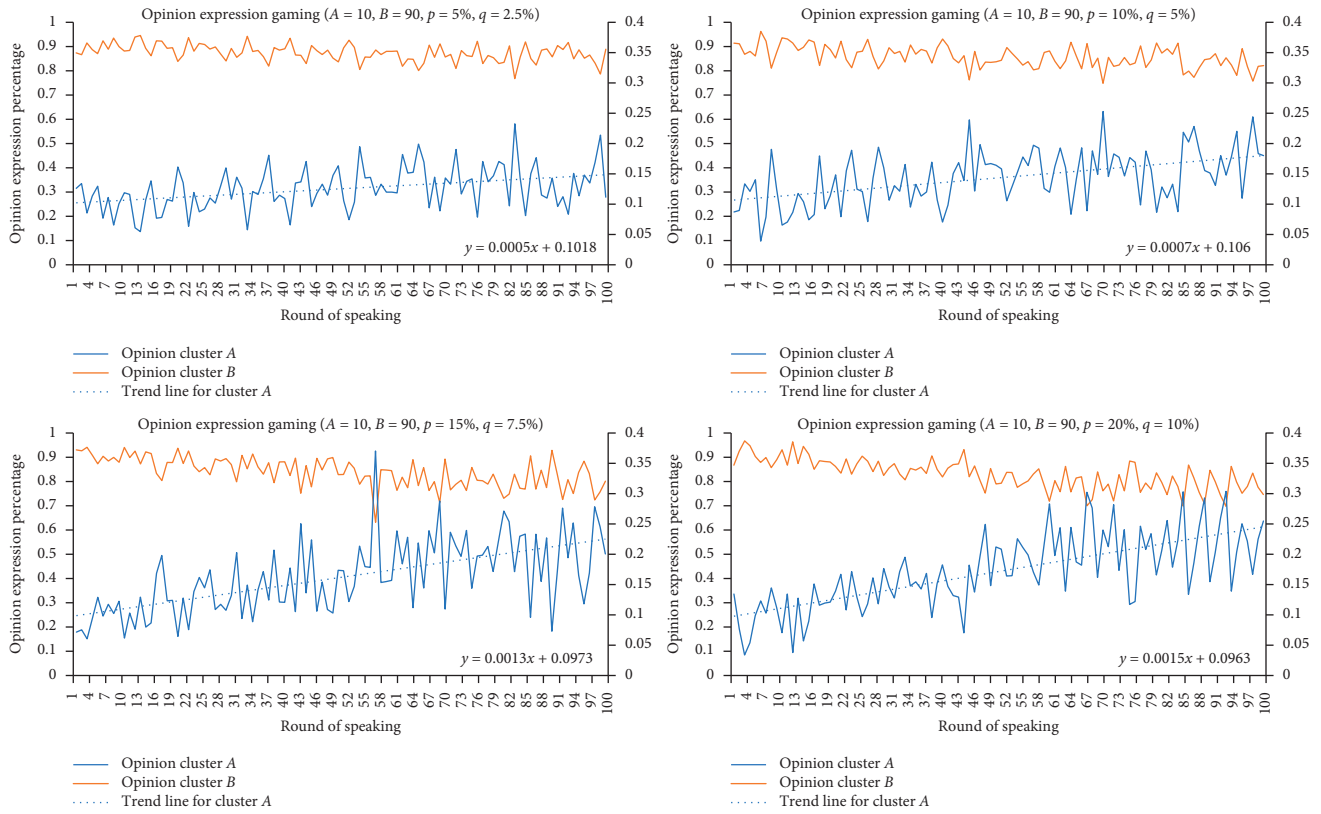


FIGURE 4: Minority/majority opinion expression. (For example, A = 10, B = 90). *Note.* Opinion expression percentage, with A (minority cluster) = blue thick line and B (majority cluster) = red thin line. Round of speaking represents time because each discussion allows 100 rounds of speaking. All the results were shown in the average score of 100 times' simulation.

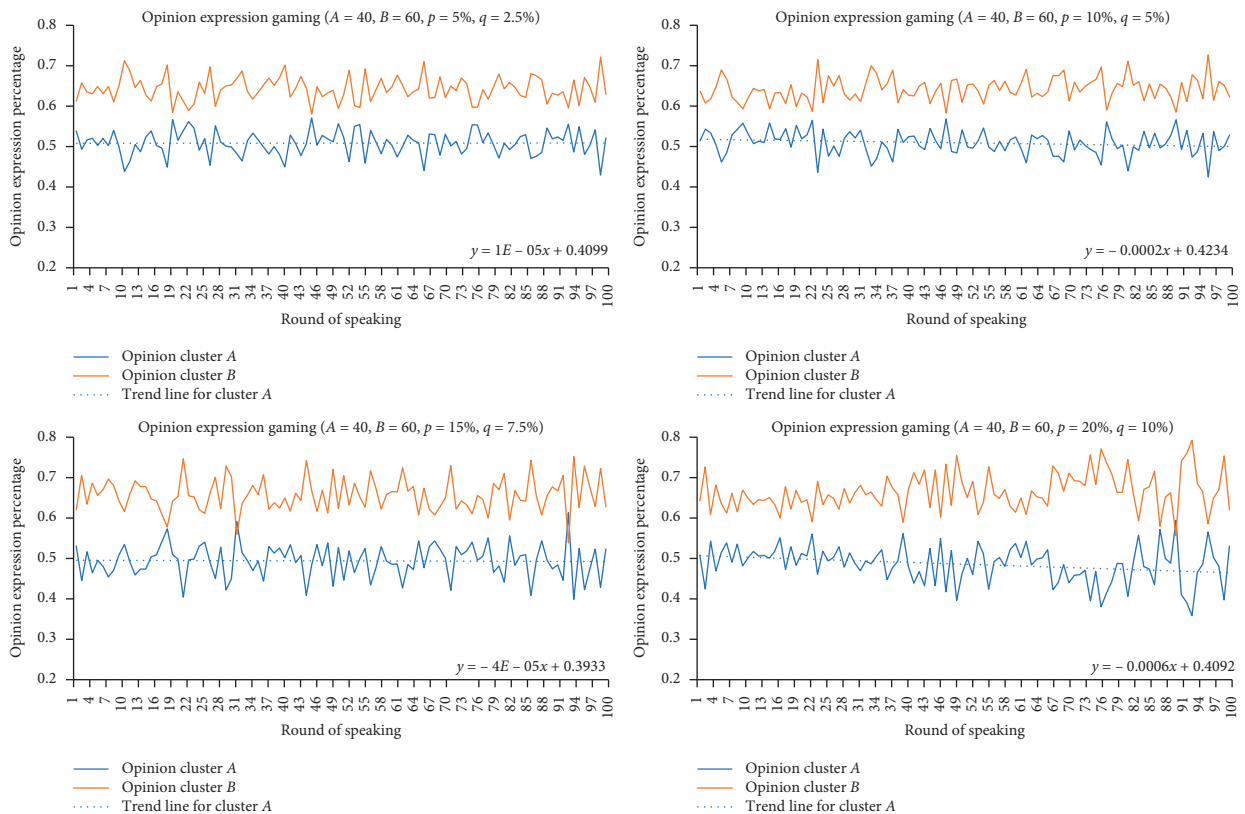


FIGURE 5: Minority/majority opinion expression. (For example, A = 40, B = 60). *Note.* Opinion expression percentage, with A (minority cluster) = blue thick line and B (majority cluster) = red thin line. Round of speaking represents time because each discussion allows 100 rounds of speaking. All the results were shown in the average score of 100 times' simulation.



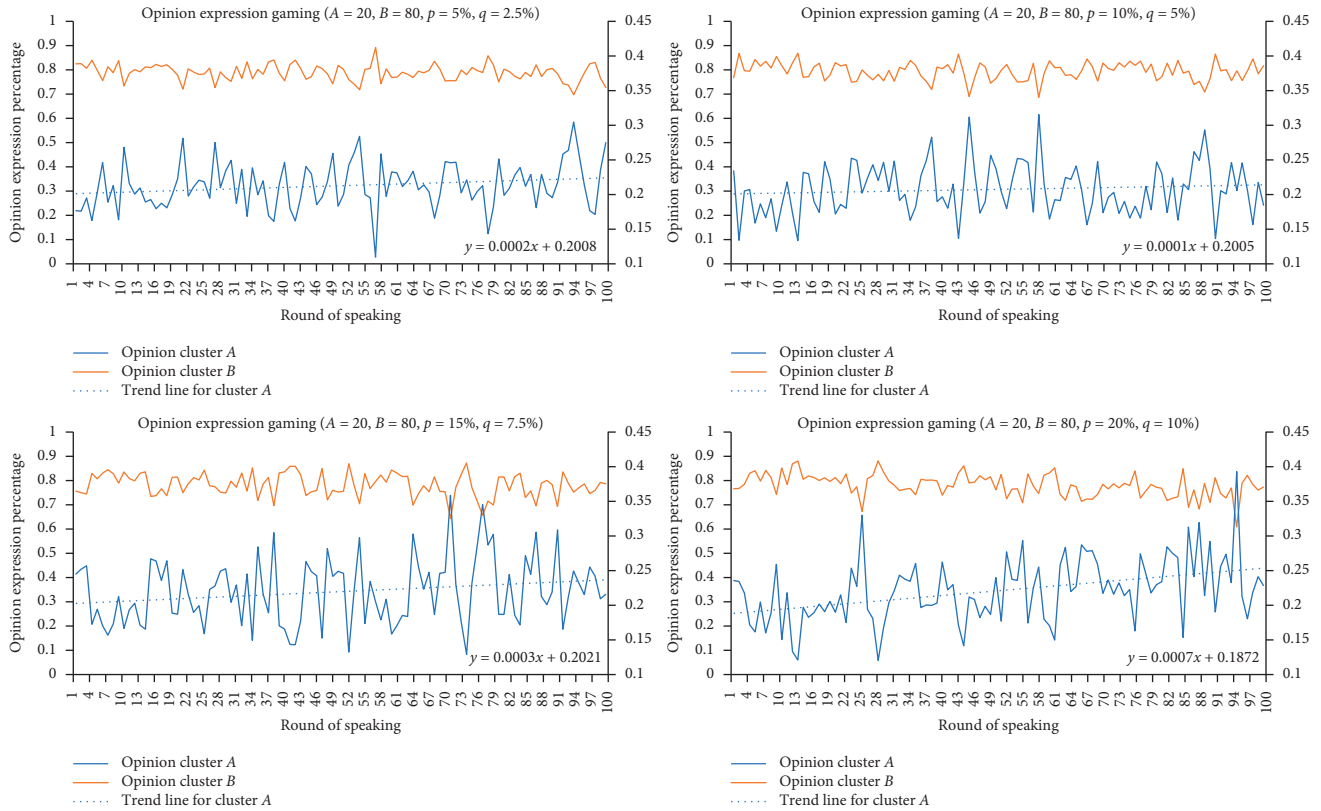


FIGURE 6: Minority/majority opinion expression. (For example,  $A = 20, B = 80$ ). *Note.* Opinion expression percentage, with  $A$  (minority cluster) = blue thick line and  $B$  (majority cluster) = red thin line. Round of speaking represents time because each discussion allows 100 rounds of speaking. All the results were shown in the average score of 100 times' simulation.

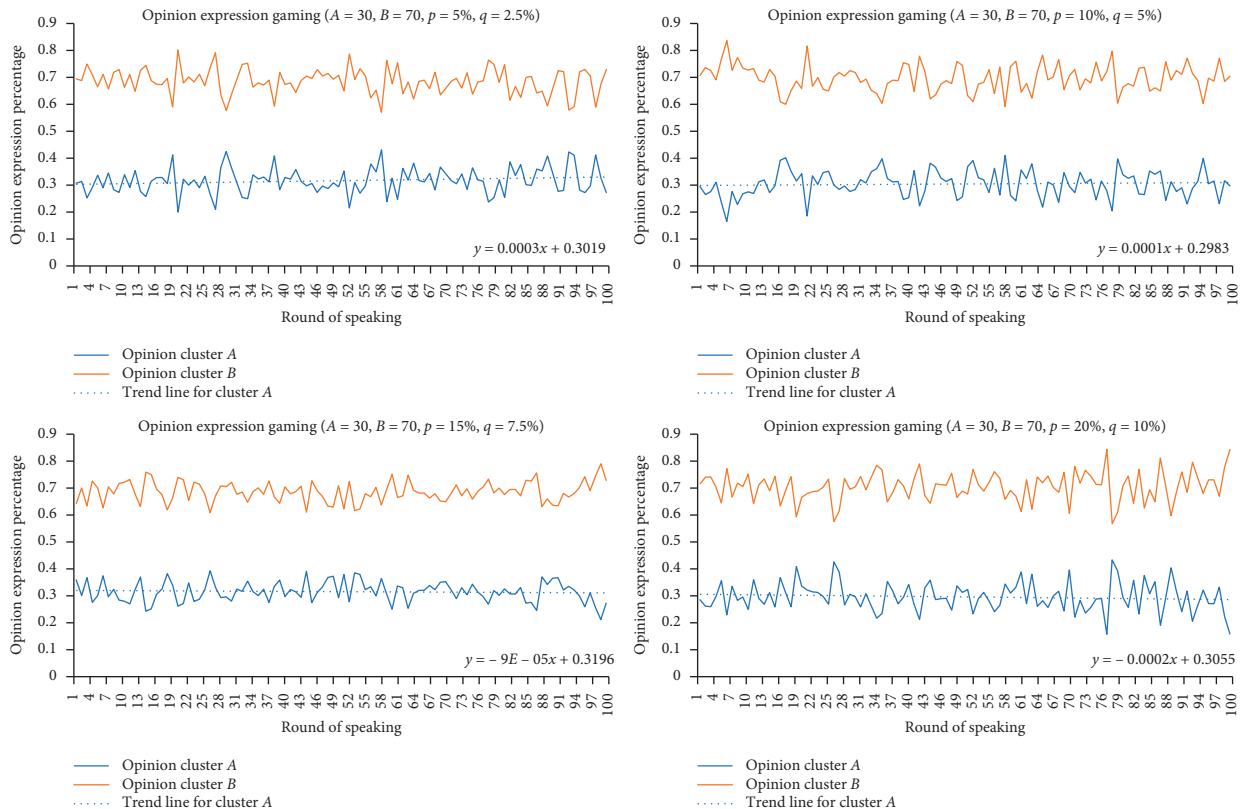


FIGURE 7: Minority/majority opinion expression. (For example,  $A = 30, B = 70$ ). *Note.* Opinion expression percentage, with  $A$  (minority cluster) = blue thick line and  $B$  (majority cluster) = red thin line. Round of speaking represents time because each discussion allows 100 rounds of speaking. All the results were shown in the average score of 100 times' simulation.

## 5. Conclusion and Discussion

The current research concluded that the minority opinion members were more sensitive to the change of peer support than the majority opinion members. Notably, an individual's willingness of opinion expression could be encouraged while the perceived peer support decreased, which was significantly different from the traditional spiral of silence studies as well as the simulation studies. One potential explanation of this finding is, the fear of isolation is the premise of the spiral of silence [27], while people tend to find peers in social media groups, so they may not feel enough isolated even if they receive less support. Besides, the studies of corrective action and hostile media have proposed that when people feel their opinions are unfairly suppressed, they will feel more indignation and then protecting their opinions rather than perceiving fear [51, 52]. Thus, less support in social media groups can encourage the willingness of opinion expression.

New phenomena emerged during the simulation of opinion expression dynamics in the social media group. The simulation model based on the quasi-experimental results showed that the increase of the willingness of opinion expression on an individual level could encourage the real opinion expression and prevent the whole cluster of minority opinion from falling into silence. That means it is quite difficult to achieve complete suppression of one opinion in social media groups. However, findings also indicate that enhancing willingness of opinion expression is not a panacea for the minority opinion members to demand greater influence. Because results showed that the minority's opinion expression percentage could only behave relatively better than the minority opinion supporters' population percentage, then floated up and down. Reaching the final one-side win or reversing the minority opinion to the majority may both need time and luck, which is hard to achieve in real life.

The current research findings have practical implications. First, people should be vigilant if some specific interest groups intend to manipulate their willingness of opinion expression. Because it can help seek greater influence not matching their capacity and then mislead the public opinion to be more beneficial to themselves. Second, those who support a minority opinion should have the confidence to survive in social media groups. Third, it is probably not necessary for majority opinion members to pursue a complete suppression of opinions; on the contrary, it is necessary to acknowledge that most changes of pluralistic public opinions are ordinary and respectful. Finally, social media chat groups are widely existed nowadays but are different from those open social media platforms. Therefore, more researches are encouraged to study further about the cause, effect, and mechanism of opinion dynamics in social media groups.

Here are some limitations of the current research. First, in order to maintain a good external validity of the results, the quasi-experiment tries to simulate a real environment of a social media chat group and uses behavioral data to operationalize psychological concepts. However, it does

not test the difference between intention and behavior, such as perceived peer support and received peer support, willingness to speak out, and the actual speak out. We would suggest future studies to come up with a better design to measure the intensions. Second, the quasi-experiment needs more participants in the observation group; and the percentage changes of the observational group between the majority, balance, and minority situations need to be controlled more strictly. Third, we do not have enough empirical data to support that people's willingness of opinion expression is normally distributed, but the researchers assume that, it is in the agent-based model. Although it is a relatively common assumption that human behavior is following Gaussian distribution, it could have possibilities and may cause bias in the simulation. Therefore, researchers plan to do further studies focusing on exploring more theoretically-supported communication strategies to simulate the collective level phenomenon and design more rigorous experiments to test the individual level presumptions.

## Data Availability

The quasi-experimental data used to support the findings of this study are available from the corresponding author upon request. The agent-based model codes are included within the supplementary information file.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Supplementary Materials

The supplementary file includes codes of conducting the agent-based model. (*Supplementary Materials*)

## References

- [1] M. Jing and C. Zang, "On the information flow mode, demassification communication and the influence of mass media over the social cohesion under the age of media convergence," *Journalism & Communication*, vol. 5, pp. 34–42, 2011.
- [2] Y. Chen, "Comparison of communication mechanism between micro-blog and WeChat under different identity," *Press Circles*, vol. 3, pp. 54–57, 2015.
- [3] G. M. Yu, "Some ideas about the future development of the media industry," *News Front*, vol. 1, 2014.
- [4] K. N. Hampton, I. Shin, and W. Lu, "Social media and political discussion: when online presence silences offline conversation," *Information, Communication & Society*, vol. 20, no. 7, pp. 1090–1107, 2017.
- [5] Z. A. Zhang and K. R. Shu, "Research on the public opinion of WeChat: relationship network and ecological characteristics," *Shanghai Journalism Review*, vol. 20, no. 6, pp. 29–37, 2016.
- [6] J. Habermas, *The Structural Transformation of the Public Sphere*, Xue Lin Publishing House, Shanghai, China, 1962.
- [7] D. Goldie, M. Linick, H. Jabbar, and C. Lubienski, "Using bibliometric and social media analyses to explore the "echo chamber" hypothesis," *Educational Policy*, vol. 28, no. 2, pp. 281–305, 2014.

- [8] G. H. Guo, "On the "group polarization" tendency of the main body of network public opinion," *Journal of Social Science of Human Normal University*, vol. 6, pp. 110–113, 2004.
- [9] E. Noelle-Neuman, *The Spiral of Silence: Public Opinion-Our Social Skin*, Peking University Press, Beijing, China, 1974.
- [10] D. A. Scheufele and P. Moy, "Twenty-five years of the spiral of silence: a conceptual review and empirical outlook," *International Journal of Public Opinion Research*, vol. 12, pp. 3–28, 2000.
- [11] L. Willnat, W. P. Lee, and B. H. Detenber, "Individual-level predictors of public outspokenness: a test of the spiral of silence theory in Singapore," *International Journal of Public Opinion Research*, vol. 14, no. 4, pp. 391–412, 2002.
- [12] M. J. Lee and J. W. Chun, "Reading others' comments and public opinion poll results on social media: social judgment and spiral of empowerment," *Computers in Human Behavior*, vol. 65, pp. 479–487, 2016.
- [13] G. Neubaum and N. C. Krämer, "Opinion climates in social media: blending mass and interpersonal communication," *Human Communication Research*, vol. 43, no. 4, pp. 464–476, 2017.
- [14] J. W. Chun and M. J. Lee, "When does individuals' willingness to speak out increase on social media? Perceived social support and perceived power/control," *Computers in Human Behavior*, vol. 74, pp. 120–129, 2017.
- [15] T. Zerback and N. Fawzi, "Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out," *New Media & Society*, vol. 19, no. 7, pp. 1034–1051, 2017.
- [16] A. F. Hayes, "Exploring the forms of self-censorship: on the spiral of silence and the use of opinion expression avoidance strategies," *Journal of Communication*, vol. 57, no. 4, pp. 785–802, 2007.
- [17] P. Moy, D. Domke, and K. Stamm, "The spiral of silence and public opinion on affirmative action," *Journalism & Mass Communication Quarterly*, vol. 78, no. 1, pp. 7–25, 2001.
- [18] E. Katz, "Publicity and pluralistic ignorance: notes on "the spiral of silence"," in *Public Opinion and Social Change. For Elisabeth Noelle-Neumann*, H. Baier, Kepplinger, and K. Reumann, Eds., pp. 28–38, Westdeutscher Verlag, Wiesbaden, Germany, 1981.
- [19] E. Wartella, C. Whitney, *Mass Communication Review Yearbook*, Sage, Vol. 4, Beverley Hills, CA, USA, 1983.
- [20] H. Oshagan, "Reference group influence on opinion expression," *International Journal of Public Opinion Research*, vol. 8, no. 4, pp. 335–354, 1996.
- [21] C. J. Glynn, A. F. Hayes, and J. Shanahan, "Perceived support for one's opinions and willingness to speak out: a meta-analysis of survey studies on the "spiral of silence"," *Public Opinion Quarterly*, vol. 61, no. 3, pp. 452–463, 1997.
- [22] W. P. Davison, "The third-person effect in communication," *Public Opinion Quarterly*, vol. 47, no. 1, pp. 1–15, 1983.
- [23] H. Rojas, "'Corrective' actions in the public sphere: how perceptions of media and media effects shape political behaviors," *International Journal of Public Opinion Research*, vol. 22, no. 3, pp. 343–363, 2010.
- [24] M. Duncan and D. Coppini, "Party v. The People: testing corrective action and supportive engagement in a partisan political context," *Journal of Information Technology & Politics*, vol. 16, no. 3, pp. 265–289, 2019.
- [25] M. Duncan, A. Pelled, D. Wise et al., "Staying silent and speaking out in online comment sections: the influence of spiral of silence and corrective action in reaction to news," *Computers in Human Behavior*, vol. 102, pp. 192–205, 2020.
- [26] Y. Tsfati, N. J. Stroud, and A. Chotiner, "Exposure to ideological news and perceived opinion climate: testing the media effects component of spiral-of-silence in a fragmented media landscape," *The International Journal of Press/Politics*, vol. 19, no. 1, pp. 3–23, 2014.
- [27] P. J. Shoemaker, M. Breen, and M. Stamper, "Fear of social isolation: testing an assumption from the spiral of silence," *Irish Communication Review*, vol. 8, pp. 65–78, 2000.
- [28] N. Pang, S. S. Ho, A. M. R. Zhang, J. S. W. Ko, W. X. Low, and K. S. Y. Tan, "Can spiral of silence and civility predict click speech on Facebook?" *Computers in Human Behavior*, vol. 64, pp. 898–905, 2016.
- [29] G. W. Yun, S.-Y. Park, and S. Lee, "Inside the spiral: hostile media, minority perception, and willingness to speak out on a weblog," *Computers in Human Behavior*, vol. 62, pp. 236–243, 2016.
- [30] C. Wells, K. J. Cramer, M. W. Wagner et al., "When we stop talking politics: the maintenance and closing of conversation in contentious times," *Journal of Communication*, vol. 67, no. 1, pp. 131–157, 2017.
- [31] C. J. Glynn and J. M. McLeod, "Implication of the spiral of silence theory for communication and public opinion research," in *Political Communication Yearbook*, D. R. Sanders, L. L. Kaid, and D. Nimmo, Eds., pp. 43–65, Southern Illinois University Press, Carbondale, IL, USA, 1984.
- [32] D. G. McDonald, C. J. Glynn, S.-H. Kim, and R. E. Ostman, "The spiral of silence in the 1948 presidential election," *Communication Research*, vol. 28, no. 2, pp. 139–155, 2001.
- [33] L. Willnat, "Mass media and political outspokenness in Hong Kong: linking the third-person effect and the spiral of silence," *International Journal of Public Opinion Research*, vol. 8, no. 2, pp. 187–212, 1996.
- [34] R. K. Sawyer, "The mechanisms of emergence," *Philosophy of the Social Sciences*, vol. 34, no. 2, pp. 260–282, 2004.
- [35] M. Pineda and G. M. Buendía, "Mass media and heterogeneous bounds of confidence in continuous opinion dynamics," *Physica A: Statistical Mechanics and Its Applications*, vol. 420, pp. 73–84, 2015.
- [36] W. B. Wang, "The present situation and prospect of the research on evolutionary game theory," *Statistics and Decision*, vol. 25, no. 3, pp. 158–161, 2009.
- [37] N. Jung, E. S. Cho, J. H. Choi, and J. W. Lee, "Agent-based models in social physics," *Journal of the Korean Physical Society*, vol. 72, pp. 1272–1280, 2018.
- [38] Z.-D. Zhang, "Conjectures on the exact solution of three-dimensional (3D) simple orthorhombic Ising lattices," *Philosophical Magazine*, vol. 87, no. 34, pp. 5309–5419, 2007.
- [39] F. W. S. Lima, "Three-state majority-vote model on square lattice," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 4, pp. 1753–1758, 2012.
- [40] Y. J. Liu, Q. Li, and W. Y. Niu, "A summary of the model of public opinion dynamics," *Management Review*, vol. 25, no. 1, pp. 167–176, 2013.
- [41] Y. Wu, Y.-J. Du, X.-Y. Li, and X.-L. Chen, "Exploring the spiral of silence in adjustable social networks," *International Journal of Modern Physics C*, vol. 26, no. 11, Article ID 1550125, 2015.
- [42] B. Ross, L. Pilz, B. Cabrera, F. Brachten, G. Neubaum, and S. Stieglitz, "Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks," *European Journal of Information Systems*, vol. 28, no. 4, pp. 394–412, 2019.
- [43] L. Deng, Y. Liu, and F. Xiong, "An opinion diffusion model with clustered early adopters," *Physica A: Statistical*

- Mechanics and Its Applications*, vol. 392, no. 17, pp. 3546–3554, 2013.
- [44] A. Boutyline and R. Willer, “The social structure of political echo chambers: variation in ideological homophily in online networks,” *Political Psychology*, vol. 38, no. 3, pp. 551–569, 2017.
- [45] D. Trilling, M. van Klingeren, and Y. Tsfati, “Selective exposure, political polarization, and possible mediators: evidence from The Netherlands,” *International Journal of Public Opinion Research*, vol. 29, pp. edw003–213, 2017.
- [46] D. Sohn and N. Geidner, “Collective dynamics of the spiral of silence: the role of ego-network size,” *International Journal of Public Opinion Research*, vol. 28, no. 1, pp. 25–45, 2015.
- [47] J. J. H. Zhu, Y. Huang, and X. Zhang, “Dialogue on computational communication research: origins, theories, methods, and research questions,” *Communication & Society*, vol. 44, pp. 1–24, 2018.
- [48] P. L. East, L. E. Hess, and R. M. Lerner, “Peer social support and adjustment of early adolescent peer groups,” *The Journal of Early Adolescence*, vol. 7, no. 2, pp. 153–163, 1987.
- [49] C.-L. Dennis, “Postpartum depression peer support: maternal perceptions from a randomized controlled trial,” *International Journal of Nursing Studies*, vol. 47, no. 5, pp. 560–568, 2010.
- [50] J. Z. Bakdash and L. R. Marusich, “Repeated measures correlation,” *Frontiers in Psychology*, vol. 8, p. 456, 2017.
- [51] C. S. Carver, “Negative affects deriving from the behavioral approach system,” *Emotion*, vol. 4, no. 1, pp. 3–22, 2004.
- [52] H. Hwang, Z. Pan, and Y. Sun, “Influence of hostile media perception on willingness to engage in discursive activities: an examination of mediating role of media indignation,” *Media Psychology*, vol. 11, no. 1, pp. 76–97, 2008.

## Research Article

# Social Networks through the Prism of Cognition

Radosław Michalski <sup>1</sup>, Bolesław K. Szymanski <sup>2,3</sup>, Przemysław Kazienko <sup>1</sup>,  
Christian Lebiere <sup>4</sup>, Omar Lizardo <sup>5</sup>, and Marcin Kulisiewicz <sup>1</sup>

<sup>1</sup>Department of Computational Intelligence, Faculty of Computer Science and Management,  
Wrocław University of Science and Technology, Wrocław, Poland

<sup>2</sup>Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, USA

<sup>3</sup>Spółeczna Akademia Nauk, Łódź, Poland

<sup>4</sup>Carnegie Mellon University, Pittsburgh, PA, USA

<sup>5</sup>University of California, Los Angeles, CA, USA

Correspondence should be addressed to Radosław Michalski; [radoslaw.michalski@pwr.edu.pl](mailto:radoslaw.michalski@pwr.edu.pl)

Received 11 June 2020; Revised 18 November 2020; Accepted 24 December 2020; Published 8 January 2021

Academic Editor: Fei Xiong

Copyright © 2021 Radosław Michalski 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.

Human relations are driven by social events—people interact, exchange information, share knowledge and emotions, and gather news from mass media. These events leave traces in human memory, the strength of which depends on cognitive factors such as emotions or attention span. Each trace continuously weakens over time unless another related event activity strengthens it. Here, we introduce a novel cognition-driven social network (CogSNet) model that accounts for cognitive aspects of social perception. The model explicitly represents each social interaction as a trace in human memory with its corresponding dynamics. The strength of the trace is the only measure of the influence that the interactions had on a person. For validation, we apply our model to NetSense data on social interactions among university students. The results show that CogSNet significantly improves the quality of modeling of human interactions in social networks.

## 1. Introduction

There is a fundamental difference between the ways social events are currently represented by network analysts and the way they are perceived and cognitively processed by humans. In the former, the discrete nature of events is retained and the weights of social network edges are updated once per each relevant event. In human memory, by contrast, perception of events changes continuously over time. Moreover, the initial strength of a trace depends on cognitive factors defined by states of mind of participants and specific aspects of their interactions. Decisions about whether to initiate, maintain, or discontinue social relations involve cognitive processes operating on all relevant information stored in memory traces over specific time-scales.

To date, however, there are no models that accurately represent the dynamics of social relations strictly through their corresponding traces in human memory. To address

this gap, we introduce a novel cognition-driven social network (CogSNet) model that captures impact of human memory on perception of accumulated events and on decisions to form, maintain, or dissolve social relations. The model explicitly represents some of the human memory dynamics, such as the gradual decay of memory traces over time. With suitable data, it can be extended to include additional cognitive aspects, such as individual levels of sensitivity to relevant events, emotions, or distractions during perception of events. Hence, the model is capable of capturing the dynamics of social interactions in natural settings from the cognitive perspective of each participant.

To evaluate the empirical performance of the proposed model, we compare model's results obtained with parameters fitted using behavioral communication event data, to the ground truth perceptual data collected from human participants with regard to their most important relations. The comparisons reveal that the perception of the depth of

interactions between people is well captured by the CogSNet model. At any given point in time, the model can compute the current strength of memory traces, including the impact of discrete events creating or reinforcing these traces. The computed state of each social relation (e.g., salient versus decayed) can then be compared against the externalization of cognition in the form of self-reports. The CogSNet model with parameters fitted using recorded event data achieves high accuracy in fitting ground truth data for the NetSense network. This success demonstrates the importance of incorporating cognitive processes and memory dynamics for adequate modeling of the dynamics of social relationships.

Memory is considered to be one of the most important components of human cognition. This is especially the case given the necessity to efficiently retrieve large amounts of knowledge and to select information from a noisy environment. Hence, one of the fundamental challenges for cognitive science has been to understand the mechanisms involved in managing information in human memory [1]. The exact details of these mechanisms are difficult to firmly establish since the human brain is a highly complex system featuring strong differences across individuals based on experimental tuning [2]. Yet, there have been significant advances in the direction of developing good quality working models of memory and other core cognitive processes. For example, the ACT-R cognitive architecture does a good job of modeling core features of human declarative memory, successfully replicating a large number of well-known effects. For our purposes, the most important among these empirical regularities are the well-established primacy and recency effects for list memory [3].

Limited capacity and the gradual decay of traces in memory over time have been confirmed by many studies analyzing the relationship between time and recall [4–6]. The forgetting mechanism is modeled by a parametric function describing how well a given item will be recalled as a function of time. It appears that humans tend to underestimate the number of events they experience, i.e., they actually forget faster than they think they do [5]. Obviously, any reminiscence of a particular event primed by an external situation allows people to remember the information longer. Yet ultimately, the limited capacity paradigm and the forgetting function are chiefly responsible for controlling the lifetime of such reminiscence in memory [7]. There are relatively many papers within the broader domain of link

predictions, but they treat human communication as a series of fixed time windows that reflect neither human cognition nor continuity of forgetting. For example, Reader et al. [8] identified over a dozen of predictive factors quantifying phone communication like time of the last call or frequency of calls within a 4-week period that would enable to forecast persistence of social relationships, i.e., communication, also in the following 4-week period. The influences among culture, cognition, and network thinking are mentioned in [9] but in a very qualitative way, without any hint how to quantify such influence into a model, which is essential for our approach. Some other works also study the problem of recalling the names of their peers in a variety of activities [10–12], but they focus on a detailed analysis of personal attributes differentiating subjects in terms of forgetting, while here we concentrate on the models of forgetting.

## 2. Materials and Methods

The human brain records events as they arrive, but only a small fraction of the incoming information is stored in long-term memory due to its limited capacity. The forgetting mechanism dictates that the chance to recall a given event decreases gradually as we move forward from the time of first exposure [6]. In some sense, it is similar to graph-streaming [13] and feed-based social media network cascades [14] scenarios where incoming events are ordered by their arrival time, but only some of them are kept.

Accordingly, the CogSNet model uses a *forgetting function*  $f$  to account for the decreasing probability of keeping aging of memory traces over time. Forgetting is thus a monotonically nonincreasing function of time with  $f(0) = 1$  and  $f(t) \geq 0$  for all  $t > 0$ . It is defined by three parameters: *reinforcement peak*  $0 < \mu \leq 1$ , *the forgetting intensity*  $\lambda$  (both used in the ACT-R model), and additional third parameter, *forgetting threshold*  $0 < \theta < \mu$ .

For clarity, we present here the basic CogSNet model; a more general version is defined in Appendix and Appendix equations (A.1) and (A.2). This model processes an event happening at time  $t$  as follows. If the event involves a pair of unconnected nodes  $i$  and  $j$ , an edge  $(ij)$  linking these nodes is created. This edge is assigned weight  $w_{ij}(t)$  equal to the reinforcement peak value  $\mu$ . Otherwise, if the nodes involved are already connected, the forgetting function  $f$  is used to set the weight of the edge connecting these nodes:

$$w_{ij}(t) = \begin{cases} \mu, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ \mu + w_{ij}(t_{ij})f(t - t_{ij})(1 - \mu), & \text{otherwise,} \end{cases} \quad (1)$$

where  $t_{ij}$  is the time of the preceding event for this edge. Processing of each event ends with advancing  $t_{ij}$  to  $t$ . Initially,  $t_{ij}$  and  $w_{ij}(0)$  are set to 0. Equation (1) and limit on value of  $\mu$  ensure that a weight of any edge is at most 1.

The weight  $w_{ij}(t)$  of an edge  $(ij)$  between two nodes at any user selected time  $t$  is computed as follows. Once all the

relevant events up to time  $t$  are processed, we simply set  $w_{ij}(t) = w_{ij}(t_{ij}) * f(t - t_{ij})$ . If the result is less than the forgetting threshold  $\theta$ ,  $w_{ij}(t)$  is reduced to zero and the edge is removed. A threshold is needed with forgetting functions, such as power and exponential forgetting, that are positive for nonnegative arguments. Otherwise, an edge would get

positive weight at creation and would always stay positive; i.e., all created memory traces would never cease to exist. The reinforcement peak  $\mu$  defines the impact of an event on the weight of the edge relevant to this event. This peak is a global model parameter here. In principle, the peak can be adjusted according to the event or node type to allow for individualized event perception.

In Figure 1, we compare the CogSNet model with the previous proposals for representing temporal network dynamics. The most common approach for representing social network dynamics is to use interaction sequences [15]. Under this method, each event is time-stamped and the weights are added to the edges connecting nodes involved in this event cf. Figure 1 ((I) and (IIa)). Moreover, a given edge is active (exists) only at a given time  $t$ . This is the most granular approach as it is capable of tracking all the events occurring between nodes while preserving the temporal order of events.

In contrast, a static binary network representation, as shown in Figure 1 (IIb), aggregates all events by making all edges time-independent. Consequently, an edge existing between a pair of nodes corresponds to an event between these nodes occurred at least once in the whole observed period [16]. Such an edge representation throws away information on the temporal ordering of events, making it impossible to study dynamic processes in static networks.

The incremental network solution accumulates events only up to the current time  $t$  of analysis. The classical approach, used early in [17–19], views a dynamic network as a series of time-ordered sequences of static graphs (see Figure 1 (IIc)). More recently, this method was applied to modeling network and community evolution [20, 21]. The drawback of this approach is that it does not preserve the ordering of interactions within time slices. Applying a simple frequency-based aggregation creates a frequency-based (FQ) metric cf. Figure 1 (IIId). Taking into account only a given number of the most recent events leads to the recency-based (RC) model cf. Figure 1 (IIe). Both of these models are used here as baseline models.

Figure 1 (IIIf) shows an example of dynamic social network generated using the CogSNet model. All other social network models presented in Figure 1 ((IIa)–(IIe)) can also be represented by CogSNet by setting appropriate parameter combinations to achieve, as needed, no decay, instant decay, and so forth. In this way, CogSNet can be thought of as a universal generative dynamic model for temporal social networks, encompassing previous approaches as special cases.

In general, the forgetting function  $f(\Delta t)$  over time interval  $\Delta t$  can be of any type (linear, power, and logarithmic), but here, informed by work in the cognitive psychology of memory [5], we evaluate only two such functions: the exponential function  $f^{\text{exp}}$  and the power function  $f^{\text{pow}}$  defined as follows:

$$f^{\text{exp}}(\Delta t) = e^{-\lambda \Delta t}, \quad (2)$$

$$f^{\text{pow}}(\Delta t) = \max(1, \Delta t)^{-\lambda}, \quad (3)$$

where  $\lambda$  denotes the forgetting intensity; typically,  $\lambda \in [0, 1]$ . The use of  $\max$  in the power function ensures that perception of events that happened less than a time unit ago is not changed by forgetting. The time unit in which the forgetting function is expressed scales the values of the parameters. Our experiments use one hour as the time unit.

To simplify optimal parameters search, we aggregate all three parameters into the *trace lifetime*  $L$  defining the time after which an unreinforced memory trace is forgotten, i.e., is too weak to be recalled. In the model,  $L$  is the time over which the forgetting function reduces the edge weight from  $\mu$  to  $\theta$  causing the edge to be removed cf. Figure 2. For the exponential forgetting function (equation (2)), trace lifetime  $L^{\text{exp}}$  is

$$L^{\text{exp}} = \frac{1}{\lambda} \ln\left(\frac{\mu}{\theta}\right), \quad (4)$$

while for the power function, the formula is

$$L^{\text{pow}} = \left(\frac{\mu}{\theta}\right)^{1/\lambda}. \quad (5)$$

Please note that according to the CogSNet model the relationship weights belong to range  $[\theta; 1) \cup \{0\}$ . Edge's weight value tends to 1 when the node encounters large number of tightly successive contacts. The larger the number of events is and the closer to each other they are, the closer to 1 the weight is.

### 3. Results and Discussion

In this section, we assess the empirical validity of the propose model by estimating the accuracy with which the CogSNet model reproduces the dynamics of social relations in a dynamic social network. Our computational experiments use the NetSense dataset introduced in [22].

The data contain 7,096,844 human mobile phone communication events, including both calls and text messages. These are augmented by 578 surveys containing self-reports on top contacts. It is worth underlining that this dataset provides a unique opportunity to evaluate how the communication frequencies are recalled by the study participants because students reported their perception of the frequency—detailed information on how this information has been reported and how data were processed is provided in Appendix.

We use this dataset to study the evolution of two coupled social networks of university students. The first is a behavioral network representing interactions between individuals in the form of the records of their mobile calls and text messages. The second one has perceptual edges defined by the personal nominations. These nominations are based on students' perception of the corresponding relations as one of the top twenty most interacting peers in the surveys administered to participants. These surveys cover the first four semesters of the student's college experience (beginning of freshman year to the end of the sophomore year). We compare the list of nominations predicted from the CogSNet network model purely from the communication event data with the list of nominations collected in a given survey.

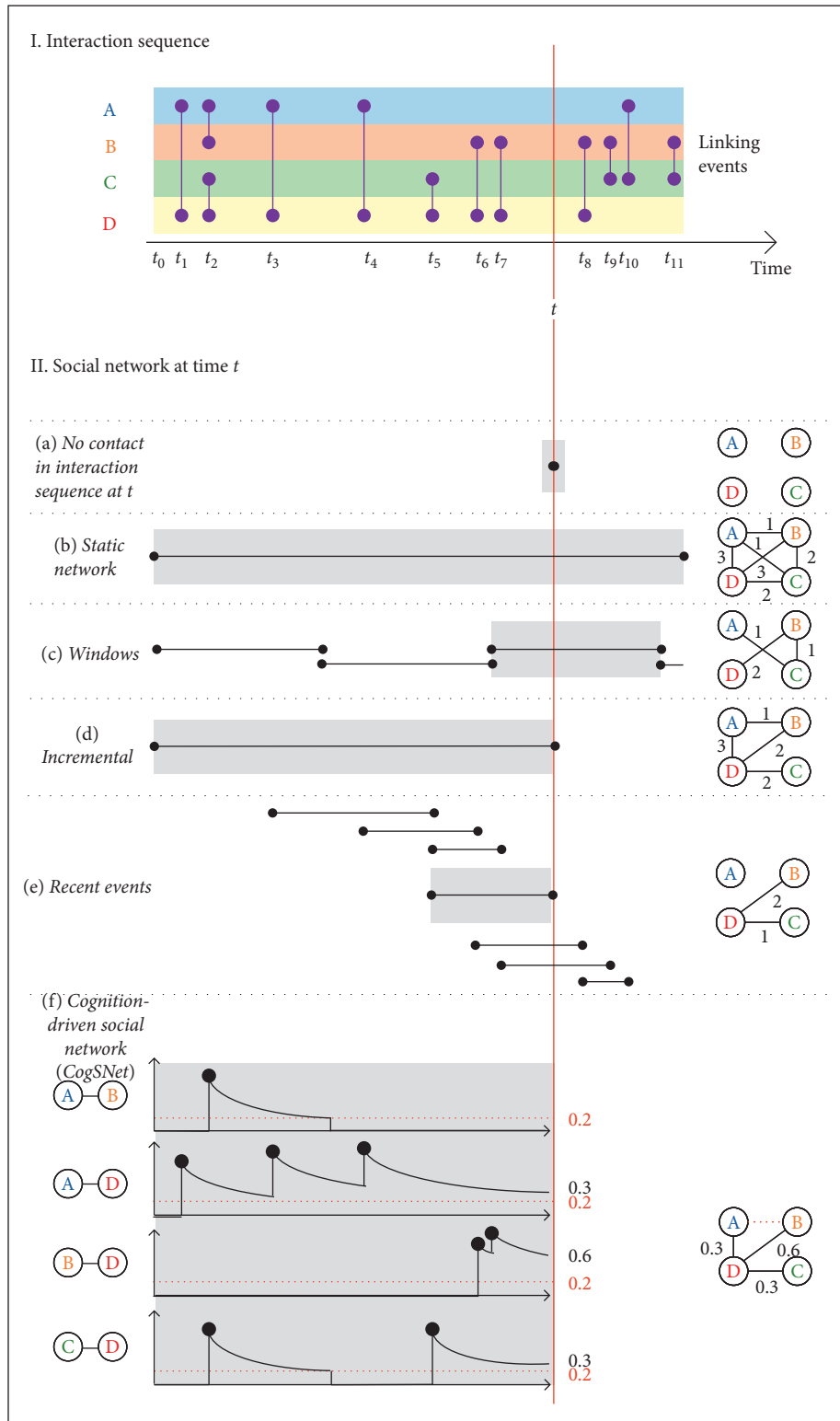


FIGURE 1: Various approaches to modeling dynamic networks and edge weighting (relation strengths) for the 4-node network at a given time  $t$ : (I) and (IIa) interaction sequence; (IIb) static (time-aggregated) network; (IIc) sliding windows; (IId) incremental network, all events from the beginning time  $t_0$  to the current time  $t$  are considered, and the frequency of interaction in the period constitutes the frequency-based (FQ) reference; (IIe) network based on  $n = 3$  recent events, used for recency-based reference RC; (IIff) cognition-based social network model (CogSNet), introduced here.



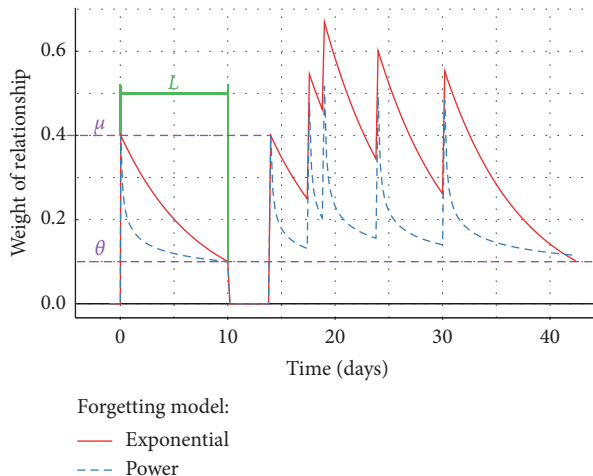


FIGURE 2: Dynamics of relations in CogSNet network with exponential and power functions and with parameters set to  $\mu = 0.4$ ,  $\theta = 0.1$ , and  $L = 10$  days.

Figure 3 shows an example of dynamic social network generated from a subset of NetSense data using the CogSNet model.

For comparison, we also implement three baseline models using the mobile phone communication data. The first is a *frequency-based* (FQ) model which orders the peers by the number of interactions, regardless of their time order. Under this model, the highest ranking peers whose number equals the number of the nominations listed in the corresponding survey are selected on the basis of their communication frequency. The *recency-based* (RC) model selects the top interacting peers who for each individual given a number of recent events. Finally, the *random sampling* (RND) model creates the list by randomly selecting peers from those who are recorded in the history of interactions of a given node. While random baseline can be considered as a simplistic lower bound, the two aforementioned are actually regarded to be state of the art since they express two cognitive phenomena. The recency effect was uncovered by the experiments of Hermann Ebbinghaus on recall. The forgetting curve he proposed is based on studies in [23] and recently confirmed in [24]. In these papers, the authors demonstrate that the learning curve rises the chances for recalling at the beginning and at the end of the contact period, while the chances to recall events are lower in the middle of it. However, as the question in surveys mainly refers to the present, we focused on the recency instead of primacy. The second baseline, frequency-based, is an implementation of scientific results taken from the domain of lexical memory [25]. The authors of this study show that the frequently repeated words are being recalled faster. Unfortunately, to the best of our knowledge, there are no other approaches suitable to be used in similar tasks.

To compare the performance of all the models, we use a *Jaccard* metric, equation (C.1), in Appendix, which measures the ratio of the number of nominations produced by the model that are also ground truth nominations listed in the corresponding survey divided by the number of unique nominations on both lists. Due to the nature of the question formulation in the survey, it was impossible to use any approach based on ordering such as Kendall's or Pearson's ranks since we had to consider the list as unordered.

Figure 4 shows the results of this comparison over the range of parameters corresponding to reported values for forgetting of one to 14 weeks. As reported in [26], the ability to recall information about social interactions starts to degrade after about one week. The experiments using NetSense dataset reveal that the performance is the highest when the forgetting of unreinforced memory traces happens after two weeks. The results remain satisfactory for forgetting thresholds as low as one week. With the two-week threshold, the CogSNet model is with either power or exponential forgetting statistically significantly better than any baseline model. In the literature, the power forgetting was found to perform significantly better than the exponential forgetting when a range of parameters was limited [3]. We do not observe such superior performance of power forgetting here. Both functions have similar peak of Jaccard metric, albeit for the different lifetime values.

When comparing the results of surveys with the states of the CogSNet network at the times of the surveys, the Jaccard metric is as high as 29.98% for exponential function for  $L = 3$  days. The distant second is recency-based RC model which delivers much lower Jaccard metric of 17.8%. Nevertheless, when looking at relative simplicity of the RC metric, one can also say that it performs quite good—this is due to the fact

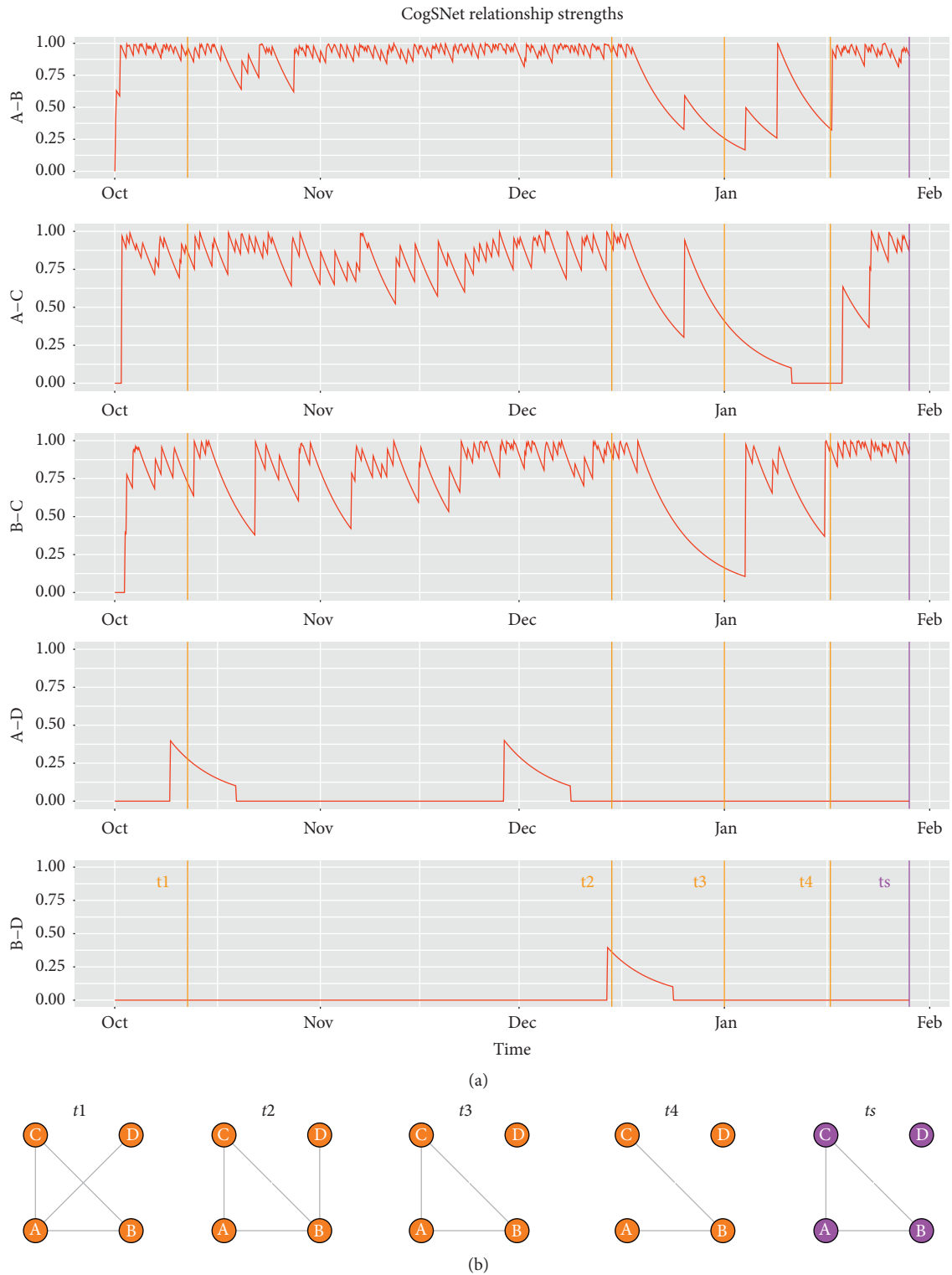


FIGURE 3: Four-node CogSNet network for the sample of real NetSense data:  $\mu = 0.4$ ,  $\theta = 0.1$ , and  $L = 10$  days; nodes A, B, C, and D correspond, respectively, to participants with ids 40997, 11360, 10841, and 1232. (a) Relation strengths according to the CogSNet model over 4-month period (one term); (b) network snapshots at four timestamps  $t_1$ – $t_4$  and at the survey time  $t_s$ .

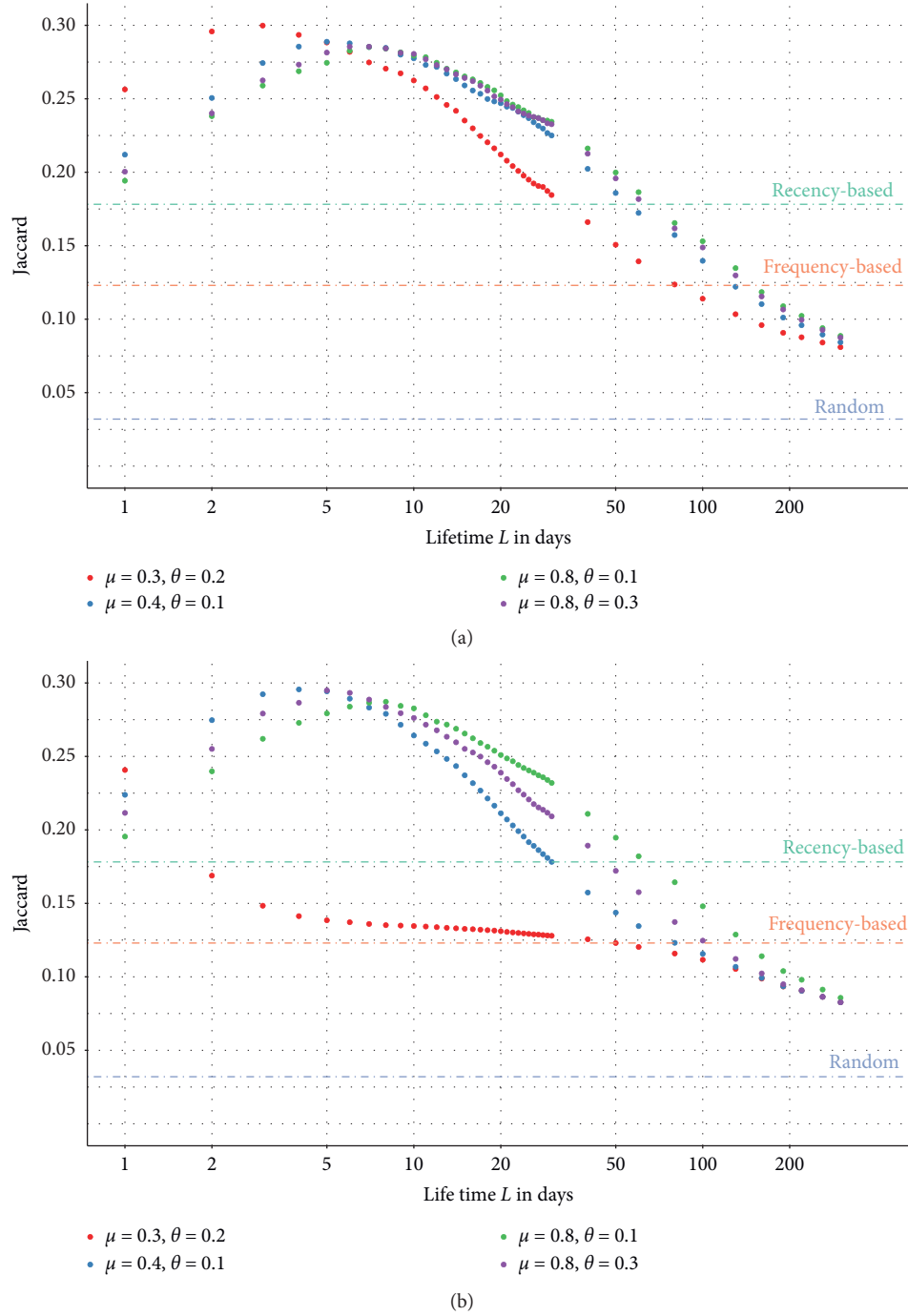


FIGURE 4: The plot of the Jaccard metric as a function of lifetime  $L$ . The metric measures the overlap between the two sets of peers, one identified by CogSNet and the other listed by students in the survey. These plots are compared to the results achieved by the three baseline models: recency-based with the best results obtained with the number of recent events set to 400, frequency-based, and random. The results are plotted for the CogSNet running with various parameters for (a) exponential and (b) power forgetting functions.

TABLE 1: Numerical and symbolic  $p$  values adjusted by Simes–Hochberg step-up method of nonparametric pairwise post hoc Nemenyi test comparing values of Jaccard metric for results obtained by each method with parameters fitted using recorded event data with the ground truth based on surveys taken over semesters 1–4.

	Power	Exponential	Recency	Frequency	Random
Power	—	***	***	***	***
Exponential	0.0000	—	***	***	***
Recency	0.0000	0.0000	—	***	***
Frequency	0.0000	0.0000	0.0000	—	***
Random	0.0000	0.0000	0.0000	0.0000	—

These results show how each method in the leftmost column performs versus the method in the first row. The part of table under the diagonal shows numerical values of the test, while the part above the diagonal presents  $p$  values coded as follows: \*\*\*  $p < 0.00005$  represents high confidence in the performance differences between the methods since typically values of  $p > 0.05$  are considered indicative of nonsignificance.

TABLE 2: Setting of parameters for power forgetting function in computational experiments.

$\lambda$ dependence on $\mu, \theta$ , and $L$ for power forgetting			
Lifetime $L$ days	$\mu = 0.4, \theta = 0.1$	$\mu = 0.8, \theta = 0.3$	$\mu = 0.8, \theta = 0.1$
1	0.43621	0.30863	0.65431
2	0.3581	0.25337	0.53716
3	0.32415	0.22934	0.48623
4	0.30372	0.21489	0.45558
5	0.28957	0.20487	0.43435
6	0.27894	0.19736	0.41841
7	0.27055	0.19142	0.40583
8	0.26368	0.18656	0.39552
9	0.2579	0.18247	0.38685
10	0.25294	0.17896	0.37942
11	0.24862	0.1759	0.37293
12	0.2448	0.1732	0.3672
13	0.24139	0.17079	0.36208
14	0.23831	0.16861	0.35747
15	0.23552	0.16663	0.35328
16	0.23297	0.16483	0.34945
17	0.23062	0.16317	0.34592
18	0.22844	0.16163	0.34267
19	0.22643	0.1602	0.33964
20	0.22455	0.15887	0.33682
21	0.22278	0.15762	0.33418
22	0.22113	0.15645	0.3317
23	0.21957	0.15535	0.32936
24	0.2181	0.15431	0.32716
25	0.21671	0.15333	0.32507
26	0.21539	0.15239	0.32309
27	0.21414	0.15151	0.3212
28	0.21294	0.15066	0.31941
29	0.2118	0.14985	0.3177
30	0.21071	0.14908	0.31606
40	0.20188	0.14283	0.30282
50	0.19553	0.13834	0.29329
60	0.19062	0.13487	0.28594
80	0.18337	0.12974	0.27506
100	0.17811	0.12602	0.26717

TABLE 3: Setting of parameters for exponential forgetting function in computational experiments.

$\lambda$ dependence on $\mu, \theta$ , and $L$ for exponential forgetting			
Lifetime $L$ (days)	$\mu = 0.4, \theta = 0.1$	$\mu = 0.8, \theta = 0.3$	$\mu = 0.8, \theta = 0.1$
1	0.05776	0.04087	0.08664
2	0.02888	0.02043	0.04332
3	0.01925	0.01362	0.02888
4	0.01444	0.01022	0.02166
5	0.01155	0.00817	0.01733
6	0.00963	0.00681	0.01444
7	0.00825	0.00584	0.01238
8	0.00722	0.00511	0.01083
9	0.00642	0.00454	0.00963
10	0.00578	0.00409	0.00866
11	0.00525	0.00372	0.00788
12	0.00481	0.00341	0.00722
13	0.00444	0.00314	0.00666
14	0.00413	0.00292	0.00619
15	0.00385	0.00272	0.00578
16	0.00361	0.00255	0.00542
17	0.0034	0.0024	0.0051
18	0.00321	0.00227	0.00481
19	0.00304	0.00215	0.00456
20	0.00289	0.00204	0.00433
21	0.00275	0.00195	0.00413
22	0.00263	0.00186	0.00394
23	0.00251	0.00178	0.00377
24	0.00241	0.0017	0.00361
25	0.00231	0.00163	0.00347
26	0.00222	0.00157	0.00333
27	0.00214	0.00151	0.00321
28	0.00206	0.00146	0.00309
29	0.00199	0.00141	0.00299
30	0.00193	0.00136	0.00289
40	0.00144	0.00102	0.00217
50	0.00116	0.00082	0.00173
60	0.00096	0.00068	0.00144
80	0.00072	0.00051	0.00108
100	0.00058	0.00041	0.00087

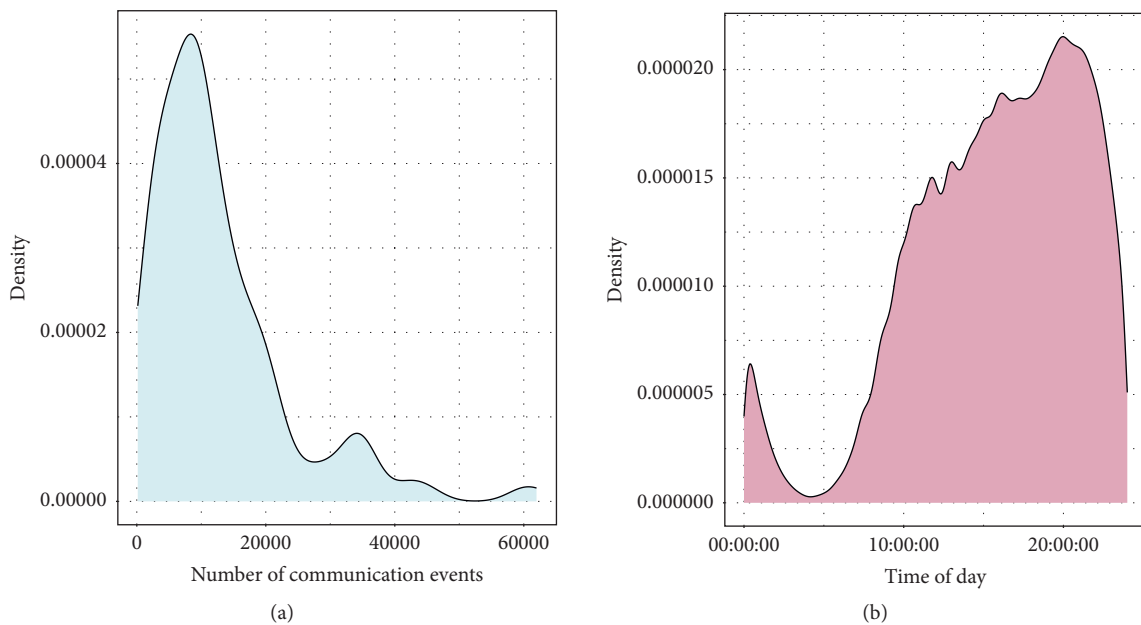


FIGURE 5: Continued.

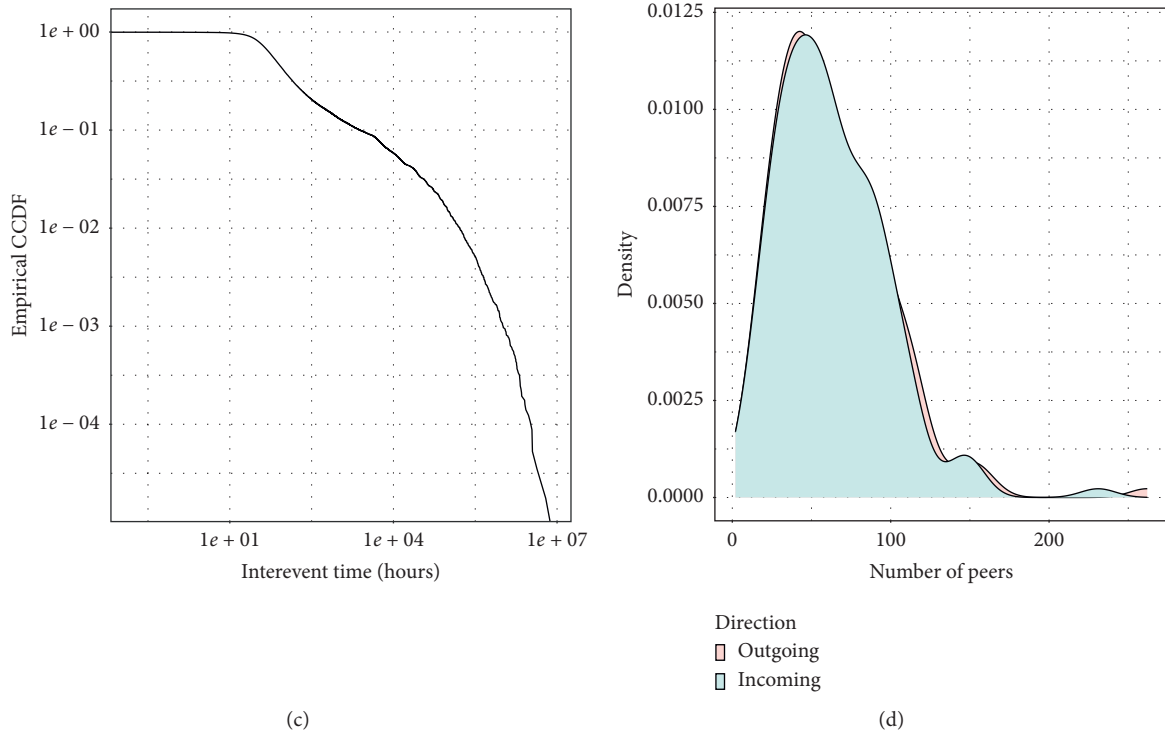


FIGURE 5: The exploratory analysis of the communication data for students participating in the NetSense study at the University of Notre Dame: (a) density plot for the number of communication events for the students; (b) the distribution of phone activity through the course of the day; (c) the cumulative distribution function of the interevent times in hours; (d) the number of incoming and outgoing communication peers for the students in NetSense, where there are at least twenty communication events with the peer during the analyzed period [30].

that primacy and recency effects come into play when listing peers [27].

#### 4. Conclusions

The results demonstrate that accounting for event perception and memory dynamics is essential for faithfully modeling the interaction dynamics in social networks. The CogSNet performed statistically significantly better than any of the base temporal models (see Table 1 in Appendix). With suitable data, the model can be extended to represent individual sensitivity to events and the impact of emotions and distractions on event perception to further improve modeling quality.

The dataset on which we operate contains data limited to a relatively small fraction of human social activities. Hence, it cannot be stated that there were no other events happening that could have affected participant memory about others. Examples of such events are face to face interactions, indirect references to someone else, and personal reminiscences about others. In addition, emotions can also play a significant role in these processes, and they can be potentially identified with some confidence from available data [28].

As a result, the reinforcement peak value can be adjusted for a given person and for an individual event cf. also equation (A.1) in Appendix. What is observed here is a partial manifestation from human memory. However, even taking into account that the model was built based on a

single data source, over 6 million telephone calls and messages among the NetSense study participants, the model fitted with the recorded event data yields results well matching the salience of social contacts over all 578 surveys completed by 184 participants. It is observed that even when using the same values of model parameters for all participants, the results are satisfactory. Nevertheless, this accuracy most likely could have been increased if the parameters had been individually adjusted for each participant. Apart from that, one should remember that the recall task is highly contextual, so the individuals will provide different sets of peers depending on the situation [11, 12].

The findings of this work can be also considered as an important building block on how to maintain healthy relationships or how to discover anomalies. For instance, based on the results, it is anticipated that, in order to keep a living relationship with a peer, the contacts should not be sparser than lifetime  $L$ , preferably even more frequent. On the contrary, these findings also demonstrate on how long a grieving period after breaking the relationship could potentially last and how to handle the recovery process afterwards in order to make it effective. Lastly, small modifications making the model closer to ACT-R would enable its use for instance in the case of Alzheimer disease where one could monitor whether the peers recalled by patients are in line with what a CogSNet model generates based on the contacts or visits. If there will be any discrepancy found between the model and observations based

on interview, this could be an early sign or disease emerging or becoming more active. Another interesting case would be to see whether the strengths of new links are decaying slower than the old ones.

In future work, we plan to extend the model in the directions outlined in Appendix including accounting for distractions during interactions, individualized strength, asymmetric of interactions of significance to participants (e.g., hierarchical relationships), and the impact of forms of interactions and of associated emotions. Hence, the CogSNet model represents an important first step towards modeling social network dynamics through the prism of human cognition. In literature, there are three approaches that are used widely to understand human cognition: cognitive bandwidth (CB), dual-process morality (DPM), and implicit association tests (IAT). In [29], those methods are compared and their common drawback is pointed out—the lack of perspective that takes a cultural background into account. Our approach makes such considerations possible, and we see this challenge as one of the future work directions. Lastly, we would love to see more datasets that are so extensive as the NetSense dataset that could allow us to evaluate our model.

$$w_{ij}(t) = \begin{cases} \mu_{ijc_{ij}+1}, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ \mu_{ijc_{ij}+1} + w_{ij}(t_{ij})f(t - t_{ij})(1 - \mu_{ijc_{ij}+1}), & \text{otherwise,} \end{cases} \quad (\text{A.1})$$

where  $\mu_{ijk}$  is the value of reinforcement peak that results from the  $k^{\text{th}}$  event that impacts the edge  $(v_i, v_j)$ . Here, the value of reinforcement peak  $\mu_{ijk}$  depends on the engagement and emotions invoked by the event that is either directly or indirectly related to the edge  $(v_i, v_j)$ . An example of an event indirectly related to this edge could be node  $v_i$  talking about node  $v_j$  or any situation that reminds node  $v_i$  about node  $v_j$ . The values of  $\mu$  can be individualized to node  $v_i$  perception of relation with node  $v_j$  at event  $k$ . The  $\mu$ 's values may also be dependent on event types:  $\mu_{ijk} \in \{\mu_1, \mu_2, \mu_3, \mu_4, \dots\}$ , e.g.,  $\mu_1 = 0.5$  for emails,  $\mu_2 = 0.55$  for phone calls,  $\mu_3 = 0.8$  for meetings, and  $\mu_4 = 0.9$  for joint collaboration in projects.

Finally, the processing of the current event updates both variables associated with the updated edge  $(v_i, v_j)$  as follows:  $t_{ij} = t, c_{ij} = c_{ij} + 1$ .

At any time  $t$  of the model evolution, the user can obtain the value of the weight of an arbitrary edge  $(v_i, v_j)$  by computing the following equation:

$$w_{ij}(t) = \begin{cases} 0, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ w_{ij}(t_{ij})f(t - t_{ij}), & \text{otherwise.} \end{cases} \quad (\text{A.2})$$

## Appendix

### (A) Formal Definition of the CogSNet Model

A social network can be represented by a graph  $G = (V, E)$ , where  $V = \{v_1, \dots, v_n\}$  denotes the set of  $n$  nodes and  $E = \{e_1, \dots, e_m\}$  is the set of directed  $m$  edges between pairs of nodes. Each edge  $e_{ij}$  from node  $v_i$  to node  $v_j, i \neq j$ , is assigned weight  $w_{ij}(t)$  and represented by a triple  $(v_i, v_j, w_{ij}(t))$ .

The model evolves in discrete steps as follows. For each pair of nodes  $(v_i, v_j)$ , the system maintains two variables:  $t_{ij}$ , which represents the time of the most recent event for this pair of nodes and  $c_{ij}$  which holds the count of events processed for this pair of nodes. Initially, both  $t_{ij}$  and  $c_{ij}$  are set to 0, as are the weights of all edges, i.e., for all pairs of nodes  $(v_i, v_j)$ ,  $w_{ij}(0) = 0$ .

When an event happens at time  $t$  in the modeled social network, it is processed in chronological order by the model. First, the weight of the corresponding edge is updated according to the following equation:

### (B) Experimental Parameter Space

To analyze the CogSNet model and compare its results to those obtained with the reference models, many combinations of the CogSNet parameters were tested. First, different values of  $\mu$ ,  $\theta$ , and  $L$  were specified, and then,  $\lambda$  coefficient was calculated using equations (4) and (5), which were also used to define CogSNet parameters for experimental verification. The complete list of parameters used is listed in Tables 2 and 3.

### (C) Quality Measures

The values of Jaccard metric for a single surveyed student participant  $v_i$  have been computed as follows:

$$\text{Jaccard}(v_i) = \frac{|V_i^{\text{CogSNet}} \cap V_i^{\text{survey}}|}{|V_i^{\text{CogSNet}} \cup V_i^{\text{survey}}|}, \quad (\text{C.1})$$

where  $V_i^{\text{survey}}$  is the set of up to 20 peers enumerated in the survey by the participating student  $v_i$  and  $V_i^{\text{CogSNet}}$  is the set of neighbors of this student in the CogSNet network with the

largest nonzero weights on edges to this student on the day on which the given survey was administered.

### (D) Recency-Based Baseline

The recency-based baseline model was tuned by finding the best parameter representing the number of the most recent events to be taken into account in computing the performance. The resulting value of 400 events was used in Figure 4 for computing the results for the recency-based baseline model.

### (E) Statistical Tests

To confirm that there is a statistical difference between the proposed CogSNet model and baselines, we performed a number of statistical tests. The results are presented in Table 1 which shows that CogSNet approach is statistically significantly better from all baseline methods. Moreover, power forgetting performs statistically significantly better than exponential forgetting.

### (F) NetSense Dataset

The dataset consists of two parts. The first includes the timestamps and duration/length of phone calls and text messages collected for each student participating in the study. Each student phone device recorded all connections/messages, including those to the phones of people outside of the test group, so recording was done on both sides of communication, by sending and receiving calls/messages, if both belonged to the study. In total, the mobile phone communication data consist of 7,575,865 activities. Out of these, 7,096,844 have been text messages (93.7% of all events) and 479,021 phone calls (6.3%). Due to the fact that the dataset covers millions of mobile phone interactions that have been recorded by a special application installed on students' phones, some issues discussed below arose during the preliminary data analysis phase.

Since events involving each student possessing the mobile phone were recorded independently from other mobile phones, some inconsistencies arose among which the most important was that not in all cases, the event was recorded on both sides. Since usually a couple of seconds pass between sending text message and receiving it by the recipients, we identified the same message by its length. However, sometimes, the recorded lengths were different, so in such cases, we were forced to make somewhat arbitrary decision based on the length difference whether the message was the same or separate one for which the corresponding event on the other side of communication was not recorded.

Next, when the students phoned each other and the call was not answered, the recipient's voice mail was reached, if it was enabled. Such case has been recorded as a regular phone call even if it has not been actually answered by the recipient. In order to solve this problem, we had to look for orphan records that do not have their corresponding records on the recipient's side.

This part of data preprocessing removed 12,482 records from our dataset. Afterwards, we performed an exploratory analysis of events. In Figure 5, we show the basic statistics regarding the communication patterns of the participants of the study.

The second part of the dataset includes surveys containing peers enumerated by the participants at the end of each term in response to the following question: "*In the spaces below, please list up to 20 people (friends, family members, acquaintances, or other people) with whom you spend time communicating or interacting.*" To be more specific on how the process of obtaining data looked like, the students have been presented with a form containing twenty empty fields with the request in question—they had to fill the names by themselves without any suggestions or generators. As already mentioned in the main part of the manuscript, the ordering of the enumerated peers was not imposed, but even if the students would start with listing the most important and not necessary most frequent contacts, we could not assume that.

Lastly, the students were asked to enumerate in any order up to twenty peers with whom they interacted recently. However, the system recorded each name with the accompanying number representing name's position on the list. For some surveys, there were some holes in this numbering, so either the student did not list the peers accordingly (left some blank lines in between names) or partial results were not recorded. This is why we consider listed peers as the sets without any implied ordering, and the Jaccard sets comparison metric was used for evaluation purposes instead of any other that are based on ordering, e.g., Pearson or Spearman rank correlation coefficients.

### Data Availability

The code for computing the state of the social network built using the CogSNet model has been made public. A fast C implementation prepared by the authors of this manuscript is available as a Code Ocean capsule (<https://doi.org/10.24433/CO.6088785.v1>). The NetSense dataset used for the model validation is available for other researchers; the readers interested in accessing the data should contact the coauthor Professor Omar Lizardo.

### Conflicts of Interest

The authors declare that there are no conflicts of interest.

### Authors' Contributions

All authors participated in the design of the research and computational experiments. RM and MK ran the computational experiments and collected the results. All authors participated in analysis of the results. All authors wrote and edited the manuscript.

### Acknowledgments

The authors would like to thank Prof. Tomasz Kajdanowicz for his valuable remarks regarding statistical tests. This work



was partially supported by the National Science Centre, Poland, by the following projects: 2015/17/D/ST6/04046 (RM) and 2016/21/B/ST6/01463 (PK, MK), Office of Naval Research (ONR) (Grant no. N00014-15-1-2640), Army Research Office (ARO) (Grant no. W911NF-16-1-0524), European Union's Marie Skłodowska-Curie Program (Grant no. 691152), and Polish Ministry of Science and Higher Education (Grant no. 3628/H2020/2016/2). Calculations have been carried out using resources provided by Wrocław Centre for Networking and Supercomputing (<https://wcss.pl>) (Grant no. 177). Collection of the NetSense data was supported by the U.S. National Science Foundation (Grant no. IIS-0968529). Further support for data collection was provided through University of Notre Dame and Sprint.

## References

- [1] D. B. Alan, "The psychology of memory," *The Essential Handbook of Memory Disorders for Clinicians*, pp. 1–13, John Wiley & Sons, Hoboken, NJ, USA, 2004.
- [2] P. R. Huttenlocher, *Neural Plasticity: The Effects of Environment on the Development of the Cerebral Cortex*, Harvard University Press, Cambridge, MA, USA, 2009.
- [3] J. R. Anderson, D. Bothell, C. Lebiere, and M. Matessa, "An integrated theory of list memory," *Journal of Memory and Language*, vol. 38, no. 4, pp. 341–380, 1998.
- [4] R. S. Burt, "Decay functions," *Social Networks*, vol. 22, no. 1, pp. 1–28, 2000.
- [5] P. Jenkins, G. Earle-Richardson, D. T. Slingerland, and J. May, "Time dependent memory decay," *American Journal of Industrial Medicine*, vol. 41, no. 2, pp. 98–101, 2002.
- [6] J. T. Wixted and E. B. Ebbesen, "On the form of forgetting," *Psychological Science*, vol. 2, no. 6, pp. 409–415, 1991.
- [7] D. A. Norman, *Memory and Attention: An Introduction to Human Information Processing*, John Wiley & Sons, Hoboken, NJ, USA, 1976.
- [8] T. Raeder, O. Lizardo, D. Hachen, and N. V. Chawla, "Predictors of short-term decay of cell phone contacts in a large scale communication network," *Social Networks*, vol. 33, no. 4, pp. 245–257, 2011.
- [9] P. McLean, *Culture in Networks*, John Wiley & Sons, Hoboken, NJ, USA, 2016.
- [10] D. C. Bell, B. Belli-McQueen, and A. Haider, "Partner naming and forgetting: recall of network members," *Social Networks*, vol. 29, no. 2, pp. 279–299, 2007.
- [11] M. E. Brashears, E. Hoagland, and E. Quintane, "Sex and network recall accuracy," *Social Networks*, vol. 44, pp. 74–84, 2016.
- [12] M. E. Brashears and E. Quintane, "The microstructures of network recall: how social networks are encoded and represented in human memory," *Social Networks*, vol. 41, pp. 113–126, 2015.
- [13] J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang, "On graph problems in a semi-streaming model," in *Proceedings of the International Colloquium on Automata, Languages, and Programming*, pp. 531–543, Springer, Turku, Finland, July 2004.
- [14] S. Sreenivasan, K. S. Chan, A. Swami, G. Korniss, and B. K. Szymanski, "Information cascades in feed-based networks of users with limited attention," *IEEE Transactions on Network Science and Engineering*, vol. 4, no. 2, pp. 120–128, 2017.
- [15] P. Holme and J. Saramäki, "Temporal networks," *Physics Reports*, vol. 519, no. 3, pp. 97–125, 2012.
- [16] P. Holme, "Network dynamics of ongoing social relationships," *Europhysics Letters (EPL)*, vol. 64, no. 3, p. 427, 2003.
- [17] T. Y. Berger-Wolf and J. Saia, "A framework for analysis of dynamic social networks," in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 523–528, ACM, Philadelphia, PA, USA, August 2006.
- [18] G. Kossinets and D. J. Watts, "Empirical analysis of an evolving social network," *Science*, vol. 311, no. 5757, pp. 88–90, 2006.
- [19] J. Moody, "The importance of relationship timing for diffusion," *Social Forces*, vol. 81, no. 1, pp. 25–56, 2002.
- [20] V. Sekara, A. Stopczynski, and S. Lehmann, "Fundamental structures of dynamic social networks," *Proceedings of the National Academy of Sciences*, vol. 113, no. 36, pp. 9977–9982, 2016.
- [21] J. Xie, M. Chen, and K. S. Boleslaw, "LabelRankT: incremental community detection in dynamic networks via label propagation," in *Proceedings of the Workshop on Dynamic Networks Management and Mining*, pp. 25–32, ACM, New York, NY, USA, May 2013.
- [22] A. Striegel, S. Liu, L. Meng, C. Poellabauer, D. Hachen, and O. Lizardo, "Lessons learned from the NetSense smartphone study," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 51–56, 2013.
- [23] H. Ebbinghaus, *Über Das Gedächtnis: Untersuchungen Zur Experimentellen Psychologie*, Duncker & Humblot, Berlin, Germany, 1885.
- [24] J. M. J. Murre and J. Dros, "Replication and analysis of ebbinghaus' forgetting curve," *PLoS One*, vol. 10, no. 7, Article ID e0120644, 2015.
- [25] D. L. Scarborough, C. Cortese, and H. S. Scarborough, "Frequency and repetition effects in lexical memory," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 3, no. 1, p. 1, 1977.
- [26] E. Nathan, A. S. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences*, vol. 106, no. 36, pp. 15274–15278, 2009.
- [27] N. Miller and D. T. Campbell, "Recency and primacy in persuasion as a function of the timing of speeches and measurements," *The Journal of Abnormal and Social Psychology*, vol. 59, no. 1, 1959.
- [28] R. A. Calvo and S. D'Mello, "Affect detection: an interdisciplinary review of models, methods, and their applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, 2010.
- [29] M. Lamont, L. Adler, B. Y. Park, and X. Xiang, "Bridging cultural sociology and cognitive psychology in three contemporary research programmes," *Nature Human Behaviour*, vol. 1, no. 12, p. 866, 2017.
- [30] M. Nurek, R. Michalski, and R. Marian-Andrei, "Hawkes-modeled telecommunication patterns reveal relationship dynamics and personality traits," 2020, <https://arxiv.org/abs/2009.02032>.

## Review Article

# A Review of the Research Progress of Social Network Structure

Ning Li <sup>1</sup>, Qian Huang <sup>1</sup>, Xiaoyu Ge <sup>2</sup>, Miao He <sup>1</sup>, Shuqin Cui <sup>3</sup>, Penglin Huang <sup>4</sup>,  
Shuairan Li <sup>1</sup> and Sai-Fu Fung <sup>5</sup>

<sup>1</sup>*Xi'an Physical Education University, Xi'an 710065, China*

<sup>2</sup>*Shanghai University of Sport, Shanghai 200444, China*

<sup>3</sup>*Sport College, Xi'an University of Architecture and Technology, Xi'an 710311, China*

<sup>4</sup>*Shandong Sport University, Jinan 250063, China*

<sup>5</sup>*City University of Hong Kong, Kowloon, Hong Kong*

Correspondence should be addressed to Qian Huang; [huangqian168@126.com](mailto:huangqian168@126.com)

Received 29 October 2020; Revised 30 November 2020; Accepted 27 December 2020; Published 8 January 2021

Academic Editor: Fei Xiong

Copyright © 2021 Ning 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.

Social network theory is an important paradigm of social structure research, which has been widely used in various fields of research. This paper reviews the development process and the latest progress of social network theory research and analyzes the research application of social network. In order to reveal the deep social structure, this paper analyzes the structure of social networks from three levels: microlevel, mesolevel, and macrolevel and reveals the origin, development, perfection, and latest achievements of complex network models. The regular graph model, P1 model, P2 model, exponential random graph model, small-world network model, and scale-free network model are introduced. In the end, the research on the social network structure is reviewed, and social support network and social discussion network are introduced, which are two important contents of social network research. At present, the research on social networks has been widely used in coauthor networks, citation networks, mobile social networks, enterprise knowledge management, and individual happiness, but there are few research studies on multilevel structure, dynamic research, complex network research, whole network research, and discussion network research. This provides space for future research on social networks.

## 1. Introduction

Social network is a form of social organization based on “network” (the interconnection between nodes) rather than “group” (clear boundary and order). It is an analytical perspective emerging from western sociology in the 1960s. Although “social networks” have been proposed for more than 60 years, social networks are still an important research method. Social network refers to the relatively stable relationship system formed between individual members of the society due to interaction. Social network focuses on the interaction and connection between people, and social interaction can affect people's social behavior.

This paper reviews the theoretical and empirical literature on social networks. Firstly, the network structure view, relationship strength theory, social capital theory, and structural hole theory of social network are discussed. Then,

the methods of social network analysis are discussed. Social network analysis is a new social science research paradigm based on social network theory, and it is an important method to study social structure. It is a tool in the theory of social networks to quantify the relationship between actors in social networks. As more and more scholars study social structure and think seriously about the “network structure” of social life, various network concepts (such as centrality, density, structural hole, and clustering coefficient) are crowding in. Then, it reviews the development of network theory: regular network, random network, and complex network. A large number of studies have proved that a complex network is more suitable to the characteristics of the real network. Social support and social discussion networks are the last part of this article; they study social network structure in the perspective of social support and social discussion. Social support network refers to a person

through mutual exchanges to maintain social status and get emotional support and material assistance. And social discussion networks define the network relationship by whether you discuss important issues with him.

## 2. Social Network Theory

Social network is a perspective of sociological analysis that emerged in the 1960s. At present, the main theories include strong ties and weak ties theory, social capital theory, and structural hole theory. After years of development, the social network theory has been widely used in the fields of computing science, economics, and information dissemination. And it has become an important theory and method for research.

Network refers to the relationship between things, and social network was first put forward by the famous British anthropologist Radcliffe Brown in his attention to structure. It focuses on how culture stipulates the behavior of internal members of a bounded group, while actual interpersonal behavior is much more complex. Barnes believed that the informal connection between individuals is the essence of the meaning of social networks [1]. A more mature definition of a social network is Wellman's idea in 1988 that "a social network is a relatively stable system composed of social relations among individuals," that is, he treats "the network" as a series of social ties or social relations linking actors, whose relatively stable patterns make up the social-structure [2]. With the expansion of application scope, the concept of the social network has exceeded the scope of interpersonal relationship. Network actors are no longer limited to individuals; it can also be collectives, such as families, departments, organizations, institutions, and countries. With the development and improvement of social network theory, social network, which is originated from sociology began to be applied in pedagogy, economics, and other fields. More and more scholars construct the concept of "social structure," and social network theory was gradually recognized by scholars [3].

"Social network" is a collection of social actors which is treated as nodes and their relationships [4]. Among them, the actor may be an individual, the organization, or the state. The relationship between actors is the specific content between two actors, including the relationship between actors (friendship and partnership), the interaction between actors (benefit-exchanging and information transmission), and the evaluation of actors (trust and respect) [5]. Some scholars think that people often communicate and form relationships through the network. But social network analysis studies the social phenomenon and social structure from the perspectives of the "relationship." It is a research method to study interpersonal relationships and interpersonal interaction among group members. As a kind of quantitative analysis tool, it has strong analytical and graphical characteristics [6]. Social network analysis is different from traditional research, which starts from the relationship between individuals and analyzes the interaction and influence among network members, while traditional research mainly focuses on the attributes of individuals [7].

The research of social networks began in the 1960s. By the mid-1970s, the social network becomes a new field of sociology. In recent years, the social network research is widely used in computer science, mathematics, education, neuroscience, economics, and other fields with many scholars' contributions. A large number of relevant studies use social network analysis methods to analyze social phenomena and social problems, such as coauthor network and citation network, mobile social network, enterprise knowledge management, career mobility, individual happiness loneliness, group decision making, urban sociology, mental health, and gerontology. There are some representative social network theories: the perspective of network structure, relationship strength theory, social capital theory, and structural holes theory [8].

Network structure view and status structure view are completely two different research paradigms. According to the status structure view, individuals in the society have certain attributes, and we can classify them by the attributes of individual members. Also, their social behaviors are explained according to the categories they belong to. However, the network structure view holds that the link between any subjects (person or organization) and other subjects is an objective social structure, which will have an impact on the subject's behavior [9]. White, Granovetter, and some others are also supporters of the view of network structure. They oppose the idea that individuals are classified by their categorical attributes and using their attributes to explain their actions. Status structure view and network structure view have the following several differences: firstly, the status structure view emphasizes the individual attribute (age, gender, class, and status), and the individual is sorted by their characteristics. But network structure theory emphasizes the link between individual and individual (family and friends); it is classified by individual social networks. Secondly, the status structure view focuses on the individual's identity and sense of ownership, while the network structure view focuses on the individual's social embeddedness. Thirdly, the status structure view emphasizes whether an individual possesses resources or not and how much resources he possesses and attribute it to the status of an individual, while the network structure view holds that the position of an individual in the network and the scale of the network determines the rules of the social network. The two views seem to be opposite, but in fact, they complement each other [10].

The concept of tie power was first proposed by American sociologist Granovetter in his article "The power of Weak tie" published in 1973, and the ties were divided into strong ties and weak ties. He believes that the strength of the ties determines the nature of the available information and the likelihood that an individual will achieve his goal. In terms of the measurement of the strength of the tie, most studies refer to the four dimensions classified by Granovetter: interaction frequency, the emotional strength, the intimacy, and the reciprocal exchange behavior [11]. Strong tie refers to the strong homogeneity of the individual social network (that is, the work that the communicating group is engaged in and the information they grasp is similar), the close relationship

between people, and there is a strong emotional factor maintaining the interpersonal relationship. Weak tie is characterized by a strong heterogeneity of individual social network (i.e., a wide range of contacts, contacts may come from all walks of life, so the information available is also multifaceted), and interpersonal relationship is not that tight, and there is not too much emotional maintenance [12]. Although Granovetter pointed out the theory of strong and weak ties, he focused on elaborating and developing the theory of weak ties accordant to American society. Bian and Zhang, a Chinese scholar, developed the theory of strong tie based on China's actual situation and defined and developed relational social capital. He viewed social phenomena and social problems from the perspective of social relations, not limited to the perspective of the social network [13].

In the 1980s, Bourdieu explained social capital. He believed that social capital is a collection of actual resources or potential resources, which are related to the network of mutual acquaintance and recognition. In other words, it is related to the identity of group members. The amount of social capital is affected by an individual's economic status and social status, and the amount of social capital an individual has is usually related to his or her relationship network, economic status, and cultural quality. The definition of social capital concept by Lin what absorbing the concept of Marx's capital, Schultz's human capital concept, as well as Bourdieu, Coleman, and Putnam was most comprehensive. It emphasizes that the social capital is a kind of source invested in social relations and for acquiring the harvest in the market, embed in social structures, and obtain resources by a purposeful action [14]. When Lin defines social capital, he emphasizes the preexistence of social capital, which exists in a certain social structure, and people must follow the rules in order to obtain the social capital. Meanwhile, the definition also explains the initiative of human action, and people can obtain social capital through purposeful actions. Lin Nan believed that the theoretical model of social capital should include three processes: firstly, investment in social capital; secondly, the involvement and mobilization of social capital; the third is the return of social capital. He divided social action into instrumental action and emotional action. Instrumental action is understood as obtaining resources not owned by the actor, while emotional action is understood as maintaining resources already owned by the actor [14]. Existing literature has conducted in-depth research on social capital theory from the perspectives of basic concepts, measurement methods, and indicators as well as social and economic performance. The development of social capital has been relatively perfect. In recent years, researches have focused on the combination of social network theory and specific problems. For example, Cook and Hole investigated the survival of homeless people in the street from the perspective of Bourdieu's social capital theory [15]. Valente and Pitts analyzed and evaluated the social network theory and applied it to public health problems [16]. Chinese scholars Zhou et al. studied the online participation behavior of health community users based on the social capital theory [17]. Zhang and Wang studied the reading

interest relationship network in public libraries from the perspective of social network theory [18].

In 1992, Ronald Burt proposed the "structural hole theory" in *Structural Holes: Competitive Social Structures*. The "structural hole" refers to the gap in the social network, that is, some individuals in the social network have direct contact with some individuals, but they do not have direct contact with other individuals and there is no direct relationship. So it seems that there is a hole in the network structure from the whole perspective of the network [19]. Burt pointed out that structural holes in social networks not only have greater access to nonrepetitive resources but also can control resource flows between a group of nodes connected by structural holes, thus bringing greater benefits to participants located in the competitive field [20]. Structural hole theory is a new member of the family of interpersonal network theory, which emphasizes that the structural holes existing in the interpersonal network can bring advantages in information and other resources to organizations and individuals in this position. Compared with other theories, it emphasizes the structural analysis and the use of the interpersonal network, which is easier to grasp and operate.

With the development in the past almost 70 years, scholars have made an in-depth analysis of the application of social networks in practical fields. Social network, as a unique perspective, provides the academic research of scholars with a unique perspective. It plays a guiding role to some degree in almost each field of human behavior and has been widely recognized by scholars all over the world. The vitality of social network theory lies in its profound theoretical foundation of relational sociology and its scientific and empirical research. From the perspective of studying Chinese society, ties sociology is a set of theoretical knowledge with the main characteristics of ethics-oriented and relationship-oriented. From the perspective of methodology, relational sociology is a kind of mindset and research methodology used to explore and analyze social behavior patterns, or it can be said that it is a knowledge to study various social types from the perspective of relational theory [21].

Relational sociology is theory-oriented empirical research of observation, measurement, and grasp of human behavior, also the combination of qualitative research and quantitative research. Based on the analysis of individuals and their relationships, social network theory emphasizes the initiative of individuals and pays attention to the social structure formed by social networks and the restriction of social systems on individuals, and the interaction between individuals may change the social structure that restricts individuals [22]. Social network theory also builds a bridge between microbehavior and macrophenomena. Society is made up of a variety of microcosmic individuals, all of whom interact to produce a macroscopic picture. In the process of individual interaction, the behavior of individuals influences the social structure through the relationship, and the social structure in turn influences the decision-making and behavior of individuals. Social networks are the organic integration of individuals and their interrelationships [23].

*2.1. Social Network Structure Analysis.* Social network analysis is a new paradigm of social science research and methods. It is different from traditional sociology research who conducts the research in the perspective of individual characteristics and is used to measure the extent of the individual accepted by the group and organization. It is a tool to discover the existing relationships between people in a group and reveal the structural characteristics of the organization itself. It is mainly used to evaluate and measure the relationships of interpersonal attraction or interpersonal rejection in a social group, as well as a method to study interpersonal selection, information exchange, and interactive relationships within a group [24]. Some scholars believe that society is a huge network composed of various relationships [25]. The structure of the social network and its influence mode on social behavior are the research objects of social network. The purpose of studying social network is to reveal the deep social structure, that is, the fixed network mode hidden under the complex appearance of the social system.

Since the 1960s, social network research has been developed along with the two directions that differ and connect mutually. One is the egocentric network analysis, the other is the whole network analysis. The former analyzes the relationship between individual/node attributes and individual networks, while the latter mainly studies the network structure, the properties of graph theory, and location attributes [26]. Egocentric network analysis defines social networks from the perspective of individuals. Their concern is how individual behavior reacts under the influence of the interpersonal network and how to form a society through the network [27]. Representatives of this field include Mark Granoveter, Harrison White, Lin Nan, and Ronald Burt. Individual network measure includes scale, type of relationship, network density, pattern of relationship, homogeneity, and heterogeneity of members. Whole network analysis studies the integrated structure of the whole network and the internal relations of a small group. A series of network analysis concepts and indicators, such as network density, group centrality and faction, are generated by using the matrix method to analyze interpersonal interaction and exchange patterns [28]. The representative of this field is Linton Freeman.

The individual network locates the network by regarding the individual as the center. The resources in the individual network affect the behavior of the individual through the network structure. The research on the individual network mainly focuses on the attributes of the individual and the explanation of the network's influence on the individual's concept and behavior. In theory, we regarded individuals as part of the social structure, analyzing the influence of individual network characteristics on individual characteristics changes, and the results of individual ideas and behaviors affected by the network [29].

Whole Network is made up of a group of specific individuals and their mutual relations, and network members have relatively obvious boundaries [30]. The research on the whole network is also based on the theory of social network, taking the individual as a part of the social structure, and

analyzing the influence of the network on the individual. The whole network takes the whole network as the direct research object, rather than the specific individuals. It uses graph theory as tools in combination with other methods and focuses on the group internal relationship. The patterns of interpersonal communication and social interaction are analyzed by network density, group centrality, faction, and other network characteristics. By analyzing the characteristics of the whole network, scholars explore the whole characteristics of individuals in the network after being affected by the network. The whole characteristics explain the individual of the whole network being constantly influenced by other individuals and finally completes the embeddedness of social environment and social structure. Whole network research can have a relatively comprehensive study of the whole network and can reveal the whole network of various structural characteristics. For example, the reciprocity index of the network can be calculated, the tripartite relationship structure of the network and the whole network structure can be revealed, the subdivision distribution in the whole network can be found, and the density of the network can be calculated. Obviously, this kind of analysis has surpassed individual network research [31]. As a matter of fact, the whole network is a research method based on dynamics, which not only studies the influence of individual ideas and behaviors by social network factors but also reflects the influence process of the social network.

By studying network relations, social structure studies combine individual relations, "micro" networks and "macro" structures of large-scale social systems. The microanalysis of the social structure focuses on the most basic social relations-role relations. Mesostructural analysis focuses on the relationship between social components, which is not reflected in individual activities. Macrostructural analysis studies the overall social structure as well as the functions and effects of the overall structure. The microstructure, mesostructure, and macrostructure analysis all study the social structure, but at different levels.

*2.1.1. Microstructure.* Centrality is one of the microstructure characteristics of social networks, and it is also one of the key points in the analysis of the microstructure of social networks. The measurement index of centrality includes degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality [32]. Degree centrality refers to the sum of the number of direct connections between a point and other points and is the most direct measure to depict node centrality in network analysis. The greater the degree of a node means the higher the degree centrality of the node, and the node in the network is more important [33]. Closeness centrality reflects the closeness between a node and other nodes in a network. The reciprocal of the sum of the shortest path distances of a node to other nodes represents the closeness centrality. That is, for a node, the closer it is to other nodes, the greater its closeness centrality is. Closeness centrality measures an indicator of an actor independent from other actors' control [34]. Betweenness centrality refers to the number of times a node serves as the

shortest bridge between the other two nodes. The more a node acts as a “Betweenness,” the greater its betweenness centrality. If a node has a higher betweenness centrality, it acts as a bridge between other individuals. If the betweenness centrality of a node is zero, it means that the node cannot control any actor and is on the edge of the network. If the middle centrality of a node is one, it can completely control other actors. It is at the core of the network and has great power [35]. Eigenvector centrality, the importance of a node depends both on the number of its neighbor nodes (i.e., their degree) and their importance [36].

“Structural holes” is another measurement index of the microstructure of social networks. Structural hole is a chasm between nonredundant contacts, which is a sparse zone between dense areas. From the whole network, it seems that a cave appears in the network. Burt believes that the position of an individual in the network is more important than the strength of the relationship. Its position in the network determines the information, resources, and power of an individual. Therefore, regardless of the strength of the relationship, if there is a structural hole, a third party that links two actors who have no direct connection has an information advantage and a control advantage. So, individuals or organizations must establish a wide range of contacts, occupy more structural holes, and acquire more information in order to maintain an advantage in the competition [37]. Common measurement index of structural holes are network constraint index, network effective size, network efficiency, centrality, PageRank, and local clustering coefficient. The network constraint index measures the network closure and structural hole. This index describes the compact degree to which a node in the network is directly or indirectly connected with other nodes. The higher the index is, the higher the network closure is and the fewer structural holes are. Network effective size reflects the whole influence of nodes, which can measure the importance of structural hole nodes, to a certain extent [38]. Efficiency index is used to describe the degree of influence nodes bring on other related nodes in the network. That is to say, the efficiency of the nodes in the structure hole is generally at a high level. The degree of the hierarchy can characterize some features of structural hole nodes. If a certain node has a high degree of hierarchy, it indicates that constraints concentrate on the node within its neighborhood. The centrality is used to measure the social status of individuals in society. The higher the centrality is, the higher possibility may the nodes be in the structural hole position. PageRank (PR) algorithm is a famous page ranking algorithm. According to the principle of the PR algorithm, it can be speculated that the node with higher PR score may be the structure hole node. Local cluster coefficient (CC) reflects the tendency of the adjacent nodes of a node to gather into groups, while the location of structural hole nodes in the network is quite special, so only nodes with a small clustering coefficient can become structural hole nodes [39, 40].

The development of human society is the result of the interaction between human and society and maintaining the coordinated development of human and society is the sound basis for the existence and development of society.

Therefore, sociology should take the two-way movement of human and society as the starting point of study [41]. As one of the research paradigms of sociology, social network analysis method is analyzed on the basis of the human being in the social research object, rather than the whole function, structure, and development characteristics of the society composed of human beings. The microstructural perspective of social network analysis method takes the individual in the society as the research object and evaluates the influence, function, and value of a person through various measurement indicators.

*2.1.2. Mesostructure.* “Birds of a feather flock together” is a universal law of nature and human society. This law is reflected in the structure of social network, that is, a part of individuals have a close relationship, while the other part has a relatively sparse relationship. The mesolevel of the social network structure refers to the relationship structure among individuals in the whole network.

At present, there are two-party relations, three-party relations, and cohesive subgroups analysis in mesolevel social network research. A two-party relationship studies the relationship between two individuals. There are three isomorphic types of directed-two-party relationship: one-way relationship, reciprocal relationship and irrelevant relationship [5]. The ratio of the number of one-way and reciprocal relationships reflects one-way communication and two-way interaction between network members. Therefore, the higher the ratio is, the more the one-way communication and interactive connection individuals have; irrelevant relationship are the opposite: the higher the ratio is, the less connection and interaction individual members have. The tripartite relationship is similar to the two-party relationship, and the research focuses on the relationship between small group members. The relationship between any two individuals in a tripartite relationship is a two-party relationship, so it can be said that the tripartite relationship is composed of the two-party relationship. In a directed-network, there are 64 possible relationship structures of the trio, of which only 16 are isomorphic. The more the two-party relationships in the trio, the stronger the sense of identity and belonging of network members in the group have, the easier their ideas spread within the group or influenced by others [5]. At present, the main theoretical research on the two-party relationship and three-party relationship is *Introduction to Social Network Analysis* written by a scholar, Liu Jun. The applied research includes that Li et al. applied the two-party relationship and three-party relationship theory to study the characteristics of China’s peasant workers’ network [42], and Gu used this theory to study the interaction of enterprises [43].

Cohesive subgroups generally have more than three members, which refers to the subsets of members with relatively strong, direct, close, frequent, or positive relationships [44]. Cohesive subgroups analysis analyzes how many such subgroups exist in the network, the characteristics of the

relations among the members within the subgroups, and the characteristics of the relations between the members of one subgroup and another subgroup. Due to the close relationship among members of cohesive subgroups, cohesive subgroups analysis is also called “clique analysis” [45]. Cohesive subgroups include Cliques,  $N$ -cliques,  $N$ -clan,  $K$ -plex,  $K$ -core, and Lambda set. [46–48]. The purpose of cohesive subgroups research is to find the subgroups that can be “dispatched.”

With the development of complex network research, scholars have found that a large number of networks are nonuniform, that is, they are composed of many subnetworks. The relationship between individuals within these subnetworks is relatively close, while the relationship between individuals among the subnetworks is relatively sparse. Community structure is one of the main contents of complex network research and also a new direction of social network analysis [49]. Traditional community structure detection methods include the spectral bisection method and Kernighan–Lin algorithm in computer science and hierarchical clustering and nonhierarchical clustering in social science. The traditional algorithm has obvious shortcomings, which is difficult to meet the needs of community structure analysis. So it promotes the updating of community structure analysis methods. Modern methods mainly include the Newman–Girvan algorithm (NG algorithm) [50], simulated Annealing algorithm, greedy algorithm, hierarchical clustering algorithm, and extremal optimization. These methods generally take the modularity index as the evaluation basis of community structure division. And to some extent, it overcomes the defect of traditional methods that cannot determine the number of communities. Although modern methods have overcome some shortcomings of traditional methods to some extent, they still need to be revised, such as the indeterminacy of the number of communities and the evaluation index of community structure with high time complexity is single. The research of the community mainly focuses on the development and optimization of community structure detection methods and algorithms. For example, Dao et al. studied the comparative evaluation of community detection methods [51]. Zarei and Meybodi research on detecting the community structure in complex networks based on genetic algorithm [52]. Domestic scholars research less on the community structure. Zhang and Du Haifeng redefined the community structure of the weighted network based on the relationship between the individual attribute. At the same time, inspired by the detection methods of linked communities and overlapping communities, they proposed a probability model for weighted network community structure detection and the corresponding optimization calculation [53]. Du et al. proposed a Community Structure detection algorithm for dynamic networks in their research, and the algorithm can detect the changes of dynamic networks’ community structure [54].

**2.1.3. Macrostructure.** Whole network analysis is a kind of relationship analysis method with integral and global characteristics, which can show the whole network structure,

density, and the relationship between the whole network members. The measurement indexes of the whole network structure mainly include network density, network centrality, clustering coefficient, and the average shortest path length. Network density is the ratio of the actual number of edges in the network to the maximum number of edges that can be accommodated [55], which is used to describe the density of interconnected edges between nodes in the network [56]. In social networks, network density is often used to measure the density and evolution of social relations. The level of network centrality indicates how much weight a node (individual) has in the whole network. In other words, it also indicates how much capacity a node has to block the transmission of information or distort the content of the information in the process of information transmission, thereby affecting the behavior of other nodes. The clustering coefficient is a coefficient indicating the degree of clustering of nodes in a network [57]. Studies have shown that in real networks, due to the relatively high density of connection points, nodes always tend to establish a set of tight organizational relationships, which is often more likely than a random connection between two nodes. The clustering coefficient includes the global clustering coefficient, local clustering coefficient, and average clustering coefficient [58], where the global clustering coefficient is based on node triples and calculated by the number of closed triples/the number of all triples. The local clustering coefficient of a node indicates the degree of compactness of its adjacent nodes to form a cluster. The average clustering coefficient is the mean value of the local clustering coefficient of all nodes defined by Watts and Strogatz [59]. According to the social network theory, an efficient network should be like any node in the network that can quickly reach other nodes through the shortest path. In this way, individuals in the network can quickly obtain resources, such as information flow. The average shortest path length is the average of this path [60].

In 1966, Friedman, in his book *regional development policy*, formally proposed the “core-periphery theory,” also known as the “center-periphery theory” [61]. The theory afterward became the main analytical tool for developing countries to study the space economy. Later, some scholars applied the “center-periphery theory” to the analysis of the social network structure in sociology. The core-periphery structure in a social network refers to a unit structure composed of a large number of individuals with a close connection in the center and sparsely dispersed periphery. The central part of the node plays an obvious leading role in the network, while the peripheral nodes are marginalized [62]. Core-periphery structure exists in many social phenomena, such as the citation network of scientific journals. The core area and the periphery area together constitute a completed space system; the two functions are different but interdependent and develop together.

At the beginning of the study of social networks, researchers believed that random networks were the most suitable to describe real networks [63]. However, with the deepening of the study, they found that a large number of real networks were neither regular networks nor random networks, but complex networks with unique characteristics. Qian Xuesen gave a strict

definition of the complex network, which refers to a network with self-organization, self-similarity, attractor, small world, and part or all of the network with scale-free. The complex network models mainly include small-world network, scale-free network, and random network. Small-world network [59] and scale-free network [64] were proposed by Watts, Strogatz, and Barabasi in 1998 and 1999, respectively. Both networks are between regular and random networks. Small-world networks are characterized by a short average path length and large clustering coefficient [65]. Scale-free networks have serious heterogeneity, and their degree distribution conforms to power-law distribution [66]. Compared with scale-free networks, scholars pay more attention to small-world networks. Small-world network theory develops rapidly and has been widely applied in physics, computational science and technology, biology, and other fields. For example, Taylor used the concept of the small world to study protein structure and function [67]. Ebel et al. also established a complex network dynamic model based on acquaintance networks [68]. Domestic scholar Cao Rui used the small-world property of complex network theory to construct and analyze EEG functional brain network of alcoholic patients [69]. In addition, small-world networks are also applied to relevant researches of enterprises [70] and weibo [71]. Generally speaking, there are relatively few researches related to complex networks in China, especially small-world networks.

There are also other indicators for the study of the macrolevel of social networks. For example, Peng et al. proposed a hierarchical structure in 2017. Hierarchy structure represents the degree of orderliness in a relationship, usually determined by identity or prestige. When computing the hierarchical structure of a network, UCINET software is usually first used to halve the network matrix, and then, the hierarchical routine program is run to measure the specific values [72]. In 2016, Gupta et al. proposed structural equivalence. Structural equivalence means that two actors in a network have an equal relationship with other actors, and the measurement is carried out by calculating Euclidean distance of paired actors. The smaller the Euclidean distance between actors, the more equal they are in the structure. If the Euclidean distance between the two actors is 0, they are completely equivalent in structure [73].

The relationship in the Chinese society is a special, emotional social bond with the function of the exchange of favors among the actors. Different from the weak linkage of the western relationship, the social capital of China's relationship is a strong relation bond [13]. However, Chinese and Western social relation networks also have something in common. They are not regular networks or random networks, but complex networks with a core-periphery structure.

*2.1.4. Comprehensive Study of Microstructures, Mesostructures, and Macrostructures.* Although micro, meso, and macrostudies can reveal the structure of social networks at corresponding levels, they cannot reflect the evolution process of networks and have certain limitations. Therefore, the complex network model arises at the historic moment and evolves.

(1) *Regular Graph Network.* The regular network refers to the translational symmetry lattice, and the number of neighbors of any lattice in the network is the same. Some scholars call one-dimensional chain and two-dimensional square lattice the regular network, and they think that the regular network can be applied to the relationship research of real systems. However, with the deepening of complex network research, scholars find that regular network does not conform to the characteristics of the complex network, so a random graph model is proposed.

(2) *Random Graph Model.* Compared with the regular network, the random network is the other extreme. Given a fixed number  $n$  and the probability  $P$  of generating an edge between two points, the network generated by this method is  $G(n, P)$  random network [74]. In 1960, Erdos and Rényi first proposed the theory of random graphs. The model they proposed is also known as the Erdos and Renyi model (ER model) [75], which views the observed network as the implementation of a random process, that is, a specific implementation of a series of networks with similar important characteristics. Although randomness conforms to the main characteristics of the formation of most real networks, it is difficult for people to have a visual understanding of the formation of complex networks and the interaction between different nodes. There are two kinds of random network models. The first is that there are  $N$  nodes in the random network, and the possibility of each node group connecting with each other is  $P$ . The random network model with fixed  $P$  is called the  $G(N, P)$  model. The other random network model is the  $G(n, L)$  model, which specifies the number of connections in the network. Nevertheless, the stochastic graph model has made an important contribution to the social network model. Later scholars developed the P1 model, P2 model, and exponential random graph model on the basis of the random graph model.

P1 model, also known as the binary relation model, was proposed by Holland and Leinhardt in 1981 [76]. The binary model assumes that some pairs of nodes in physics are reciprocal for directed networks, that is, the binary relation is independent of each other.

P2 model is a random effect model of covariates of network directed graph, which aims to associate binary network data with covariates for analysis while considering a specific network structure [77].

Markov model is made up of Frank and Strauss in 1986 after a summary of the development of the early social network analysis model, proposed by Frank and Strauss [78]. It is the earliest exponential random graph model, Markov model is the basic idea of "under the condition of the given current only or information, (i.e., before the current state of history) in the past, is irrelevant to predict the future (i.e., after the current state of the future)." Then, Wasserman and Pattison [79] and Frank [80] put forward a more general exponential random graph model based on the Markov model, but the model still has some shortcomings. Therefore, Snijders et al. [81] proposed three new statistics of the model, using alternative  $k$ -stars, triangles, and paths, using



independent Markov models, which could also solve the problem of degradation of the Markov model.

The social network analysis model is constantly developing in order to better reveal the social process and characteristics reflected by the network structure. At present, domestic applications of stochastic graph models mainly focus on related fields. For example, the formation mechanism of the patent reference relationship is explained from the perspective of the exponential stochastic graph model [82], network news media analysis and research are carried out based on the exponential stochastic graph model [83]. There is a lack of optimization research on the stochastic graph model itself. The international study of social network theory and stochastic graph model is in the leading position.

(3) *Small-World Network Model*. Small-world network was first proposed by Watts and Strogatz [59], who pointed out that such network graphs could be identified by two independent structural features, namely, the clustering coefficient and the average distance between nodes (also known as the average shortest path length). In the subsequent research, Watts and Strogatz proposed a new graph model (namely, the current WS model), which has two features: (1) the average shortest path length is small and (2) the aggregation coefficient is large. After years of development, small-world networks have been widely used in the studies of sociology, Earth science, and brain [84–88].

(4) *Scale-Free Network*. Scale-free network was proposed by Barabási in 1999. Scale-free network is between regular and random network [64]. A network is said to be scale-free if most of the “normal” nodes have few connections and a few of the “hot” nodes have extremely many connections [89]. In more technical terms, the degree distribution (approximately) of scale-free networks satisfies the power-law distribution. In addition to the power-law distribution of node degree, scale-free networks also have robustness and vulnerability [90]. Robustness and vulnerability are some of the basic characteristics of large-scale networks and also important topological characteristics that show the significant difference between random graph networks and scale-free networks. Due to the existence of hub nodes, the scale-free network has strong fault tolerance to random failures. If errors occur randomly, the scale-free network almost will not be affected with a small number of hub nodes, and the deletion of other nodes has little impact on the network structure. However, if the hub nodes are deliberately attacked, the network structure will be easily damaged and become discrete and broken [91].

According to the scale and level of the research object, this paper summarizes and expounds the social network structure from three levels: micro, meso, and macro. The study of the single hierarchical network structure of the complex network system is not enough to explain the complex sociological problem. Therefore, scholars begin to pay attention to the relationship between complex networks at different levels, that is, to discuss the two-way influence relationship and mechanism of different hierarchical

networks. At present, research on complex networks at home and abroad has achieved certain results [92–95], but there are still few research on the relationships between complex networks at different levels, which is also a direction worthy of research.

## 2.2. Review of Relevant Studies on Social Network Structure.

Social network analysis is a set of an analytical method that combines the sociology, the graph theory, mathematics, and metrology, which based on the study of the interaction between social actors, conducting precise quantitative analysis on the relationship between the social network structure and properties, thus provide quantitative tools for the construction of social network theory and empirical propositions inspection. Social network analysis is not only a set of techniques for the analysis of relationships or structures but also a theoretical method of structural analysis, which is widely used.

Social network is a collection composed of social actors as nodes and their relationships. Social support and social discussion are the main causes of such connections. The former is based on the material and spiritual support between individuals to quantify the relationship network, and the latter is based on whether the individuals discuss important issues to quantify the relationship network. Therefore, the social support network and social discussion network are two important contents of social network research.

2.2.1. *Social Support Network*. The concept of social support first appeared in social etiology of mental disorders in the 1970s. In the beginning, social support was associated with the individual’s physiological, psychological, and social adaptability. Then, it gradually developed into the research on the relationship between the individual’s social relationship network and social support network and the individual’s health. Cobb believes that social support consists of three aspects. One is that individuals believe that they are cared for and loved; the other is that individuals believe that they are respected and valuable; and the third is that individuals believe that they belong to a common information network and assume responsibility in this network [96]. Huang believed that social support refers to the material and spiritual help that individuals receive from the society (such as relatives, friends, and units) [97]. Zhou and Feng hold that social support includes the recipient, the giver, the content, and methods. The recipient can be any individual and also the vulnerable group in the society. The giver may be the state, society, family, friends, and colleagues. Content and methods include many aspects, such as emotion, material and information [98]. Scholars’ definitions of social support have two points in common: first, social support takes place among social actors; second, social support transmits spiritual and material information. Then, a social support network is the interaction between individuals, through which individuals can maintain their social identity and get emotional support and material assistance.

Scholars have different definitions for the types of social support. Van del. P divides social support into three aspects: practical support, emotional support, and social interaction support [99]. Xiao and Yang believed that social support involves three aspects. First, it is objective and visible, such as social network, group relationship, and direct material assistance. The second is a subjective experience, which mostly refers to the subjective feeling and experience of being respected, understood, and supported in social practice. The third is the utilization efficiency of social support. Sometimes, social support may not be well accepted and utilized for some reasons, although there are appropriate social support resources and conditions [100]. The author believes that social support can be divided into two categories: first, objective, and practical support, that is, actual social support, including material assistance and direct service; second, received social support. Subjective experienced or emotional support refers to individuals' emotional experience and satisfaction in feeling respected, supported, and understood in society.

Social support can relieve individual psychological pressure and eliminate individual psychological barriers and plays an important role in promoting individual mental health [101]. Therefore, it is widely used in the study of happiness and physical and mental health. In 2016, Fang Liming conducted the study of social support and subjective well-being of rural elderly people. The study showed that, in the formal social support factors, medical insurance and pension and other social security system construction made a significant effort in increasing the subjective well-being of the elderly. For the different living arrangements of the rural elderly, the influence of different social support on the subjective well-being of them [102] is different. Tao and Shen studied the impact of social support on the physical and mental health of the rural elderly and pointed out that both informal social support and formal social support had positive effects on the physical and mental health of the rural elderly, and the addition of formal social support had no significant crowding-out effect on the informal social support. Domestic relevant studies mainly focus on exploring the relationship between social support of college students, the elderly, rural population and other groups, and psychological concepts such as happiness and physical and mental health [103].

Due to the strong relationship between China's social capitals, domestic scholars have carried out new year greetings network research with Chinese characteristics [104]. New year greetings are a traditional Chinese folk custom, a way to express good wishes to each other, people bid farewell to the old, and welcome the new; it indirectly reflects human relations. Using data from The Spring Festival Greetings website, Bian Yanjie found that the quality of social capital among Chinese urban residents is influenced by class status and occupational activities.

In addition to its role in individual psychology, social support networks also contribute to the individual process of job-hunting [105]. As the size of a job seeker's network structure (the number of network members) increases, job seekers get more information that can help them find better

jobs. Weiyan et al. found in their research that subjective support and social support can predict employment intensity, short-term or long-term job-hunting, and work status, and subjective support in social support is more conducive to students' job-hunting behavior [106]. Granovetter found in his study that job seekers in the United States rely mainly on weak connections to find jobs and that those with more connections to high earners receive more benefits than those with less connections [107].

*2.2.2. Social Discussion Network.* Social discussion network is the core relationship discussion network. It is a clear and specific operational definition to define the network relationship by asking surveyors who they have discussed important issues with in the past six months [108]. As a part of social network analysis, social discussion network studies the interaction between individual behavior and network [109]. The concept of a discussion network was put forward by Marsden in his research in 1985. He found that the discussion network of Americans is relatively small, relative-centered, relatively dense, and homogeneous [110]. In 1986, Ruan and Lu conducted research and investigation on the discussion network of urban residents in Tianjin and found that compared with the discussion network in the United States, Tianjin had a larger scale, lower heterogeneity, higher homogeneity, and stronger closeness [111]. Based on the whole network data of the special survey of rural floating population, Li et al. analyzed the network characteristics of different types of migrant workers' social support networks and social discussion networks. The study found that the two-party relationship, three-party relationship, and whole network characteristic indexes of the social support network for migrant workers were significantly higher than that of the social discussion network [42].

In social discussion networks, individuals' ideas and opinions may influence the studies of other individuals through social networks [112]. Li et al. used the scoring matrix of user projects in their research to find out the importance of user ratings and determine their impact on the whole social network [113]. In their study, Xiong et al. mentioned that, after users reposted a topic, the influence of the users on their neighbors was always present [114].

Social discussion network also includes marriage discussion network, fertility discussion network, contraception discussion network, and elderly-care discussion network. The research of Ren et al. pointed out that, in social discussion networks, migrant workers formed more subgroups in the discussion of marriage, childbirth, and the elderly, while few subgroups formed in the discussion of contraception, a privacy issue [115]. Through the analysis of the social support network and social discussion network, it is found that migrant workers are less willing to discuss more private topics, and relatively speaking, they are more likely to have supportive behaviors [115]. Relevant studies on marriage networks also found that the opinions and opinions of members in the network would affect the individual's satisfaction and risk degree of

marriage. In their research, Wang and Zhou found that the quality of social networks was lower, and marital happiness was also lower [116]. Jin et al. pointed out that the attitudes and concepts of social network members of migrant workers towards unmarried pregnancy and extramarital love can be transmitted to migrant workers through social networks and have an impact on them. Social network variables that may have an impact on the concept of love and marriage of migrant workers include the comprehensive effect of scale, weak relationship, and the influence of network members [117]. Similarly, social networks influence the spread of ideas about fertility, contraception, and elderly-care among individuals. Buyukkececi found in his research that the birth, marriage and divorce of siblings all have certain influences on the birth, marriage and divorce of individuals [118]. As discovered in S Pink's book, *Fertility and social interaction—a simulation approach*, parents, brothers and sisters, and her colleagues have a great influence on the result of social interaction effect on bearing.

By combing the domestic and foreign literature, we found that social support network and discussion network is mainly used in the subjective feeling of different groups (such as satisfaction and loneliness) and ideas (birth, contraception, and pension). It involves a wide range of research objects, but most of the studies start from the individual level and lack the research from the overall network perspective, as well as the visualization and dynamic evolution of the network, which is a future research direction.

### 3. Conclusion

In this paper, the theory of the social network is sorted out, and then, the research methods of the social network structure are reviewed from the micro, meso, and macro-levels. Finally, the research on the social support network and social discussion network in the social network structure are reviewed.

As a sociological research perspective, social network research method has been widely used in other fields, providing a new way of thinking for the research of relation and network problems research. The visualization and dynamic study of the network also enables us to see the structure and dynamic evolution of the society more intuitively, which deepens our understanding of the research problem.

It can be summarized as follows:

- (1) There are many researches on the single hierarchy structure of social network, but a few researches on the combination of multiple-level structure. There are few researches on the micro, meso, and macro-level of the social network structure, especially on the model of an exponential random graph.
- (2) There are more studies on the static perspective of social network, but less on the dynamic perspective. The analysis and research of social network mainly focus on the static study, analyzing the network structure and characteristics, and seldom grasp the evolution of the network from the dynamic perspective. Dynamic

network research can capture the characteristics and phenomena that static network research cannot.

- (3) There are many researches on individual network, but few researches on whole network. According to different research objects, social networks can be divided into individual networks and whole networks. Individual network research focuses on the relationship between the attributes of individuals/nodes and individual networks, while whole network research focuses on the structure, graph-theoretic nature, and location attributes of networks. At present, there is a lack of relevant research on the whole network.
- (4) There are many researches on support networks of social networks, but few researches on discussion networks. The development of social support network theory is relatively complete. Although social discussion network has developed, it generally lags behind social support network. At present, the research mainly focuses on social support networks, and the research on the social discussion network needs to be further expanded and deepened.
- (5) There are many traditional social network analysis methods, but few studies from the perspective of complex networks. Once the theory of the complex network was put forward and developed rapidly, it soon became the key research direction of social network theory. However, there are relatively few in-depth studies and applications of complex network.

In short, the existing research on social networks lacks systematicness and comprehensiveness. Most of the research are just simple statistical analysis, which are not in-depth enough and lack the whole network research in the social network structure. The research direction is relatively single, focusing on the social support network and lacking the theoretical and applied research on the social discussion network. Since the development of social networks, complex networks have gradually become the research focus of social networks. However, in the academic research field, there is a lack of attention paid to the development and improvement of complex network models and the research of complex network analysis methods, and the research of complex networks has not been widely applied into the research of various fields. All these provide space for future research.

### Data Availability

No data were used to support this study.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Authors' Contributions

Ning Li was responsible for writing the original draft, review and editing, and literature collection. Qian Huang was responsible for determining the framework of paper, funding

acquisition, and review and editing. Xiaoyu Ge was responsible writing the original draft. Miao He was responsible for review and editing. Shuqin Cui was responsible for review and editing. Penglin Huang and Shuairan Li were responsible for literature collection. Sai-fu Fung was responsible for funding acquisition.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant nos. 61440036 and 61040029 and the National Social Science Foundation of China under Grant no. 16ATY002.

## References

- [1] J. A. Barnes, "Class and committees in a Norwegian Island Parish," *Social Networks*, vol. 7, no. 1, pp. 39–58, 1954.
- [2] B. Wellman, "Structural analysis: from method and metaphor to theory and substance," in *Social Structures: A Network Approach*, B. Wellman and S. D. Berkowitz, Eds., Cambridge University Press, Cambridge, UK, 1988.
- [3] J. O. Kereri and C. M. Harper, "Social networks and construction teams: literature review," *Journal of Construction Engineering and Management*, vol. 145, no. 4, pp. 1–10, 2019.
- [4] S. Tabassum, F. S. F. Pereira, and S. Fernandes, *Social Network Analysis: An Overview*, Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery, Hoboken, NJ, USA, 2018.
- [5] J. Liu, *Introduction to Social Network Analysis*, Social Science Press, Beijing, China, 2004.
- [6] Y. Liang, "Social security for vulnerable groups from the perspective of social network," *Financial Teaching and Research*, no. 2, pp. 71–73, 2015.
- [7] R. Leonard, D. Horsfall, and K. Noonan, "Identifying changes in the support networks of end-of-life carers using social network analysis," *BMJ Supportive & Palliative Care*, vol. 5, no. 2, pp. 153–159, 2015.
- [8] S. Zhang, S. Xu, and D. Li, "A review of social network theory and its research," *Think Tank Age*, no. 37, p. 264, 2019.
- [9] R. Ghosh and K. Lerman, "Rethinking centrality: the role of dynamical processes in social network analysis," *Discrete and Continuous Dynamical Systems—Series B*, vol. 19, no. 5, 2012.
- [10] M. Li and Z. Jia, "A review of the development and research progress of social network theory," *China Management Informatization*, vol. 17, no. 3, pp. 133–135, 2014.
- [11] M. S. Granovetter, "The strength of weak ties," *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.
- [12] S. Mukerjee, T. Yang, and S. González-Bailón, "What counts as a weak tie? A comparison of filtering techniques for weighted networks," *SSRN Electronic Journal*, 2019.
- [13] Y. Bian and L. Zhang, "On relational culture and relational social capital," *Journal of Humanities*, no. 1, pp. 107–113, 2013.
- [14] N. Lin, "Social capital: resources, motivations, and interactions," no. 4, pp. 41–54, 2001.
- [15] H. Cook and R. Hole, "Appropriately homeless and needy: examining street homeless survival through the lens of Bourdieusian social capital theory," *Journal of Human Behavior in the Social Environment*, vol. 30, no. 7, 2020.
- [16] T. W. Valente and S. R. Pitts, "An appraisal of social network theory and analysis as applied to public health: challenges and opportunities," *Annual Review of Public Health*, vol. 38, no. 1, pp. 103–118, 2017.
- [17] T. Zhou, Y. Wang, and S. Deng, "Research on user participation behavior of online health community based on social capital theory," *Journal of Information Resource Management*, vol. 10, no. 2, pp. 59–67, 2020.
- [18] Z. Zhang and Y. Wang, "Research on the construction of reading interest network in public libraries—from the perspective of social network theory," *Journal of Henan Library*, vol. 40, no. 7, pp. 10–12, 2020.
- [19] J. Zhu, L. Zhu, and C. Bao, "Top-k structure holes detection algorithm in social network," in *Proceedings of the 2018 9th international conference on information technology in medicine and education (ITME)*, IEEE Computer Society, Hangzhou, China, October 2018.
- [20] R. S. Burt, *Structural Holes*, Harvard University Press, Cambridge, MA, USA, 1992.
- [21] J. Lin, *Social Network Analysis: Theory, Method and Application*, Beijing Normal University Press, Beijing, China, 2009.
- [22] D. Hussein, S. Park, S. N. Han, and N. Crespi, "Dynamic social structure of things: a contextual approach in CPSS," *IEEE Internet Computing*, vol. 19, no. 3, pp. 12–20, 2015.
- [23] B. S. Hanson, B. Liedberg, and B. Owall, "Social network, social support and dental status in elderly Swedish men," *Community Dentistry & Oral Epidemiology*, vol. 22, no. 5, pp. 331–337, 2015.
- [24] J. Scott and P. J. Carrington, *The SAGE Handbook of Social Network Analysis*, SAGE, Thousand Oaks, CA, USA, 2011.
- [25] A. Korshunov, I. Beloborodov, and N. Buzun, *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge, UK, 2014.
- [26] C. Fang, S. Xiaoming, and X. Wang, "The research on the impact of the ego-network's changes from the key inventors to the innovation performance: the whole network as an intermediary variable," *Science & Technology Progress and Policy*, vol. 34, no. 17, pp. 80–90, 2017.
- [27] A. Biswas and B. Biswas, "Investigating community structure in perspective of ego network," *Expert Systems with Applications*, vol. 42, no. 20, pp. 6913–6934, 2015.
- [28] L. I. Bo, W. Lei, and P. E. School of, "Feasibility analysis of passing performance in football match by social network analysis," *Journal of Beijing Sport University*, vol. 40, no. 8, pp. 112–119, 2017.
- [29] V. Arnaboldi, M. Conti, and M. La Gala, "Ego network structure in online social networks and its impact on information diffusion," *Computer Communications*, 2016.
- [30] O. Casquero, M. Benito, and J. Romo, "Participation and interaction in learning environments: a whole-network analysis," in *Utilizing Virtual and Personal Learning Environments for Optimal Learning*, IGI Global, Harrisburg, PA, USA, 2016.
- [31] M. Magnani and M. Marzolla, "Path-based and whole-network measures," in *Encyclopedia of Social Network Analysis and Mining*, Springer, New York, NY, USA, 2016.
- [32] J. Zhang and Y. Luo, "Degree centrality, betweenness centrality, and closeness centrality in social network," in *Proceedings of the 2017 2nd International Conference on Modelling, Simulation and Applied Mathematics (MSAM2017)*, Bangkok, Thailand, March 2017.
- [33] L. Zhen-Ya, W. Wan-Hong, and Z. Hong-Mei, "The impact of degree centrality of social network, work engagement on organizational citizenship behavior of nurses," *Modern Preventive Medicine*, vol. 45, no. 6, pp. 1039–1042, 2018.

- [34] C. P. Hoffmann, C. Lutz, and M. Meckel, "A relational altmetric? Network centrality on ResearchGate as an indicator of scientific impact," *Journal of the Association for Information Science & Technology*, vol. 67, no. 4, pp. 1–11, 2016.
- [35] U. Brandes, S. P. Borgatti, and L. C. Freeman, "Maintaining the duality of closeness and betweenness centrality," *Social Networks*, vol. 44, pp. 153–159, 2016.
- [36] S. Chakraborty and B. K. Tripathy, "Privacy preserving anonymization of social networks using eigenvector centrality approach," *Intelligent Data Analysis*, vol. 20, no. 3, pp. 543–560, 2016.
- [37] M. Tortoriello, "The social underpinnings of absorptive capacity: the moderating effects of structural holes on innovation generation based on external knowledge," *Strategic Management Journal*, vol. 36, no. 4, pp. 586–597, 2015.
- [38] W. Yun-Ming, W. Qing-Ye, and P. Cheng-Sheng, "Method for key nodes identification in command and control network by considering structural holes," *Fire Control & Command Control*, vol. 42, no. 3, pp. 59–63, 2017.
- [39] H. Zhongming and W. Yang, T. Xusheng, L. Wen, and Y. Weijie, "Comparison and analysis of measurement index of social network structure hole nodes," *Journal of Shandong University (Engineering Science Edition)*, vol. 45, no. 1, pp. 1–8, 2015.
- [40] D. Wang, "Comparison and evaluation of structural hole algorithms," *Modern Intelligence*, no. 9, pp. 153–156, 2008.
- [41] H. C. Carey, "Principles of social science," in *Principles of Social Science*, J. B. Lippincott & Co., Philadelphia, PA, USA, 1858.
- [42] S. Li, Y. Ren, Feldman, and X. Yang, "Analysis on the characteristics of the overall social network of migrant workers in China," *Chinese Journal of Population Sciences*, no. 3, pp. 19–29, 2006.
- [43] H. Gu, *Research on Industrial Cluster Upgrading Based on Social Network Evolution Analysis*, Southeast University, Dhaka, Bangladesh, 2006.
- [44] L. Xing and Q. Xiaodong, "Research of word-of-mouth communication mechanism based on cohesive subgroup," *Application Research of Computers*, vol. 35, no. 12, pp. 3584–3587, 2018.
- [45] Y. Guichan, "The choice of cooperating team within enterprise based on cohesive subgroup analysis," *Sci-Tech Information Development & Economy*, vol. 25, no. 19, pp. 109–112, 2015.
- [46] Z. Miao and B. Balasundaram, "Approaches for finding cohesive subgroups in large-scale social networks via maximum K-plex detection," *Networks*, vol. 69, no. 4, 2017.
- [47] A. Nguyen and S. H. Hong, "K-core based multi-level graph visualization for scale-free networks," in *Proceedings of the Pacific Visualization Symposium*, IEEE, Seoul, Republic of Korea, April 2017.
- [48] M. Gjoka, E. Smith, and C. T. Butts, "Estimating clique composition and size distributions from sampled network data," in *Proceedings of the IEEE INFOCOM*, Toronto, Canada, May 2014.
- [49] D. Mehrle, A. Strosser, and A. Harkin, "Walk-modularity and community structure in networks," *Network Science*, vol. 3, no. 3, pp. 348–360, 2015.
- [50] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, no. 2, Article ID 026113, 2004.
- [51] V. L. Dao, C. Bothorel, and P. Lenca, "Community structure: a comparative evaluation of community detection methods," 2020, <https://arxiv.org/abs/1812.06598>.
- [52] B. Zarei and M. R. Meybodi, "Detecting community structure in complex networks using genetic algorithm based on object migrating automata," *Computational Intelligence*, vol. 36, 2020.
- [53] K. Zhang and H. Du, "A weighted network community structure detection model based on individual relationship attributes," *Journal of Systems Management*, vol. 25, no. 6, pp. 1083–1090, 2016.
- [54] H. Du, Z. Yue, S. Li, and Y. F. Chen, "Dynamic network community structure detection method based on modularity index," *Systems Engineering Theory and Practice*, vol. 29, no. 3, pp. 162–171, 2009.
- [55] C. Intanagonwiwat, D. Estrin, and R. Govindan, "Impact of network density on data aggregation in wireless sensor networks," in *Proceedings of the International Conference on Distributed Computing Systems*, IEEE, Vienna, Austria, July 2002.
- [56] Z. Youbing, L. A. Dalguer, and S. S. Goo, "Evaluating the effect of network density and geometric distribution on kinematic source inversion models," *Geophysical Journal International*, vol. 1, p. 5618, 2014.
- [57] R. Vargas, F. Garcea, and B. Z. Mahon, "Refining the clustering coefficient for analysis of social and neural network data," *Social Network Analysis & Mining*, vol. 6, no. 1, p. 49, 2016.
- [58] Z. Ertem, A. Veremyev, and S. Butenko, "Detecting large cohesive subgroups with high clustering coefficients in social networks," *Social Networks*, vol. 46, pp. 1–10, 2016.
- [59] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [60] T. Matsumura, K. Iwasaki, and K. Shudo, "Average path length estimation of social networks by random walk," in *Proceedings of the IEEE International Conference on Big Data & Smart Computing*, IEEE, Shanghai, China, 2018.
- [61] J. R. Friedman, *Regional Development Policy: A Case Study of Venezuela*, MIT Press, Cambridge, MA, USA, 1966.
- [62] J. Jia and A. R. Benson, "Random spatial network models with core-periphery structure," 2018, <https://arxiv.org/abs/1808.06544>.
- [63] B. Bollobás, *Random Graphs*, Academic Press Inc., London, UK, 1995.
- [64] A.-L. Barabási, R. Albert, and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.
- [65] Z. Jianyu, X. Xi, and S. Yi, "A simulation research on small world effect in knowledge flow's network evolution," *Management Review*, vol. 27, no. 5, pp. 70–81, 2015.
- [66] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, OUP Catalogue, Oxford, UK, 2013.
- [67] N. R. Taylor, "Small world network strategies for studying protein structures and binding," *Computational & Structural Biotechnology Journal*, vol. 5, no. 6, pp. 1–7, 2013.
- [68] H. Ebel, J. R. Davidsen, and S. Bornholdt, "Dynamics of social networks," *Complexity*, vol. 8, no. 2, pp. 24–27, 2003.
- [69] R. Cao, *Application of Nonlinear and Complex Network Theory in EEG Data Analysis*, Taiyuan University of Technology, Taiyuan, China, 2014.
- [70] X. Tang, X. Sun, G. Tian, H. Hu, and J. Wang, "Effect of small-world network dynamics on the creativity of key

- enterprise developers,” *Journal of Management Engineering*, vol. 32, no. 4, pp. 54–62, 2012.
- [71] Y. Liu, P. Lin, and Z. Zhao, “Research on the evolution of micro-blog rumor propagation based on small-world network,” *Science of Complex Systems and Complexity*, vol. 11, no. 4, pp. 54–60, 2014.
- [72] W. Peng, D. Jin, and Q. Zhu, “Review and prospect of team social network research,” *China Human Resources Development*, no. 3, pp. 57–68, 2017.
- [73] N. Gupta, V. Ho, J. M. Pollack, and L. Lai, “A multilevel perspective of interpersonal trust: individual, dyadic, and cross-level predictors of performance,” *Journal of Organizational Behavior*, vol. 37, no. 8, pp. 1271–1292, 2016.
- [74] V. D. Hofstad, “Random graphs and complex networks,” *Preparation*, vol. 27, no. 2, pp. 188–206, 2009.
- [75] P. Erdos and A. Rényi, “On the Evolution of Random Graphs,” *Publications de l’Institut Mathématique Hungarian Academy of Sciences*, vol. 5, no. 1, pp. 17–60, 1960.
- [76] P. W. Holland and S. Leinhardt, “An exponential family of probability distributions for directed graphs,” *Journal of the American Statistical Association*, vol. 76, no. 373, pp. 33–50, 1981.
- [77] M. A. J. Duijn, T. A. B. Snijders, and B. J. H. Zijlstra, “P2: a random effects model with covariates for directed graphs,” *Statistica Neerlandica*, vol. 58, no. 2, pp. 234–254, 2004.
- [78] P. Wang, G. L. Robins, and P. E. Pattison, *PNet: A Program for the Simulation and Estimation of Exponential Random Graph Models*, School of Behavioural Science, University of Melbourne, Melbourne, Australia, 2006.
- [79] S. Wasserman and P. Pattison, “Logit models and logistic regressions for social networks: I. An introduction to Markov graphs andp,” *Psychometrika*, vol. 61, no. 3, pp. 401–425, 1996.
- [80] O. Frank, “Statistical analysis of change in networks,” *Statistica Neerlandica*, vol. 45, no. 3, pp. 283–293, 2010.
- [81] T. A. B. Snijders, A. Lomi, and V. J. Torló, “A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice,” *Social Networks*, vol. 35, no. 2, pp. 265–276, 2013.
- [82] G. Yang, C. Liang, J. Zhang, and G. Li, “An explanatory framework for the formation of patent reference relations: from the perspective of an exponential random graph model,” *Books and Information Work*, vol. 63, no. 5, pp. 100–109, 2019.
- [83] A. Chen and Y. Yan, “Network news media analysis based on exponential random graph model,” *Journal of Jinling Science and Technology University*, vol. 28, no. 2, pp. 30–36, 2012.
- [84] A. Jiménez, K. F. Tiampo, and A. M. Posadas, “Small world in a seismic network: the California case,” *Nonlinear Processes in Geophysics*, vol. 15, no. 3, pp. 389–395, 2008.
- [85] O. Sporns, D. Chialvo, M. Kaiser, and C. Hilgetag, “Organization, development and function of complex brain networks,” *Trends in Cognitive Sciences*, vol. 8, no. 9, pp. 418–425, 2004.
- [86] P. Cohen, “Small world networks key to memory,” *New Scientist*, 2004.
- [87] A. Roxin, H. Riecke, and S. A. Solla, “Self-sustained activity in a small-world network of excitable neurons,” *Physical Review Letters*, vol. 92, no. 19, Article ID 198101, 2004.
- [88] S. C. Ponten, F. Bartolomei, and C. J. Stam, “Small-world networks and epilepsy: graph theoretical analysis of intracerebrally recorded mesial temporal lobe seizures,” *Clinical Neurophysiology*, vol. 118, no. 4, pp. 918–927, 2007.
- [89] Y. B. Xie, T. Zhou, and B. H. Wang, “Scale-free networks without growth,” *Physica A: Statal Mechanics and Its Applications*, vol. 387, no. 7, pp. 1683–1688, 2017.
- [90] W. Jian-Rong, W. Jian-Ping, and H. Zhen, “Degree distribution and robustness of cooperative communication network with scale-free model,” *Chinese Physics B*, vol. 24, no. 6, pp. 115–121, 2015.
- [91] J. Emmanuel and M. Peter, “Robustness of scale-free spatial networks,” *Annals of Probability An Official Journal of the Institute of Mathematical Statistics*, vol. 45, no. 3, pp. 1680–1722, 2015.
- [92] Li Shu, Z. Shao, and Y. Kong, “Research on complex network modeling method based on random graph,” *Miniature Microcomputer System*, vol. 41, no. 9, pp. 1935–1938, 2020.
- [93] L. Zhang, *Research on Dynamic Evolution of Complex Networks and its Functional Modules*, Shanghai University, Shanghai, China, 2018.
- [94] L. K. Gallos, C. Song, and H. A. Makse, “A review of fractality and self-similarity in complex networks,” *Physica A: Statistical Mechanics and Its Applications*, vol. 386, no. 2, pp. 686–691, 2016.
- [95] L. Linyuan, “Toward link predictability of complex networks,” in *Proceedings of the National Academy of Sciences of the United States of America*, Monroe, LA, USA, July 2015.
- [96] C. Sidney, “Social support as a moderator of life stress,” *Psychosomatic Medicine*, vol. 38, no. 5, pp. 300–314, 1976.
- [97] L. J. Huang, “A preliminary study on social support theory,” *Journal of Psychology*, vol. 15, no. 16, pp. 238–239, 2020.
- [98] L. Zhou and J. Feng, “Social support theory—a review of the literature,” *Journal of Guangxi Normal University*, vol. 33, no. 3, pp. 11–14, 2005.
- [99] P. Van del, “Delineating personal support network,” *Social Forces*, vol. 15, no. 1, pp. 49–70, 1993.
- [100] S. Xiao and D. Yang, “The impact of social support on physical and mental health,” *Chinese Journal of Mental Health*, vol. 1, no. 4, pp. 183–187, 1987.
- [101] R. Fernández-Peña, J. Molina, and O. Valero, “Personal network analysis in the study of social support: the case of chronic pain,” *International Journal of Environmental Research & Public Health*, vol. 15, no. 12, 2018.
- [102] L. Fang, “Social support and subjective well-being of rural elderly,” *Journal of Central China Normal University (Humanities and Social Sciences Edition)*, vol. 55, no. 1, pp. 54–63, 2016.
- [103] Y. Tao and Y. Shen, “The Impact of social support on the physical and mental health of rural elderly,” *Population and Economy*, no. 3, pp. 3–14, 2014.
- [104] Y. Bian, “Sources and functions of social capital of urban residents: network views and survey findings,” *Chinese Social Sciences*, vol. 3, pp. 136–146, 2004.
- [105] S. K. Oh and J. Jun, “Structural relationships between career barriers, social support levels, egoresilience, job search efficacy, and career preparation behavior of middleaged unemployed men,” *KEDI Journal of Educational Policy*, vol. 15, no. 1, pp. 21–42, 2018.
- [106] Z. Weiyang, L. Minghui, and D. Ximing, “Study on social support and job search behavior of college students,” *Journal of Jining Medical University*, vol. 38, no. 6, pp. 421–423, 2015.
- [107] A. Calvó-Armengol and Y. Zenou, “Job matching, social network and word-of-mouth communication,” *Journal of Urban Economics*, vol. 57, no. 3, pp. 500–522, 2005.
- [108] L. Sun, *An Empirical Study on the Influencing Factors of the “Three West” Immigrant Community Residents Discussion Network*, Lanzhou University, Lanzhou, China, 2010.

- [109] R. Lupton and J. Thornton, "Disagreement, diversity, and participation: examining the properties of several measures of political discussion network characteristics," *Political Behavior*, vol. 39, pp. 585–608, 2017.
- [110] P. V. Marsden, "Core discussion networks of Americans," *American Sociological Review*, vol. 52, no. 1, pp. 122–131, 1987.
- [111] D. Ruan and Z. Lu, "Preliminary analysis of Tianjin urban residents' social network-and comparison with American social network," *Chinese Social Sciences*, no. 2, 1990.
- [112] Y. Hu, F. Xiong, S. Pan, X. Xiong, L. Wang, and H. Chen, "Bayesian personalized ranking based on multiple-layer neighborhoods," *Information Sciences*, vol. 542, pp. 156–176, 2021.
- [113] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [114] F. Xiong, Y. Liu, and Z. J. Zhang, "An information diffusion model based on retweeting mechanism for online social media," *Physics Letters A*, vol. 376, no. 30-31, pp. 2103–2108, 2012.
- [115] Y. Ren, H. Du, Y. Xiao, and S. F. Li, "Analysis of cohesion subgroup structure of social network of migrant workers in China," *Society*, no. 5, pp. 20–40, 2008.
- [116] S. Wang and W. Zhou, "The unintended long-term consequences of Mao's mass send-down movement: marriage, social network, and happiness," *World Development*, vol. 90, 2016.
- [117] X. Jin, F. Ren, and Z. Yue, "Attitudes of migrant workers towards premarital and extramarital sex: a study based on social networks," *Population Research*, no. 5, pp. 67–78, 2008.
- [118] Z. Buyukkececi and T. Leopold, "Sibling influence on family formation: a study of social interaction effects on fertility, marriage, and divorce," *Advances in Life Course Research*, Article ID 100359, 2020.

## Research Article

# A Cognition Knowledge Representation Model Based on Multidimensional Heterogeneous Data

Dong Zhong <sup>1</sup>, Yi-An Zhu,<sup>1</sup> Lanqing Wang <sup>1</sup>, Junhua Duan,<sup>1</sup> and Jiaxuan He<sup>2</sup>

<sup>1</sup>School of Computer Science, University of Northwestern Polytechnical of China, Xi'an, Shaanxi 710129, China

<sup>2</sup>School of Software, University of Northwestern Polytechnical of China, Xi'an, Shaanxi 710129, China

Correspondence should be addressed to Dong Zhong; [dzhong@nwpu.edu.cn](mailto:dzhong@nwpu.edu.cn)

Received 26 September 2020; Revised 26 November 2020; Accepted 14 December 2020; Published 29 December 2020

Academic Editor: Fei Xiong

Copyright © 2020 Dong Zhong 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 information in the working environment of industrial Internet is characterized by diversity, semantics, hierarchy, and relevance. However, the existing representation methods of environmental information mostly emphasize the concepts and relationships in the environment and have an insufficient understanding of the items and relationships at the instance level. There are also some problems such as low visualization of knowledge representation, poor human-machine interaction ability, insufficient knowledge reasoning ability, and slow knowledge search speed, which cannot meet the needs of intelligent and personalized service. Based on this, this paper designs a cognitive information representation model based on a knowledge graph, which combines the perceptual information of industrial robot ontology with semantic description information such as functional attributes obtained from the Internet to form a structured and logically reasoned cognitive knowledge graph including perception layer and cognition layer. Aiming at the problem that the data sources of the knowledge base for constructing the cognitive knowledge graph are wide and heterogeneous, and there are entity semantic differences and knowledge system differences among different data sources, a multimodal entity semantic fusion model based on vector features and a system fusion framework based on HowNet are designed, and the environment description information such as object semantics, attributes, relations, spatial location, and context acquired by industrial robots and their own state information are unified and standardized. The automatic representation of robot perceived information is realized, and the universality, systematicness, and intuition of robot cognitive information representation are enhanced, so that the cognition reasoning ability and knowledge retrieval efficiency of robots in the industrial Internet environment can be effectively improved.

## 1. Introduction

Human beings understand the environment and recognize the objects in the environment according to the shape and color of the objects, and the information of environment and objects acquired by human beings is stored in the human brain in a structured and hierarchical form. With the improvement of computing capacity and the amount of stored data of computing system, the machine can construct the environment knowledge base by simulating human cognition and storing information and then has the ability to recognize the environment like human beings.

In the working environment of industrial Internet, describing the service environment information and storing it in the form of knowledge base is the premise of realizing

intelligent and personalized service. When robots perform tasks cooperatively in the industrial Internet, the characteristics of diversity, semantic, hierarchy, and relevance of the involved environmental information greatly increase the difficulty for robots to complete intelligent service tasks. By characterizing the environment information and storing it in the form of knowledge base, it can help robots not only master the semantic information of objects in the working environment but also help robots search for objects in the environment quickly and improve the ability of intelligent cooperative work of robots. However, in the field of robot knowledge base construction, there are some problems in the aspects of environment representation and environment information acquisition, such as insufficient practicability, poor human-machine interaction capability, insufficient



knowledge base expansion, high degree of manual participation in semantic information addition, and low reliability of independent information addition.

Knowledge graph technology, which has arisen in recent years, can effectively solve the above problems. The knowledge graph was put forward by Google in 2012 and used to improve search engines. Subsequently, many scholars began to study it one after another, and many achievements emerged in the field of knowledge storage and inquiry. A large number of structured knowledge data are usually stored in knowledge graphs, and the real-world scenes are directly modeled by using triplets being made of nodes and relationships. Using the universal “language” of triplets, the relationship between objects and objects in the scene can be effectively and intuitively represented. In addition, the knowledge graph emphasizes not only the concept and but also the entity relation and the entity attribute value, which enriches and expands the knowledge of ontology. As a graph-based data storage structure, knowledge graph based on the above characteristics not only has strong storage, query, and reasoning capabilities but also has advantages in real-time update and human-machine interaction.

The application of knowledge graph technology to the construction of robot knowledge base in industrial Internet can greatly improve the ability of the robot to represent and store environmental knowledge. It can effectively unify various environmental information into context information that the robot can understand. At the same time, the knowledge base based on knowledge graph stores environmental information in the form of structured network, which makes the robot knowledge base query have similar associative ability to human beings and becomes the key to improve the robot’s intelligence and realize service tasks.

## 2. Related Work

Environmental information representation and knowledge base construction are the research hotspots in the field of industrial Internet and are the key steps for robots to realize intelligent services in industrial Internet. The essential problem lies in how to encode and store the environment information in a unified form, and based on this encoding, robots can perceive and understand the environment and then realize intelligent cooperative work. Therefore, the process of building knowledge base is divided into two steps: acquiring environment knowledge and representing and storing knowledge.

Considerable achievements have been made in obtaining environmental information. Kollar et al. [1] proposed a task-based human-machine dialogue learning method for environmental knowledge and analyzed the entity mapping from each semantic element to the environment by introducing a joint probability distribution model on semantics. Lemaignan et al. [2] proposed a dialogue module to convert the natural language into symbolic facts (OWL statements) according to the intention of the original sentence and constructed a symbolic knowledge base. Lemaignan et al. [3] established a basic and shared world model by extracting,

representing, and using symbolic knowledge in the verbal and nonverbal interaction between humans and robots, which is suitable for subsequent high-level tasks, such as dialogue understanding. Schiffé et al. [4] proposed a flexible natural language interpretation system for mobile robot voice instructions, modeling language processing as an interpretation process that maps speech to entities in the environment. In addition, multimedia interaction and autonomous perception of robots are also important ways to acquire knowledge. Burghart et al. [5] proposed a two-layer embedded framework for service robots, which gradually embedded perception, learning, action planning, motion control, and human-like communication into the architecture. Mozos et al. [6] use robots to learn examples of common models of furniture on the Web. Through these models, the classification and location of unknown furniture in the environment are realized. Ko et al. [7] proposed semantic map representation and human-like navigation strategies for monocular robots, extracting semantic information on the basis of constructing visual maps and constructing semantic maps represented by nodes such as regions or landmarks and their spatial relationships. Song [8] embeds environmental information into artificial signposts, by identifying which robots can acquire environmental information. Tamas and Cosmin Goron [9] proposed a three-dimensional point cloud marking system based on physical characteristics, which can realize semantic interpretation of surrounding environment items (such as furniture, ceiling, and door). Literature studies [10, 11] construct an intelligent space suitable for robot task execution by adding robot-perceivable tags in the environment. The intelligent space distributes the semantic information needed by the service in the surrounding environment, and the robot obtains the corresponding information through corresponding sensing methods to complete the service task.

With the development of robot technology and artificial intelligence technology, more and more attention has been paid to the human-machine interaction of environmental information. At present, the commonly used representation methods of environmental information mainly include predicate logic representation, production rule representation, and semantic Web ontology representation, etc. Ontology’s structured knowledge representation has a wide range of representation, strong representation ability, and reasoning ability, which has quickly become a research hotspot. Literature studies [12–15], respectively, use ontology technology to construct corresponding knowledge frameworks to realize environmental representation. Park et al. [12] proposed a scene knowledge base system design based on domain knowledge ontology. Hao et al. [13] used the ontology knowledge base model to reorganize the original dataset, making the logical structure of the new dataset more suitable for upper application and improving the utilization rate of open data. Yang et al. [14] proposed a semiautomatic annotation framework which was represented by WoT resource metadata. The framework is based on a probabilistic graphical model, mapping the schematic of WoT resources to independent domain knowledge base and collectively inferring entities, classes, and relationships.

Das et al. [15] designed an ontology-based information sharing mechanism among robots, forming a collective knowledge base, which is convenient for the overall control and planning of the system. In addition to ontology representation, Jezek Moucek [16] designed a semantic framework and realized object-oriented environmental representation through semantic Web language. Chen [17] proposed an environment representation method based on quadtree. By designing an access code mechanism, the robot can quickly grasp the environment obstacle information to complete navigation tasks in complex scenes. Gao et al. [18] designed a three-layer representation model of the indoor environment, which represented the family environment in the form of holographic map and applied it to object-oriented task service.

These achievements not only reflect the progress in the research direction of environment information representation but also show some problems, such as lack of structure and expansibility of knowledge representation in the aspect of environment representation; the acquisition of representation information is characterized by high manual participation and low information reliability. Therefore, it is urgent to study the information representation method and representation information acquisition methods suitable for the intelligent service environment of industrial Internet.

The concept of knowledge graph was put forward by Google in 2012 [19] to improve the performance of its search engine. As soon as it was put forward, it aroused a strong response and got a widespread concern. Many organizations followed up quickly and carried out related research in succession to improve the performance of their own search engines. Many achievements have emerged, such as Sogou's knowledge cube and Baidu's bosom friend [20, 21]. Since then, knowledge graph technology has gradually spread to other fields.

In order to improve the accuracy of search, Marino et al. [22] combine the knowledge graph with neural network and introduce the graphic search neural network method to effectively merge the large knowledge graph into the visual classification pipeline to improve the accuracy of image classification with structured prior knowledge. Szekely et al. [23] proposed a method of constructing knowledge graph, which uses semantic technology to fuse data from different sources, extends crawled data to billions of triplets, and applies knowledge graph to the DIG system to combat human trafficking. Literature studies [24–29], respectively, constructed knowledge graphs of corresponding disciplines and carried out visual analysis. Among them, GDM Laboratory of Fudan University designed the Chinese knowledge graph related to books to realize the visualization of the classified query in the book field [27]. Jia et al. [29] constructed TCM (traditional Chinese medicine) knowledge graph, realized effective integration of TCM knowledge resources, and discussed the application prospect of TCM knowledge graph. Lu et al. [30] constructed the teaching knowledge map, discussed the knowledge graph with students, and compared the two, which may enable the teacher to reflect on their own work and cause the students to realize questions related to study, and thus, it is advantageous to enhance the teaching content, strategy, and activity.

In order to construct the corresponding database, Kumar et al. [31] use the structured data based on knowledge graph to construct a unified system combining speech recognition and language understanding, which is used to solve the specific problems of automatic speech recognition and language processing. Jia et al. [32] proposed a network security knowledge base and inference rules based on a five-element model. Using machine learning, entities are extracted, ontologies are constructed, and network security knowledge base is obtained. Kem et al. [33] proposed a knowledge model to describe the spatial structure of an environment. The network physics society entities it contains, and the relationship between them, are regarded as a graph, called cyberspatial graph (CSG). Collarana [34] designed a semantic data integration method called FuhSen, which uses keywords of Web data sources and structured search function to generate knowledge graphs to merge data collected from available Web data sources. Kumar et al. used a semantic-rich knowledge graph library to solve specific problems in automatic speech recognition and natural language processing.

These achievements show that the knowledge graph has a very broad development prospect, but most of the achievements and applications exist in search engine improvement, text information processing, and intelligent search. Hao et al. [35] proposed an environmental information representation mechanism based on knowledge graph in the field of intelligent service of home robots. In this paper, we further into the professional database in the Internet industry for the construction of knowledge graph based on the above research, at the same time in the process of knowledge extraction added the description information of work environment, the function information, and real-time status information of the mechanical arm into knowledge extraction triples, and make the mechanical arm cognitive external working environment and working object. And the mechanical arm can match the external environment and the object information to its own state and ability, so that it has the ability to judge whether it can complete some nonspecific complex tasks.

### 3. Machine Cognition Knowledge Graph Framework

In the environment of industrial Internet, robots can obtain the information of the robot's external environment and its internal state through various sensors and interaction mechanisms, which are collectively referred to as robot perception information in this paper. The expression and storage of perceptual information is the premise for robots to realize cognitive ability. However, when robots complete nonspecific complexity tasks, it is far from enough to simply recognize objects and perceive scenes. In order to achieve higher intelligence and autonomy, robots need to master not only various attributes of objects operated, such as location, function, operation mode, and instructions but also the categories of objects. In this section, aiming at the representation of robot perception information, considering the diversity, semantics, hierarchy, and relevance of the

representation of perception information when robots perform nonspecific complexity tasks, a robot cognitive knowledge graph framework including perception layer and cognition layer based on knowledge graph is designed to realize the unified representation and storage of semantic information, attribute information, spatial function information, and context information of operating objects in the environment, thus realizing the construction of robot cognitive knowledge base.

The knowledge representation method based on semantic Web, represented by knowledge graph, not only has the advantage that predicate logic representation is easy to understand, but also knowledge always exists in the form of triplets (entities, relationships, and entities) in the knowledge graph, which can represent knowledge in a formal and concise way and intuitively model various scenes in the real world. In the knowledge graph, the entities of the triplets are regarded as nodes and the relationships between the entities are drawn as edges, so the knowledge base containing a large number of triplets forms a huge knowledge network. The method can effectively solve the problem that the robot is difficult to learn and store the relationship between objects in complex working scenes and improve the robot's ability in human-machine interaction, object inquiry, knowledge reasoning, and task execution.

Knowledge base aims to describe various entities or concepts and their relationships in the real world. Knowledge base is modeled as a set of triplets, with entities represented by  $e$  and relationships represented by  $r$ . In the knowledge graph, entities or concepts are represented by nodes and attributes or relationships are represented by edges, and thus, entities and relationships in the real world can be formed into a huge semantic network graph. The following gives the definitions of three kinds of nodes and edges included in the robot cognitive knowledge graph.

*Entity.* It refers to instances (people or objects, etc.) in the environment, such as a person, a table, and a gear. Entities are the most basic elements in the knowledge graph. Each entity has a unique number in the knowledge base to distinguish it from other entities.

*Concept.* It is a collection of entities with the same characteristics, for example, gears and bearings are both workpieces machined, and lathes and milling machines are both production equipment.

*Attribute.* It is used to describe certain characteristics of an entity, for example, the shape parameters and performance parameters of gears or bearings. In the cognitive knowledge graph, it exists in the form of edges, and attributes are measured by attribute values.

*Relationship.* It is formalized as a function, and it exists as an edge in the cognitive knowledge graph and is used to describe the relationship between entities or concepts in the graph.

Based on the above definitions, triplets are an information representation of cognitive knowledge graphs,

namely,  $G = (E, R, S)$ , where  $E = \{e_1, e_2, \dots, e_n\}$  is an entity set and  $R = \{r_1, r_2, \dots, r_n\}$  is a relational set. The ternary combination set of representation information  $S$  is  $S \subseteq E \times R \times E$ .

The construction of the cognitive knowledge base based on the cognitive knowledge graph is divided into the construction of the perception layer and the construction of the cognition layer. The flowchart is shown in Figure 1. The perception layer is mainly composed of a series of factual data, including a series of multisource heterogeneous original data and coded structured data. The original data include environmental data collected by robots through various sensors, data acquired by robots through a universal knowledge base on the Internet, and professional data stored in some professional databases. The cognition layer is mainly composed of a series of knowledge stored in units of triplets, such as triplets with gears in the accessory cabinet (gears, stored in, accessory cabinet). The cognition layer is built on the perception layer and is the core of the knowledge graph, storing abstract knowledge (concepts).

*3.1. Formation of Perception Layer.* The construction of the perception layer needs to firstly discover and add entities (i.e., find entities), so the data source of knowledge graph is the first problem faced by the self-construction of knowledge graph. According to the working environment of the industrial Internet, the robot can sense and collect images, videos, point clouds, and other data from the environment by carrying corresponding sensors for environmental object identification and space division, etc. However, a lot of tacit knowledge (such as functional attributes and operational attributes of objects) is difficult for the robot ontology to acquire through sensors. When people ask the robot to complete a nonspecific and complicated task of "opening the door," the robot neither knows the functional attribute of the key to open the door, nor does it know the operating attribute of the key to correctly insert the key into the lock hole, rotate twice, and then turn the doorknob counterclockwise by 90 degrees, so the task cannot be completed. Robots learn relevant knowledge from encyclopedia websites (e.g., Wikipedia), electronic product manuals, and shared structured data sources (e.g., relational databases of various specialties) and add them to the knowledge base, which is an effective way to solve the above problems. Due to the different data sources of the knowledge base, we have designed a knowledge extraction module to achieve the consistency of data formats. The structure diagram is shown in Figure 2, which mainly includes a triplet downloader, a document downloader, an entity and entity relation processing submodule, a triplet filter, a robot capability description module, and an environment information extraction module. The main functions are as follows:

- (1) The document downloader is designed based on crawler technology. Its main work is to capture the text of the web page and other text knowledge provided by the Internet. The robot grabs the text of the web page through the document downloader and downloads it to the local and processes the text of the

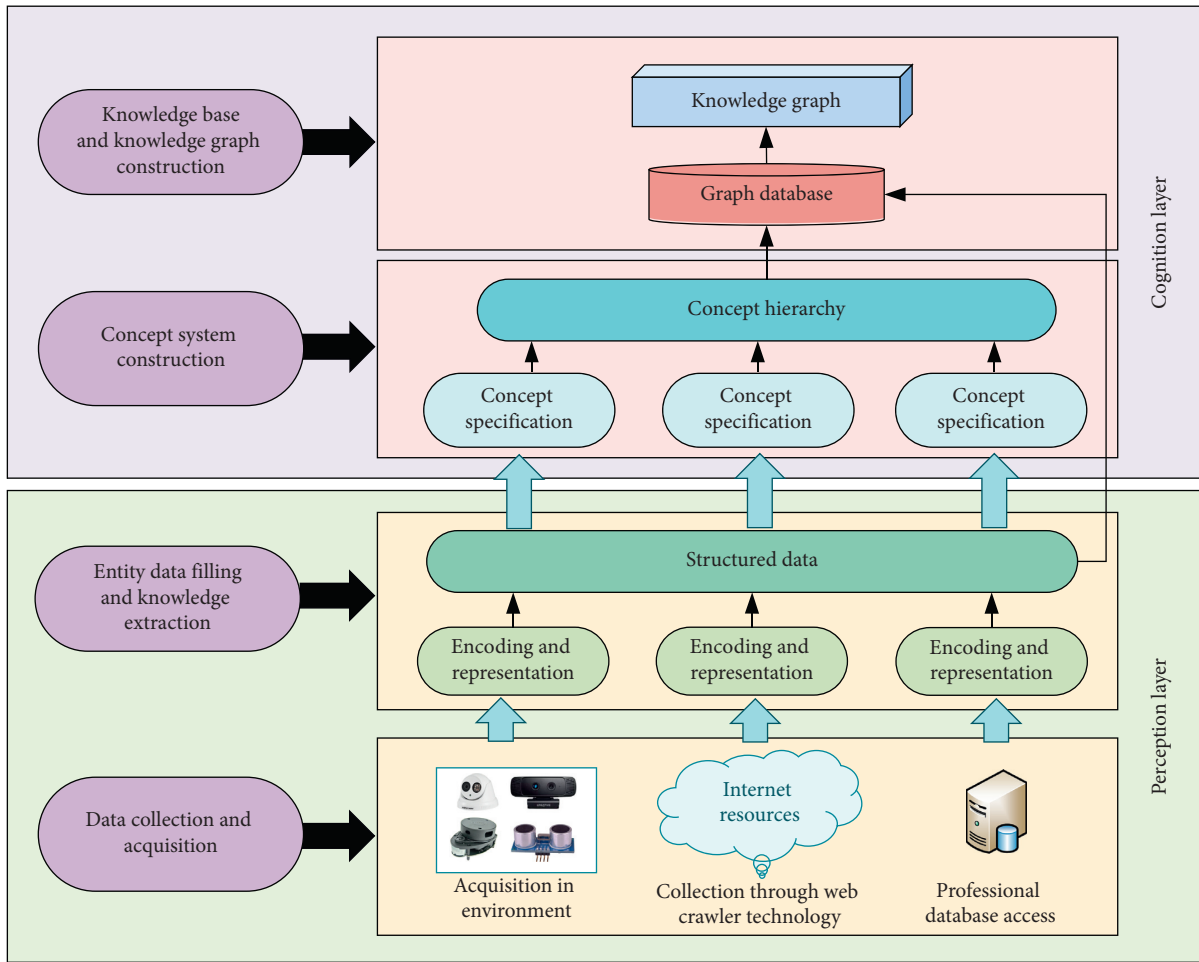


FIGURE 1: Flowchart of knowledge graph construction.

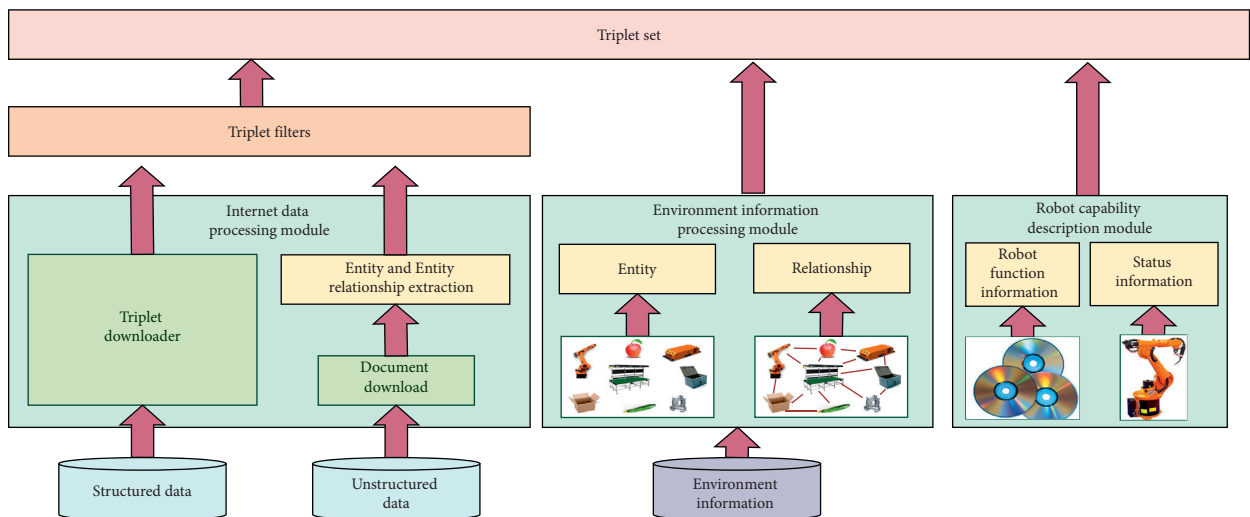


FIGURE 2: Logical structure of knowledge extraction.

web page to remove invalid data, thus obtaining the text data to be learned.

(2) Entity and entity relationship extraction module: it is based on the open-source toolkit CORENLP

developed by Stanford University for natural language processing. Through the use of the NER analysis module in CORENLP, lexical features of statements are analyzed to realize automatic

extraction of “text” entity relationships obtained by the above document downloader.

- (3) For structured data, there is no need to extract triplets through entity and entity relation module, and we can directly download triplets from the data source through the triplet downloader as candidate triplets to be added to the triplet candidate set.
- (4) Triplet filters: entities and entity relationships extracted by robots from unstructured and structured data inevitably contain repeated information. Through the triple filter module, repeated addition of information is avoided.
- (5) Environment information processing module: it obtains the location and attribution relationship between entities and entities in the environment through semantic SLAM technology [36] and spatial structured reasoning technology [37], respectively, for information perceived by the robot environment.
- (6) Robot capability description module: for the capability description information of the robot, the semantic description information of the capability of the robot in different states and the capability of operating different entities in the current environment is obtained through the description information of the robot and the current state information. For example, what objects can be operated on in the current environment and what operations can be performed on the objects under what posture.

Through the knowledge extraction module, the robot extracts entities and relationships from the working environment, its own state information, and unstructured and structured data and then generates triplets. However, the triplet set at this time is not the final knowledge graph, and entity disambiguation and entity alignment are needed for triplets in the set to realize the fusion of heterogeneous knowledge from different sources.

**3.2. The Formation of Cognition Layer.** Knowledge graph technology, like ontology technology, uses the triple set of nodes and relationships to directly model real-world scenes. Using the universal “language” triple set, we can effectively and intuitively represent the relationship between objects in the scene. In terms of content, compared with ontology, which emphasizes conceptual relationship, knowledge graph emphasizes not only the concept but also the entity relationship and the entity attribute value, which enriches and expands ontology knowledge as shown in Figure 3.

Therefore, the construction concepts in robot cognitive knowledge graph are similar to ontology, which is formed by identifying, classifying, and abstracting entities in the environment; we establish the up-down relationship between concepts according to the hierarchy and connection between concepts, thereby forming a conceptual hierarchy as shown in Figure 4. Therefore, the construction of the cognition layer in the cognitive knowledge graph is divided into two steps: the extraction of concepts and the establishment of the relationship between concepts.

**3.2.1. The Extraction of Concepts.** At present, most of the research studies on concept extraction in knowledge graph focus on the statistics of text corpus and extract candidate concepts in a certain field by calculating the frequency of words in various documents. This paper also uses statistical methods to extract candidate concepts in the working environment of industrial Internet. According to experts’ experience, a concept is a word or phrase that appears at a high frequency in the field, which can represent the characteristics of the field. Therefore, the concept of specific working environment of industrial Internet can be defined by the following two features: (1) the frequency of occurrence in specific application fields of industrial Internet is higher than that in other fields and (2) it is distributed evenly in the documents of specific application fields of industrial Internet, rather than concentrated in the documents of individual industrial Internet application fields.

For the abovementioned feature (1), we quantitatively describe it with domain relevance. Domain relevance, as its name implies, is used to describe the appropriateness of concepts and domains, and its calculation formula is as follows:

$$\text{Domain } R_{n,m} = \frac{P(n/D_m)}{\max P(n/D_i)}, \quad 1 < i < t. \quad (1)$$

Set of domain =  $\{D_1, D_2, \dots, D_m\}$ , and formula (2) shows the calculation of the correlation between concept  $n$  and domain. The conditional probability  $P(n/D_m)$  in the formula can be estimated by the following formula:

$$E\left(P\left(\frac{n}{D_m}\right)\right) = \frac{f_{n,m}}{\sum_{n' \in D_m} f_{n',m}}, \quad (2)$$

where  $f_{n,m}$  is the frequency of concept  $n$  in domain  $D_m$  and  $f_{n',m}$  is the frequency of concept  $n'$  in  $D_m$ . As shown in formula (2), the domain correlation of a concept is only proportional to its frequency in the domain and has nothing to do with other elements. When a certain concept frequently appears in individual domain documents, it will be ineffective to measure the concept only by using domain relevance, so feature (2) is needed to reflect the distribution of concepts in domain documents.

For feature (2), we use domain consistency to quantitatively describe it, and its calculation formula is as follows:

$$\begin{aligned} \text{Domain } C(n, D_m) &= H(P(n, d_j)) \\ &= \sum_{d_j \in D_m} \left( P(n, d_j) \log \frac{1}{P(n, d_j)} \right), \\ E(P(n, d_j)) &= \frac{f_{n,d_j}}{\sum_{d_j \in D_m} f_{n,d_j}}, \\ W_{n,m} &= \alpha \text{Domain } R_{n,m} + \beta \text{Domain } C_{n,m}, \quad \alpha, \beta \in (0, 1), \end{aligned} \quad (3)$$

where  $d_j$  is any document in the field,  $f_n$ , and  $d_j$  is the frequency of concept  $n$  appearing in document  $d_j$ . Formula (3) combines domain relevance and domain consistency,

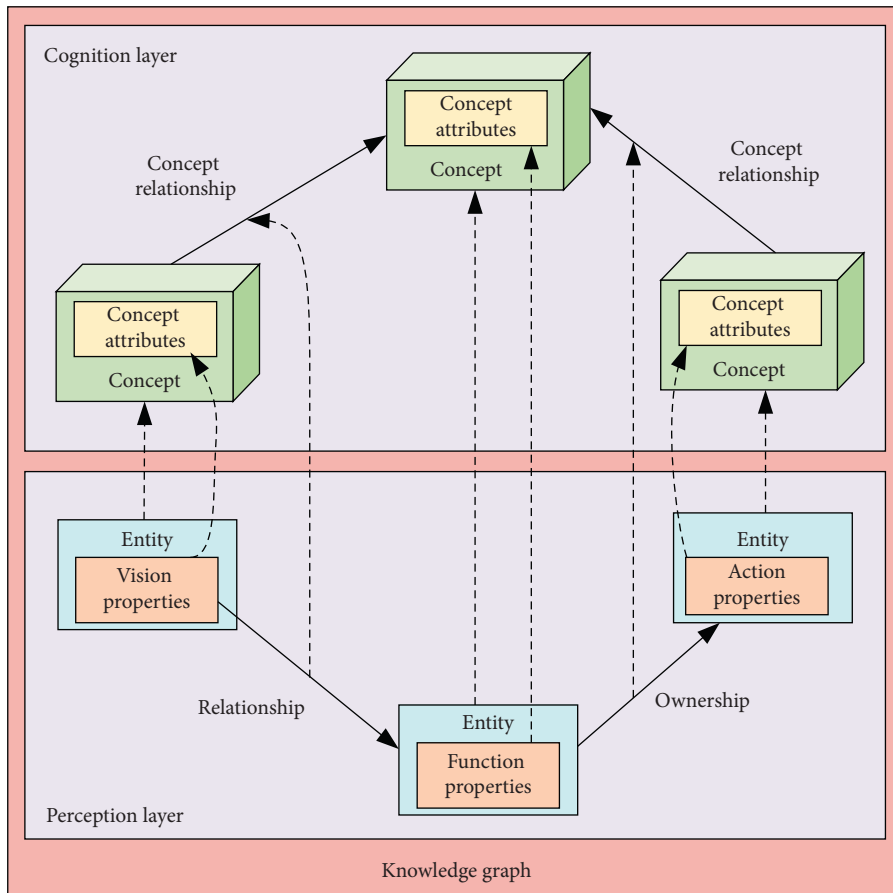


FIGURE 3: The relationship of knowledge graph and ontology.

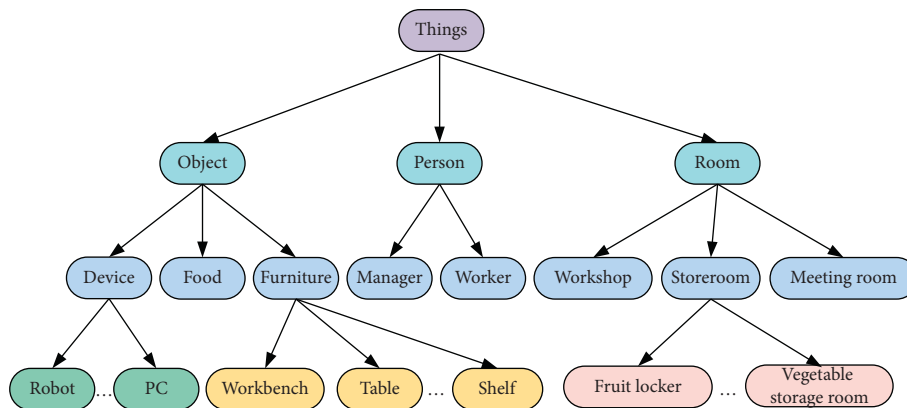


FIGURE 4: Hierarchical relationship of concepts.

and by setting related parameters  $\alpha$  and  $\beta$ , the probability of concepts unrelated occurrence to the domain can be reduced.

3.2.2. *The Establishment of the Relationship between Concepts.* In the previous section, we extracted the related concepts of specific application fields of industrial Internet, but these concepts are discrete and cannot be linked with each other. In the knowledge graph, concepts are connected by the

relationship between upper and lower levels, so we need to establish the relationship between upper and lower levels of these concepts. In people’s language usage habits, there are often such sentence patterns as “A is a B,” “A is like B and C,” and “A has B, C”. Using this kind of language model, we can establish the hyponymy relationship between A, B, and C. For example, “unmanned factories have production equipment such as lathes and milling machines”. According to the above linguistic schema, it can be deduced that the upper words of lathes and milling machines are production equipment.

Using the above linguistic sentence patterns, robots in industrial Internet can learn the upper and lower relations of concepts in encyclopedic text resources related to their specific application fields, thus connecting discrete concepts in series with the upper and lower relations to form a structured and hierarchical concept layer.

As shown in Figure 5, the robot combines the perceptual information perceived by the robot with the semantic description information such as attributes, concepts, and relationships of entities obtained from the Internet and establishes a structured reasoning knowledge base including the perception layer and the cognition layer, thereby realizing the unified representation and storage of the semantic information, attribute information, spatial function information, and context information of objects in the environment and generating the corresponding cognitive knowledge graph.

#### 4. Knowledge Fusion Algorithm Based on Vector Features

In the working environment of industrial Internet, when completing some unspecified complex tasks, robots usually need to know a variety of attributes of operated objects, such as position, function, and operation. The knowledge base with good performance and detailed content is the basis to solve this problem. In the process of building knowledge base, knowledge extraction technology realizes data acquisition. However, due to its wide data sources, the structure of these data is also very different; due to the freedom of natural language expression, polysemy often occurs. The robots based on this kind of knowledge base cannot generate unified and unambiguous cognition or complete some unspecified and complex tasks, so it is necessary to fuse the extracted knowledge.

In this paper, the differences of entity semantics and knowledge system in robot knowledge base are studied. To solve the problem of entity semantics inconsistent, this paper proposes an entity semantic fusion method, which calculates its word vectors through text information and description information of structured knowledge and determines the similarity of entity semantics according to the cosine similarity of word vectors, thus realizing the alignment and disambiguation of entity semantics. For the differences of knowledge systems, a framework of knowledge system integration based on HowNet is designed, which effectively solves the problem of knowledge system integration.

*4.1. Semantic Fusion of Multimodal Entities Based on Vector Features.* For the fusion of heterogeneous knowledge, a cross-modal data representation is needed first. In recent years, it is becoming more and more popular to represent various types of data in the form of vector features and to conduct corresponding research based on them. The most common vector feature learning is text domain, such as text embedded representation or distributed word representation. They are based on unsupervised learning and only rely on the input text corpus. The heterogeneous text entity

information is encoded into dense vector representation, and the similarity between two text entities is obtained by calculating the cosine similarity between entity vectors, thus completing the entity fusion.

Based on this idea, some structured data can also be encoded in low-dimensional vector space. For example, knowledge graph or ontology database (DBpedia and Wikidata), which is a network database, composed of entity-relation-entity or entity-predicate-entity. After the structured data composed of these entities, relations, or predicates are coded into low-dimensional vectors, it is easier to predict the links in the knowledge graph and reason the knowledge.

Therefore, it is a practical and effective method to preprocess text knowledge and structured knowledge by corresponding means and represent them in vector form and then realize multisource heterogeneous knowledge fusion by vector fusion.

*4.2. Vectorized Representation of Text Information.* For the vector transformation of text corpus knowledge, we choose word2vec model. word2vec is an open-source tool for word vector calculation released by Google in 2013, which has attracted the attention of industry and academia and become a popular application in the field of natural language.

Figure 6 shows the flow of generating word vectors based on word2vec method and obtains input and output word pairs by Skip-gram or CBOW (Continuous Bag of Words) model. At the same time, the input and output words are encoded in one-hot mode, which is taken as a training sample and brought into a neural network for training. The input matrix is multiplied by the output-hidden layer weight matrix, and the result is the word vector of the output word.

CBOW and Skip-gram are two different methods of word2vec. CBOW predicts the probability of a word according to its context. Skip-gram is just the opposite. In the process of solving word vectors, these two models actually play an important role in obtaining training samples needed for subsequent neural network model training.

Taking Skip-gram algorithm as an example, there are two parameters in this method: skip-window and num-skips, which, respectively, represent the number of contexts to be predicted and the number of output results. Taking the sentence "Robots walk around boxes to the workplace" as an example, assuming that the input word is box, skip-window and num-skips take 2 and 4, respectively, the final words containing box are (walk, around, box, to, the), and the result can be represented in the form of (input word, output word). Input all the words in the sentence as input words, and the output results are shown in Table 1.

The training samples as shown in Table 1 are obtained through the skip-gram model, but these training samples cannot be directly used for neural network training. One-hot coding is also needed to encode them to mark the position where words appear as 1 and the rest as 0. For example, the code of robots is [1, 0, 0, 0, 0, 0, 0] in the above example.

Enter the one-hot code of the input words into the neural network model as shown in Figure 7, and each neuron in the

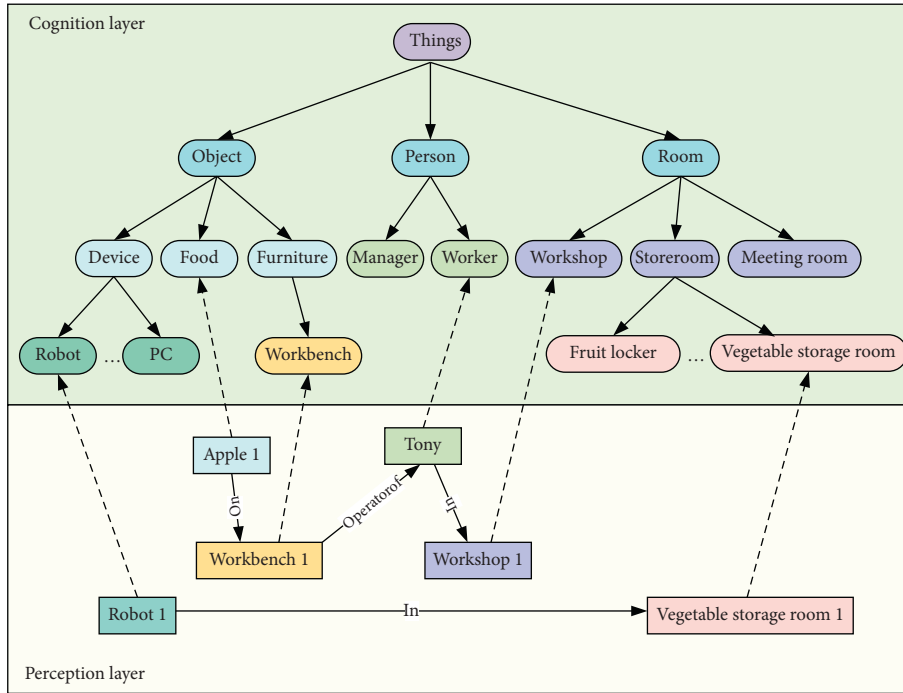


FIGURE 5: Schematic diagram of the cognition layer and perception layer in knowledge base.

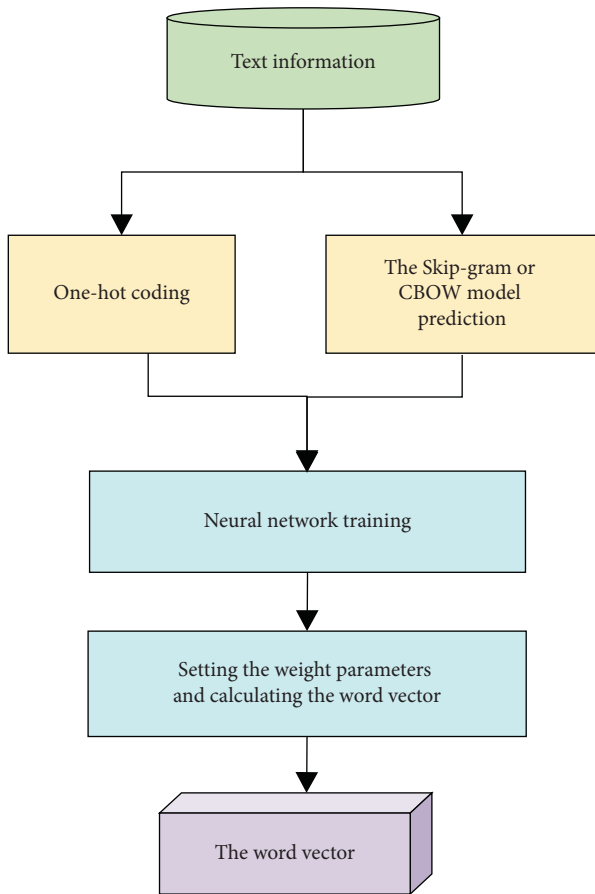


FIGURE 6: Flowchart of word vector formation.

input layer represents each bit in one-hot code. A hidden layer neural network is set up, and the one-hot coding passes through an input-hidden layer weight matrix and a hidden layer-output matrix to finally obtain the output of the neural network. We do not care about the output of the neural network, but the input-hidden layer parameter weight matrix after neural network training. The corresponding word vector can be obtained by multiplying the one-hot code of words with the parameter matrix.

**4.3. Vector Representation of Structured Knowledge.** TransE model is an algorithm to transform structured knowledge into vector features, which was proposed by Bordes [38] in 2013. Figure 8 shows a schematic diagram of the distributed vector representation model based on entities and relationships. Figure 9 shows the algorithm. The intuitive meaning of TransE is to take the relationship  $r$  in the triple set  $(h, r, t)$  as the translation from entity  $h$  to entity  $t$  and make  $(h + r)$  equal to  $t$  as much as possible by adjusting the vectors  $h, r,$  and  $t$ .

With a given triplet  $(h, r, t)$ , TransE model transforms the relation  $r$  into a vector  $r'$ , so that entities  $h$  and  $r$  can be connected in the form of vectors with less loss, and the distance function  $d(h + r, t)'$  is defined as follows:

$$d(h, r, t) = \|h + r - t'_2\|. \quad (4)$$

Formula (4) is used to measure the distance between  $h + r$  and  $t$ , and the hinge loss function is used to minimize it in the training process of the model. The hinge loss function is as follows:



TABLE 1: Training samples based on Skip-gram algorithm.

Input word	Output result 1	Output result 2	Output result 3	Output result 4
Robots	(Robots, walk)	(Robots, around)	—	—
walk	(walk, Robots)	(walk, around)	(walk, boxes)	—
around	(around, walk)	(Around, robots)	(around, boxes)	(around, to)
boxes	(boxes, around)	(boxes, walk)	(boxes, to)	(boxes, the)
to	(to, boxes)	(to, around)	(to, the)	(to, workplace)
the	(the, to)	(the, boxes)	(the, workplace)	—
workplace	(workplace, the)	(workplace, to)	—	—

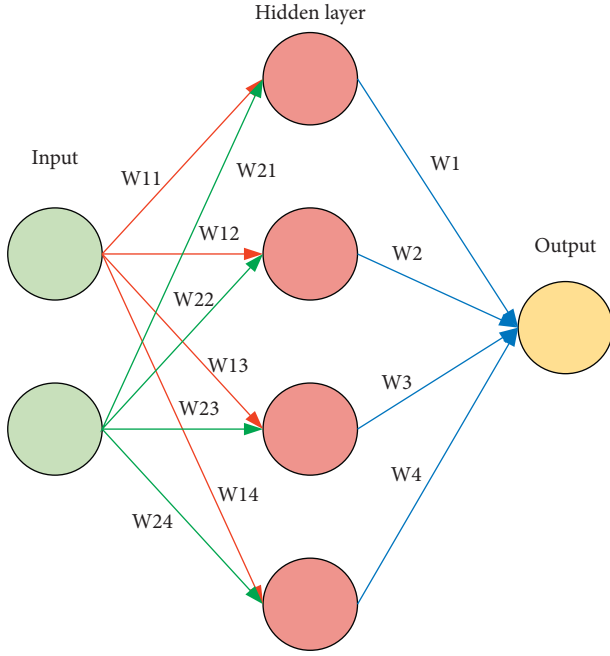


FIGURE 7: Neural network model.

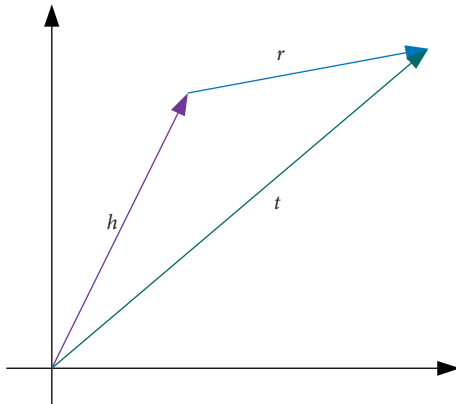


FIGURE 8: Schematic diagram of TransE model.

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r',t') \in S'} [\zeta + (h + r, t) - d(\dot{h} + r, \dot{t})]_+. \quad (5)$$

In formula (5),  $s$  is the triple set in the structured knowledge base, and  $S'$  is the triple with negative sampling.  $\zeta$  is the interval distance parameter, so the value is positive.

$[X]_+$  is a positive function, which is constant when  $x$  is a positive number and zero when  $x$  is a negative number.

4.4. *Semantic Fusion of Multimodal Entities Based on Singular Value Decomposition (SVD)*. The vector representation of entities based on word2vec and TransE model can solve the semantic problem of entities in single-modal data such as text and structuralization, but cannot solve the knowledge alignment problem between structuralization and text mode. To solve this problem, we get the vector representation set of entity description by word2vec coding the description text (such as color, shape, location, and category) of the entity in structured data as shown in Table 2 and get the text vector description of the entity by averaging the description vectors, so as to realize the entity alignment between structured data and text data, for example, the mobile phone “Apple” is distinguished from the fruit “Apple” by its description information “Block,” “Call,” and “Screen,” and its flow is shown in Figure 10.

After obtaining the multimodal vector representation, the vector dimension is large, which is not conducive to the subsequent similarity calculation. Therefore, it is necessary to reduce the dimension of vector features for calculation.

In this paper, singular value decomposition (SVD) is used to reduce the dimension of features. Input a matrix  $N$  can be decomposed into three matrices by SVD, as shown in the following formula:

$$N = A \Sigma B^T. \quad (6)$$

In formula (6),  $A$  and  $B$  are unitary matrices,  $\Sigma$  is a diagonal matrix, and the singular value of  $N$  on the diagonal decreases in descending order. By taking the first  $m$  columns and the first  $m$  singular values, we can get a new  $M$ -dimensional representation, thus realizing dimension reduction.

For multimodal vector representation, the similarity is calculated by calculating cosine similarity, and the calculation formula is as follows:

$$\text{similarity} = \cos(\theta) = \frac{X \cdot Y}{\|X\| \|Y\|} = \frac{\sum_{i=1}^n X_i \times Y_i}{\sqrt{\sum_{i=1}^n (X_i)^2} \sqrt{\sum_{i=1}^n (Y_i)^2}} \quad (7)$$

$X_i$  and  $Y_i$  in the formula are the components of vectors  $X$  and  $Y$ . The similarity of vectors calculated by the above

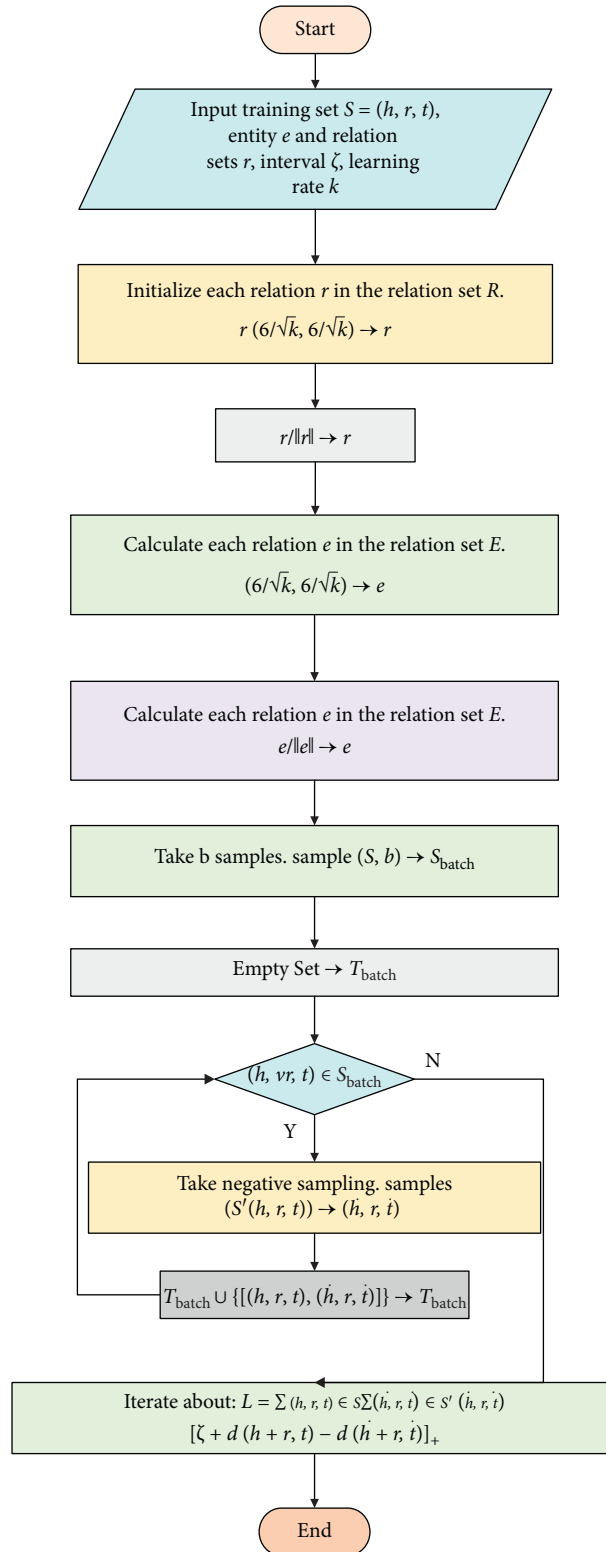


FIGURE 9: TransE algorithm.

formula ranges from  $-1$  to  $1$ . Among them,  $-1$  means that the entities represented by the two vectors are completely different and  $1$  means that they are two identical entities. The similarity is between  $-1$  and  $1$ , and its value represents the semantic similarity between the two entities.

4.5. Knowledge System Fusion Framework Based on HowNet. Semantic alignment of different names of the same entity can be solved by entity semantic fusion algorithm. However, because the knowledge description systems of different knowledge bases are different, some relational knowledge

TABLE 2: Description information of the structural database.

Entity semantic	Description information					.....
	Color	Shape	Weight (g)	Position	Category	
Cucumber	Green	Strip	110	Vegetable warehouse	Vegetable	.....
Apple	Red, green	Oblate	320	Fruit warehouse	Fruit	.....
Orange	Yellow	Oblate	160	Packaging room	Fruit	.....
Workbench	—	<i>n</i> -shaped	5	Packaging room	Equipment	.....

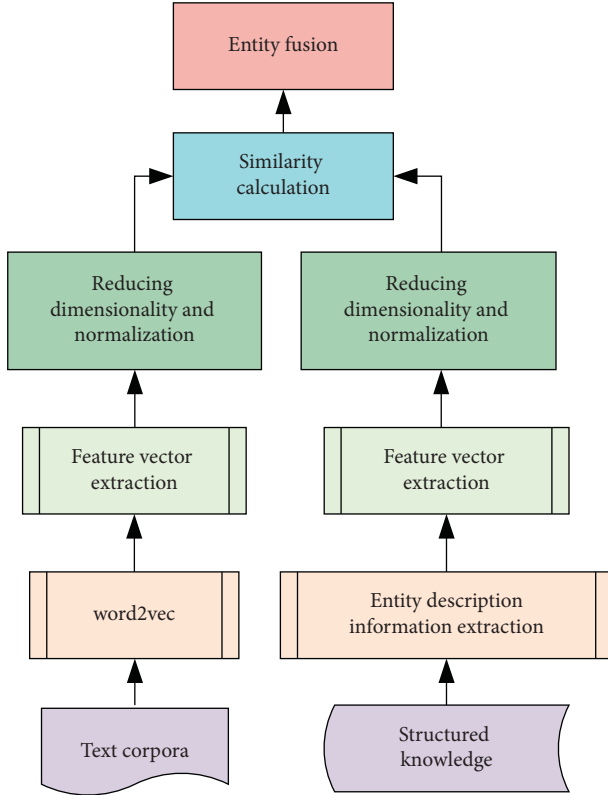


FIGURE 10: Flowchart of multimodal semantic fusion.

bases can only provide the relationship information between entities, while some text knowledge bases focus on the description of entity attributes. In view of the differences of knowledge systems in the above knowledge bases, a multisource semantic knowledge system fusion method is proposed based on Baidu Encyclopedia, Wikipedia, and HowNet semantic dictionary. By providing a unified hierarchical framework of “entity-category-attribute-attribute content” and “entity-relationship-entity,” entity-attribute template and entity-relationship template are established, which effectively solves the problem of multisemantic knowledge base fusion. According to the construction mechanism of knowledge graph mentioned above, the structure of the cognition layer and perception layer is easy to expand data, which makes the robot have the ability to build knowledge base automatically. Knowledge in the perception layer is stored in triples, so that entities, attributes, and relationships can better understand the semantic scope, improve the accuracy of robot item search, and enable robots to have the ability of related search. The flowchart of

knowledge fusion is shown in Figure 11, and its steps are as follows:

Step 1: import the entity set obtained by the knowledge extraction module into the knowledge graph perception layer, fuse the HowNet semantic dictionary, extract the semantic attribute information corresponding to the entities in the semantic dictionary, and then supplement and mark the fused entity attributes

Step 2: learn concept information such as categories from domain documents in encyclopedic knowledge base, use linguistic model to obtain the upper and lower position relationship of categories, use HowNet semantic dictionary to obtain attribute information of categories, and establish cognition layers of categories-attributes and categories-subclasses

Step 3: import the relation set to form a unified normative environmental knowledge representation specification based on the hierarchical framework of “entity-category- attribute-attribute content” and “entity-relationship-entity”

## 5. Experiments: Results and Discussion

5.1. *Multimodal Entity Semantic Fusion Experiment.* Based on the entity semantic alignment algorithm proposed in this paper, 50 groups of objects related to the robot working environment are extracted from YAGO and encyclopedia knowledge base, respectively. According to the algorithm in this paper, the entity text in encyclopedia knowledge base and the description information in the YAGO knowledge base are represented in the form of vector features. The evaluation metric of similarity calculation method of feature vectors uses the precision rate ( $p$ ), recall rate ( $r$ ), and  $F1$  values in entity semantic fusion, and their definitions are as follows:

$$P = \frac{TP + TN}{TP + TN + FP + FN},$$

$$R = \frac{TP}{TP + FN}, \quad (8)$$

$$F1 = \frac{2 \cdot RP}{R + P}.$$

The meanings of TP, TN, FP, and FN are shown in Table 3.

TP, FP, FN, and TN can be understood as follows:

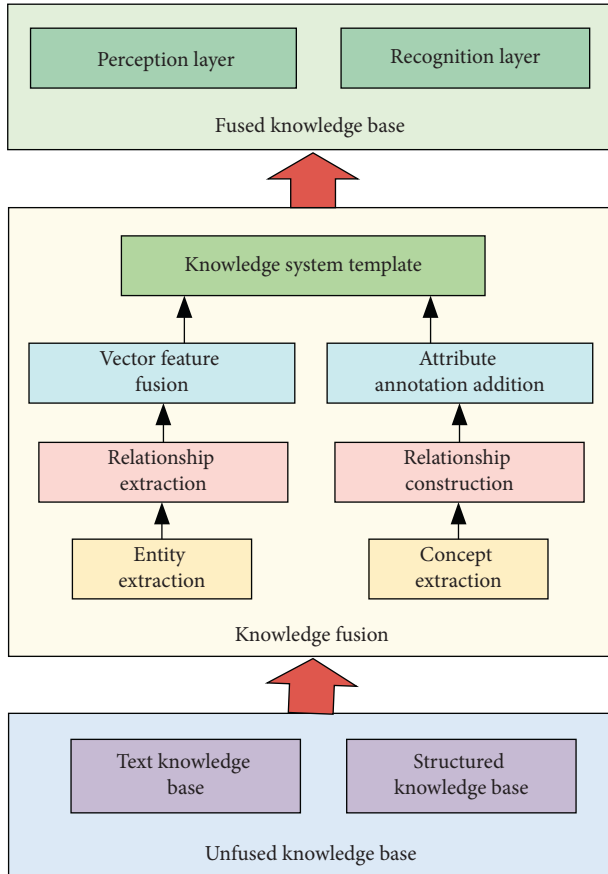


FIGURE 11: Flowchart of knowledge fusion.

- (i) TP: fusion semantic is 1 and actual semantic is 1, fusion correct
- (ii) FP: fusion semantic is 1 and actual semantic is 0, fusion error
- (iii) FN: fusion semantic is 0 and actual semantic is 1, fusion error
- (iv) TN: fusion semantic is 0 and actual semantic is 0, fusion correct

To verify the effectiveness of the algorithm in this paper, we selected the data of eight common fruits and vegetables on the fruit and vegetable sorting line for testing, set the confidence threshold to 0.9, and obtained the confusion matrix of entity semantic fusion accuracy as shown in Figure 12. It can be seen from the confusion matrix that, except tangerines and oranges, whose description information is too similar, which leads to lower accuracy, the accuracy of other semantic fusion has reached a higher level, and the overall performance is good. To further verify the algorithm in this paper, 50 sets of entity semantics and corresponding description information are selected to expand the test data. According to the above evaluation metric of accuracy and  $F1$  value, the test results obtained by experiments are shown in Figure 13, with blue broken line as the accuracy result and orange broken line as the  $F1$  value result. By observing the semantic fusion results of 50 groups of test samples, the accuracy of the semantic fusion method adopted in this chapter

and the average value of  $F1$  value reach about 70%, and the semantic fusion effect is ideal.

**5.2. Comparative Experiment on Dimension Reduction of Entity Semantic Feature Vector.** At the same time, in order to improve the efficiency of entity semantic similarity calculation, we choose the singular value decomposition method to reduce the dimension of vector and compare it with the vector without dimension reduction in running time and accuracy. The test results obtained by experiments are shown in Figure 14.

By comparing the accuracy curves in Figure 14, we can know the accuracy of entity semantic fusion without dimension reduction and with dimension reduction. Although the accuracy of vector with dimension reduction is slightly lower than that without dimension reduction, they are roughly at the same level. By comparing the running time curves, it can be seen that although the running time of some items in the dimension-reduced vector semantics exceeds that of the no-dimension-reduced vector semantics, the fusion time of the dimension-reduced semantics is much lower than that of the no-dimension-reduced vector semantics. Generally speaking, through the dimension reduction operation of vectors, the matching time is greatly reduced while the success rate of vector matching is kept, and the efficiency of entity semantic matching is improved.

**5.3. Cognition Experiment.** The main purpose of this experiment is to verify whether the intelligent decision-making ability of the mechanical arm in processing nonspecific complex tasks can be improved after integrating professional database information, functional information, and real-time status information of the mechanical arm into the industrial Internet when constructing the knowledge graph. In the experiment, we placed different fruits and vegetables in three areas around the arm.

As shown in Figure 15, area 1 is an operable area where the mechanical arm can accurately identify the type of items and grab them flexibly. Area 2 is the critical area around the farthest position where the arm can reach the object. Area 3 is the area where the arm must not be able to catch objects. One watermelon and one apple are placed in each of the three areas. We chose a mechanical arm with a grip of 3 kg for the test. The mechanical arm determines the distance of the object by using a depth camera. The instruction we gave was just "take all the fruit and vegetables". During the experiment, each time the arm made a judgment for a grab, we recorded whether it was successful or not and then changed the locations of the watermelon and apple in three areas. A total of 50 experiments were conducted, and the experimental results are listed in Table 4.

In Table 4, the success rate of the mechanical arm in judging apples and watermelons grab in different areas is, respectively, calculated. We found that no matter where the mechanical arm recognized the watermelon, it did not grab it. It shows that the arm knows that it cannot catch watermelon through its knowledge graph. Only in area 3, there were two errors in judgment because the distance was too far and the

TABLE 3: Meaning of TP, TN, FP, and FN.

		Fusion semantic	
		1	0
Actual semantic	1	TP	FP
	0	FN	TN

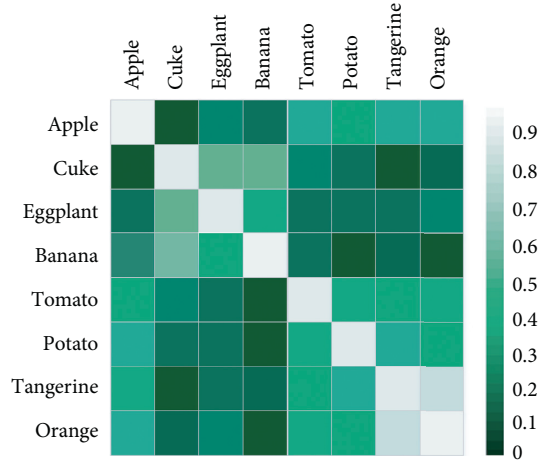


FIGURE 12: Confusion matrix of entity semantic fusion.

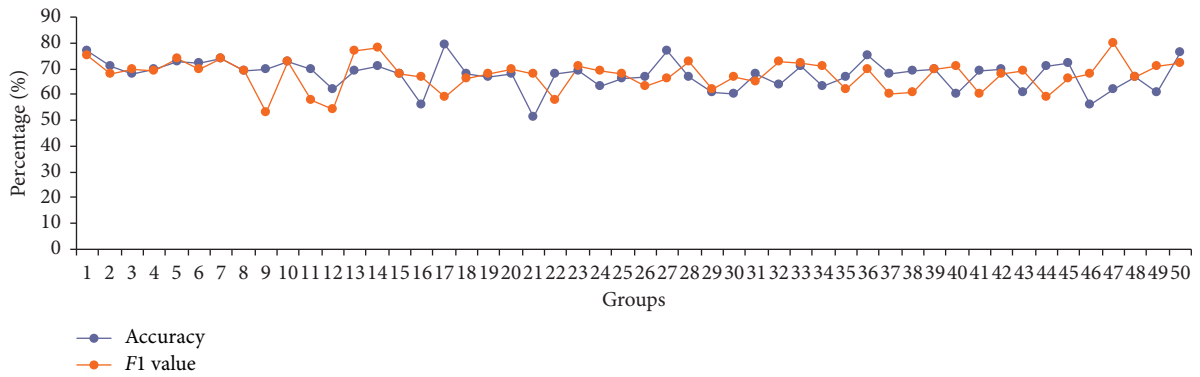
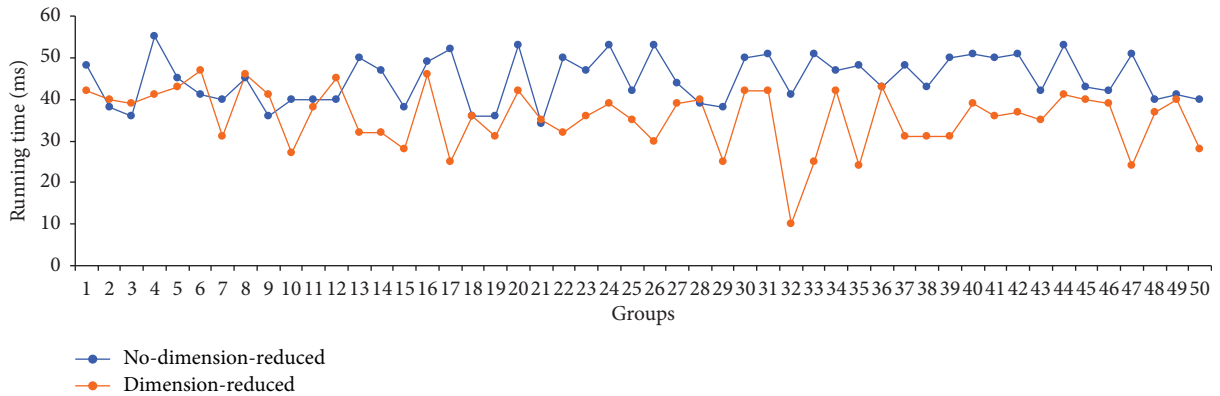
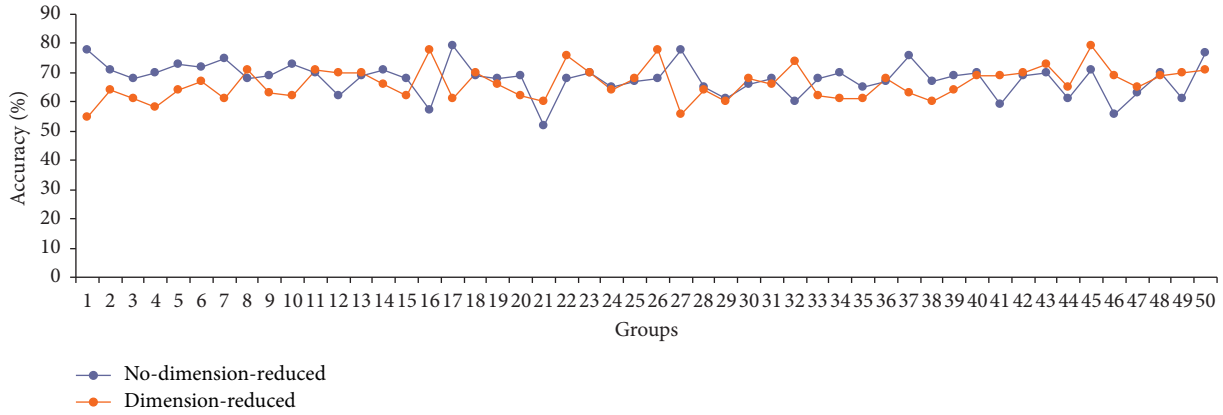


FIGURE 13: Line chart of accuracy and F1 value.



(a)

FIGURE 14: Continued.



(b)

FIGURE 14: Comparison curve of time and accuracy of entity semantic fusion without dimension reduction and dimension reduction. (a) Comparison of running time and (b) comparison of accuracy.

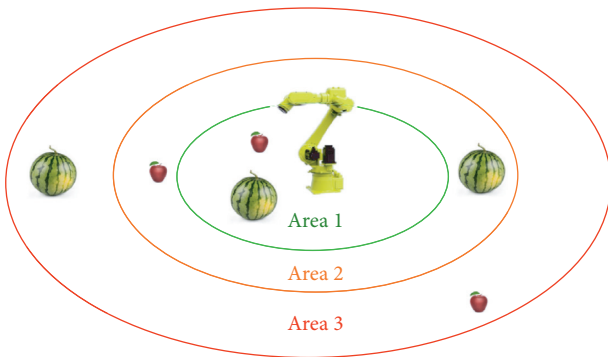


FIGURE 15: Grab the object area diagram.

TABLE 4: Success rate statistics of grab.

Area	Apple (%)	Watermelon (%)
Area 1	98	100
Area 2	68	100
Area 3	90	96

object recognition was wrong. The success rate of recognizing and grabbing apples is the highest in region 1. Only once after successful recognition, grab fails, and the apple slipped. The success rate of recognizing and grabbing apples in region 2 is only 68%. The reason is that when apples are in critical position, the distance judgment accuracy of the mechanical arm is limited, and there are often errors in position recognition. In area 3, the success rate of apple recognition and grab was also high, only 5 apples failed to grab because all the apples in the distance were judged correctly and did not try to grab, while there were 5 apples in area 3 placed close to the critical area, and the arm misjudged that the apples could be caught, resulting in the failure of grab.

## 6. Conclusions and Future Work

This paper mainly studies how to construct the representation model of environmental information in industrial Internet, proposes a representation model of cognitive information based on knowledge graph, gives the construction process of the model, and describes the detailed construction of the perception layer and cognition layer. After the environmental information is represented by the ontology awareness information and network resource information of the industrial robot and stored in the form of knowledge base, the industrial robot can plan the task according to the position, function, and operation of the items in the task and then perform intelligent operation accurately and efficiently. In addition, we also analyzed the necessity of knowledge fusion and described the problems in data fusion. A concrete fusion scheme is given, which includes entity semantic fusion and knowledge system framework building. The semantic fusion of multimodal entities based on vector representation is realized. It provides a unified hierarchical framework of “entity-category-attribute-attribute content” and “entity-relationship-entity,” which effectively solves the problem of inconsistent knowledge systems in the fusion of multisemantic knowledge base.

In addition, after integrating professional database information, function information, and real-time status information of the mechanical arm into the industrial Internet when constructing the knowledge graph, the ability of the mechanical arm to handle nonspecific complex tasks and the level of intelligence are improved.

This paper has made some achievements in environmental information representation and knowledge fusion, but in the process of knowledge fusion in knowledge base, only the semantic fusion of items is considered, and there is no research on the fusion of items’ relationships and attributes. Therefore, some algorithms related to the fusion of items’ attributes and relationships can be considered to achieve a higher level of knowledge fusion in the future.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

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

## Acknowledgments

This work was supported in part by the Key R & D Program of Shaanxi Province (no. 2019ZDLGY03-04), Xi'an Science and Technology Plan Project (Program nos. GXYD19.8 and GXYD19.7), and the National Natural Science Foundation of China (no. 61303225), and the Space Science and Technology Fund (Program no. D5120200106).

## References

- [1] T. Kollar, V. Perera, D. Nardi, and M. Veloso, "Learning environmental knowledge from task-based human-robot dialog," in *2013 IEEE International Conference on Robotics and Automation*, pp. 4304–4309, Karlsruhe, Germany, May 2013.
- [2] S. Lemaignan, R. Ros, R. Alami, and M. Beetz, "What are you talking about? grounding dialogue in a perspective-aware robotic architecture," in *2011 RO-MAN*, pp. 107–112, Atlanta, GA, USA, July 2011.
- [3] S. Lemaignan, R. Ros, E. A. Sisbot, R. Alami, and M. Beetz, "Grounding the interaction: anchoring situated discourse in everyday human-robot interaction," *International Journal of Social Robotics*, vol. 4, no. 2, pp. 181–199, 2012.
- [4] S. Schiffer, N. Hoppe, and G. Lakemeyer, "Natural language for an interactive service robot in domestic domains," *Communications in Computer & Information Science*, vol. 358, pp. 39–53, 2012.
- [5] C. Burghart, R. Mikut, R. Stiefelhagen et al., "A cognitive architecture for a humanoid robot: a first approach," in *5th IEEE-RAS International Conference on Humanoid Robots, 2005*, pp. 357–362, Tsukuba, Japan, December 2005.
- [6] O. Mozos, Z.-C. Marton, and M. Beetz, "Furniture models learned from the WWW," *IEEE Robotics & Automation Magazine*, vol. 18, no. 2, pp. 22–32, 2011.
- [7] D. W. Ko, C. Yi, and I. H. Suh, "Semantic mapping and navigation with visual planar landmarks," in *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 255–258, Daejeon, South Korea, November 2012.
- [8] C. Song, *Research on Building Intelligent Semantic Map of Indoor Robot Based on Distributed Representation*, Beijing University of Chemical Technology, Beijing, China, 2015.
- [9] L. Tamas and L. Cosmin Goron, "3D semantic interpretation for robot perception inside office environments," *Engineering Applications of Artificial Intelligence*, vol. 32, pp. 76–87, 2014.
- [10] H. Wu, G. Tian, X. Chen, T. Zhang, and F. Zhou, "Map building of indoor unknown environment based on robot service mission direction," *Robot*, vol. 32, no. 2, pp. 196–203, 2010.
- [11] H. Wu, "Characterization of large-scale unknown environmental information based on RFID technology," *Journal of Central South University (Science and Technology)*, no. S1, pp. 166–170, 2013.
- [12] W. Park, M. Han, W. J. Son, and S. J. Kim, "Design of scene knowledge base system based on domain ontology," in *2017 19th International Conference on Advanced Communication Technology*, pp. 560–562, IEEE, Bongpyeong, South Korea, February 2017.
- [13] Q. Hao, Y. Li, L. M. Wang, and M. Wang, "An ontology-based data organization method," in *2017 Fifth International Conference on Advanced Cloud and Big Data (CBD)*, pp. 135–140, Shanghai, China, August 2017.
- [14] Y. Yang, Z. Wu, and X. Zhu, "Semi-automatic metadata annotation of web of things with knowledge base," in *2016 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pp. 124–129, Beijing, China, September 2016.
- [15] P. Das, V. Hilaire, and L. Ribas-Xirgo, *An Ontology to Support Collective Intelligence in Decentralised Multi-Robot Systems*, <https://arxiv.org/abs/1806.00367>, 2018.
- [16] P. Jezek and R. Moucek, "Semantic framework for mapping object-oriented model to semantic web languages," *Frontiers in Neuroinformatics*, vol. 9, no. 3, p. 3, 2015.
- [17] Y. Chen, *Research on Indoor Service Robot Navigation in Large-Scale Complex Scenes*, University of Science and Technology of China, Hefei, China, 2017.
- [18] S. Gao, L. Kong, and P. Wu, "Autonomous processing method of Chinese service instruction for indoor intelligent robot," *Robot*, no. 4, pp. 424–434, 2015.
- [19] A. Singhal, *Introducing the Knowledgegraph: Things, Not Strings*, Official google blog, 2012.
- [20] *Sohu Encyclopedia*, 2015, <http://baike.sogou.com/lh6b61b234.htm>.
- [21] *Baidu Encyclopedia*, 2015, <http://baike.baidu.com/view/10972128.htm>.
- [22] K. Marino, R. Salakhutdinov, and A. Gupta, "The more you know: using knowledge graphs for image classification," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20–28, Honolulu, HI, USA, 2017.
- [23] P. Szekeley, C. A. Knoblock, J. Slepicka et al., *Building and Using a Knowledge Graph to Combat Human Trafficking*, pp. 205–221, Springer International Publishing, NY, USA, 2015.
- [24] Y. Zhang and M. Liu, "Visualization analysis of sports industry research in China based on knowledge map," *China Sport Science and Technology*, vol. 52, no. 1, pp. 24–29, 2016.
- [25] Q. He, "Visualization analysis of physical education research field in China based on knowledge map," *Journal of Beijing Sport University*, vol. 39, no. 2, pp. 98–103, 2016.
- [26] Z. Han, B. Li, K. Zhang et al., "Knowledge structure of China's marine economy research: an analysis based on CiteSpace map," *Scientia Geographica Sinica*, vol. 36, no. 5, pp. 643–652, 2016.
- [27] Y.-H. Xiao, K.-Z. Zhang, and W. Wang, *A Knowledge Graph Construction Method in the Field of Books Reading*, China.
- [28] G. Jin, F. Lü, and Z. Xiang, "Enterprise information integration based on knowledge graph and semantic web technology," *Journal of Southeast University (Natural Science Edition)*, vol. 44, no. 2, pp. 250–255, 2014.
- [29] L. Jia, J. Liu, Y. Tong et al., "Construction of TCM knowledge map," *Journal of Medical Informatics*, vol. 36, no. 8, pp. 51–53+59, 2015.
- [30] X. Lu, J. Zeng, M. Zhang, X. Guo, and J. Zhang, "Using knowledge graphs to optimize MOOC instruction," *Distance Education in China*, no. 7, pp. 79–80, 2016.
- [31] A. Kumar, C. Schmidt, and J. Köhler, "A knowledge graph based speech interface for question answering systems," *Speech Communication*, vol. 92, pp. 1–12, 2017.

- [32] Y. Jia, Y. Qi, H. Shang, R. Jiang, and A. Li, "A practical approach to constructing a knowledge graph for cybersecurity," *Engineering*, vol. 4, no. 1, pp. 53–60, 2018.
- [33] O. Kem, F. Balbo, A. Zimmermann, and P. Nagellen, "Multi-goal pathfinding in cyber-physical-social environments: multi-layer search over a semantic knowledge graph," *Procedia Computer Science*, vol. 112, pp. 741–750, 2017.
- [34] D. Collarana, M. Galkin, I. Traverso-Ribón, C. Lange, M.-E. Vidal, and S. Auer, "Semantic data integration for knowledge graph construction at query time," in *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, pp. 109–116, San Diego, CA, USA, January 2017.
- [35] W. Hao, J. Menglin, T. Guohui, M. Qing, and G. Liu, "R-KG: a novel method for implementing a robot intelligent service," *Artificial Intelligence (AI)*, vol. 1, no. 1, pp. 117–140, 2020.
- [36] J. McCormac, A. Handa, A. Davison, and S. Leutenegger, "Semantifusion: dense 3D semantic mapping with convolutional neural networks," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, Marina Bay Sands, Singapore, May 2017.
- [37] Z. Wang, "RGB-D scene parsing based on spatial structured inference deep fusion networks," *Chinese Journal of Electronics*, vol. 5, pp. 1253–1258, 2018.
- [38] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in Neural Information Processing Systems*, vol. 26, pp. 2787–2795, 2013.



## Research Article

# Research on the City Network Structure in the Yellow River Basin in China Based on Two-Way Time Distance Gravity Model and Social Network Analysis Method

Duo Chai , Dong Zhang, Yonghao Sun, and Shan Yang

School of Government, Central University of Finance and Economics, Beijing 100081, China

Correspondence should be addressed to Duo Chai; [chaiduobnu@126.com](mailto:chaiduobnu@126.com)

Received 8 October 2020; Revised 4 December 2020; Accepted 11 December 2020; Published 22 December 2020

Academic Editor: Fei Xiong

Copyright © 2020 Duo Chai 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.

Modern cities form city networks through complex social ties. City network research is widely applied to guide regional planning, infrastructure construction, and resource allocation. China put forward the Yellow River Basin Development Strategy in 2019, but no research has been conducted on regional social connections among cities. Based on the gravity model modified by two-way “time distance” between cities, this is the first study to empirically examine the intensity and structure of the entire city network in the Yellow River basin using the social network analysis method and ArcGIS software. The connection rules of the cross-city transfer of city officials in the basin are also investigated to illustrate the official ties between cities. The results suggest that the intensity of two-way connections between cities is generally low in the Yellow River basin and there is a positive correlation between city network development level and regional economic development level. The development gap between cities on the north and south banks is larger than that between the east and west regions, and some cities in the middle and upper reaches of the river are marginalized in the network. The status of the central cities in the Yellow River basin is distinct, but their connecting and leading abilities are not strong, showing an inverted T-shaped spatial distribution. The subgroups of city networks have strong internal connections, while the connection among subgroups is weak and the development shows a partitioned and fragmented pattern, making it difficult to form linkages among the upper, middle, and lower reaches. The “beaded chain” spatial development strategy can be adopted in the river basin planning, giving priority to strengthening the links within subgroups of cities and among adjacent subgroups, building central city chains, and reinforcing the overall basin management.

## 1. Introduction

Modern metropolises and towns of different sizes have formed “macro-” social networks with various complex ties in politics, economy, and culture and have become a special field of social network research [1, 2]. The main research contents of the analysis and graph mining of city networks include identifying the connection among cities, measuring the intensity of city network connections, and depicting the forms of city networks [3]. It is an effective method to reveal regional economic patterns, population distribution, and land development intensity and has been widely used in regional, spatial, and industrial planning, infrastructure construction layout, and other fields [4]. Over the past 40 years, cities in China have experienced rapid growth.

According to the *China Statistical Yearbook*, from 1978 to 2019, China’s urbanization rate increased from 17.9% to 60.6%, the number of cities increased from 193 to 672, the urban population increased from 170 million to 840 million, and the urban construction land area increased from 6,720 square kilometers to 56,075.9 square kilometers. Driven by infrastructure construction such as highways, railways, and civil aviation systems, connection among cities has been continuously strengthened, forming many complex city networks [5]. The rapid formation of city networks in China provides an excellent subject for studies in city networks. In particular, research on city networks in river basins is a topic of significant current interest.

China’s Yellow River basin includes nine provinces or autonomous regions, namely, Qinghai, Sichuan, Gansu,

Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan, and Shandong. Historically, it is China's earliest developed area, and the river is known as the "Mother River." However, its resources and ecology are fragile. Since China's reform and opening-up, priority has been given to supporting the development of the southeast coastal regions, while cities in the Yellow River basin have been significantly neglected, becoming one of the areas in China with the most prominent problems such as unbalanced economic development, resource depletion, environmental pollution, and poverty [6]. To revitalize the Yellow River basin, China put forward a series of regional development plans, such as the "Great Western Development Strategy" in 2000 and the "Rise of Central China Plan" in 2004, strengthened investment in ecological protection and poverty alleviation, and built high-speed railways and highways on a large scale, which resulted in gradual economic growth and improved population movement among cities in the basin [7]. In September 2019, China introduced the "Ecological Protection and High-Quality Development of the Yellow River Basin" as a national strategy [8], in which building a high-quality city system in the basin and promoting balanced urban development are key objectives [9]. However, at present, there is hardly any research on the overall city network in the Yellow River basin to provide guidance for urban system development.

By using the two-way "time distance" modified gravity model, the social network analysis method, and ArcGIS software, this study took the cities in the Yellow River basin as a whole to study the spatial structure patterns and the intensity of the complex city network in the Yellow River basin. In addition, the official transfer network characteristics among cities were brought into the city network research for the first time, to illustrate political connections and cooperation possibilities among cities.

## 2. Related Work

Previously, research on city networks mainly studied the organizational structure of cities through road connections. Ardekani and Herman [10] studied the evolution of the connection structure between Austin and Dallas, two cities in Texas in the USA, using the technology of time-lapse aerial photography and the "two-fluid model," combined with ground experiments. Smith and Timberlake, Taylor et al., and Choi et al. [11–13] gradually introduced intercity connecting factors such as business, capital, and the Internet to improve the measurement methods for the structure, hierarchy, and density of city networks. Jarmon and Vanderleeuw together with Neal [14, 15] designed the decision network analysis method, the interlocking world city network model, and the multisector network analysis method.

Among all types of city networks, the ones in river basins are unique. Because of connecting factors such as water conservancy, the environment, shipping, and population migration from the middle and upper reaches of rivers in inland areas to coastal areas, cities in river basins have a natural association with each other. Research on city networks can provide direct guidance for the overall river basin

development planning [16]. Varol et al. together with Jennifer and Angel [17, 18] studied the Tigris and Euphrates River basins in Turkey and the Mississippi River basin urban agglomeration in the USA but still used the traditional geospatial distance analysis method. In the river basins' city networks studies in China, Wang and Zhuang in addition to Li et al. [19, 20] applied the methods of comprehensive scale analysis and diffusion potential analysis to study the urban agglomerations in the Yangtze River basin and the Pearl River basin. Wu et al. [21] applied an analytic hierarchy process and spatial autocorrelation to identify the functional nodes and network structure of the Haihe River basin urban agglomeration. Wang et al. [22] replaced the "linear distance" with the "shortest arrival time" between cities as the indicator of the strength of ties between cities to modify the gravity model. By using the modified gravity model and the aid of geographic information system technology (GIS), the spatial structure of the city network in the middle reaches of the Yangtze River was studied. The modified gravity model was enlightening for this study and will be further improved by considering the two-way arrival time difference between cities and comparing the shortest arrival time of highway and high-speed railways to determine the "time distance" between cities.

At present, there is hardly any research on the overall city network in the Yellow River basin, with only a few river basin sections studied. There is significantly more research on cities in the middle and lower reaches of the basin than on inland (relatively backward) cities in the upper reaches. Research on city networks in the lower reaches of the Yellow River has mainly focused on the Shandong Peninsula. Li [23] applied the information flow of "Baidu Index" to analyze the trend of decreased centrality and "flattened" development of the city network in this peninsula. C. Wang and M. Wang [24] analyzed the trend differences of increased or decreased concentrations of the connections between cities in different regions of Shandong Province. Cheng and Wang [25] investigated the mechanisms by which spatial distance and administrative level influence the city network structure in this province. Research on the city network in the middle reaches of the Yellow River has focused on the Central Plains urban agglomeration in Henan Province. Zhao et al. [26] constructed a Voronoi diagram of prefecture-level cities in Henan Province and, based on the analysis of the number of coordinate cities to each central city and their spatial distance, identified the dense and "empty" areas of the city network. Miao and Wang, Meng and Lu, and Song and Liu [27–29] introduced transportation links and industrial links in their research on the city network of Zhengzhou urban agglomeration. In recent years, Geng and Li [30] have studied the city network structure of Shanxi Province and Shaanxi Province in the middle reaches of the Yellow River based on migration and communication data from Baidu.com. In the case of cities in provinces or autonomous regions in the upper reaches of the Yellow River, such as Gansu, Qinghai, Ningxia, and Inner Mongolia, very little research has been conducted on the connection structure between cities.

Generally, there are three shortcomings in the existing research on city networks in the Yellow River basin. First, there is a lack of research on the overall river basin city network, especially those in the middle and upper reaches. Second, most studies carried out retrospective analyses on the intercity connection intensity based on indicators such as the distance between cities, the number of traffic lines, and the number of people who migrated in the past, but seldom used indicators such as the commuting time between cities that are closer to reality and reflect the current situation and potential closeness of communication between cities. In reality, although there was little connection among cities in the past such as through population movement, due to the large-scale construction of transportation facilities (e.g., high-speed railways) and logistics and information networks in China, commuting time between cities has been shortened significantly, and there may be much room for growth in future connection. As a result, former indicators could not comprehensively measure the potential for intercity connection, nor were they sufficient to guide future regional development planning and facility construction. The optimal evaluation index for the intensity and convenience of intercity transportation links has changed from “distance and number of passages” to “shortest arrival time.” Third, existing studies have a common subjective error; that is, they assume that “City A to City B” and “City B to City A” have the same intensity of commuting connection, without distinguishing the forward and backward directions of intercity connection but calculating only a one-way connection. In reality, the lengths of intercity commuting time in the two directions are not necessarily the same, so it is necessary to consider all cities as individually equal and calculate the degree of the two-way connection.

### 3. Materials and Methods

**3.1. Theoretical Analysis.** The social connections between cities are very complex, and it is usually necessary to identify the main alternative indexes reflecting the intensity of connection [31]. By referring to the existing research, this paper makes the following inferences: the intensity of connection between two cities in terms of the economy and population movement mainly depends on two aspects: one is the size of the economy and population of the two cities. The larger the sizes of the two cities, the more the total volume and types of population and economic connection that could occur between them. The other is the commuting connection between the two cities [5, 22, 32]. The shorter the commuting time is, the stronger the connection is. Among them, commuting connection is a comprehensive indicator: first, commuting connection is the most important channel for intercity population movement and various political, economic, and cultural connections. Second, the construction of the existing commuting network between cities is based on the actual connection needs of population flow and logistics between the cities in the past and also reflects the overall arrangement and policy intention of the state for regional development. Third, intercity commuting connections can comprehensively reflect the present situation

and future growth potential of the intercity population, logistics, and trade connections [33].

In the past, most studies selected the geographical distance or road distance between cities as the index to calculate commuting connection and held that the connection intensity between cities is inversely proportional to the square of commuting distance [27, 34]. However, because the two-dimensional straight-line distance on a planar map does not take into account the factor of terrain, the fundamental changes in commuting efficiency between cities by the coverage of modern high-speed railways, and the difference in intercity commuting time in the two directions, distance between cities cannot accurately describe the actual intensity of the commuting connection between cities [35].

In view of these shortcomings, this paper takes the whole urban agglomeration of the Yellow River basin as the research object, consults the shortest arrival times of “City A to City B” and “City B to City A” one by one on Google Maps (highways) and the website of China Railway Customer Service Center, and constructs the shortest “time distance” matrix of two-way connections between cities. Then, the gravity model is constructed by introducing the indicators of urban population and economic aggregates, and the current and potential connection intensity of population, commerce, and logistics between cities in the Yellow River basin is calculated. Finally, the city network structure and status of cohesive subgroups in the Yellow River basin are measured by the social network analysis method, and the visual analysis of city network is carried out using ArcGIS software.

**3.2. Research Objects and Data Sources.** The Yellow River basin covers nine provinces or autonomous regions, namely, Qinghai, Sichuan, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan, and Shandong in China. After excluding nonurban areas, the research scope of this paper includes 78 prefecture-level cities in the nine provinces or autonomous regions. The research scope is shown in Figure 1 and the divisions of upper, middle, and lower reaches of Yellow River basin are shown in Figure 2.

The main data sources of this paper are shown in Table 1.

Table 2 shows that most of the key indicators of city development in the Yellow River basin are slightly higher than China’s average level but significantly lower than those of cities in the Yangtze River basin. In particular, the table indicates that from 2015 to 2019 the average annual GDP growth rate in the Yellow River basin cities was lower than China’s average level. In 2019, the average GDP of the 10 largest cities in the Yellow River basin was 557 billion yuan RMB and only Zhengzhou’s GDP exceeded 1 trillion yuan RMB. The average GDP of the 10 largest cities in the Yangtze River basin is 1704.8 billion yuan RMB, which is three times that of the Yellow River basin average value, and the GDP of eight cities exceeded 1 trillion yuan.

**3.3. Test Method for Intercity Connection Intensity: Gravity Model Modified by Two-Way “Time Distance”.** The idea of a gravity model originated from the universal gravitation formula of Newtonian mechanics in physics and was



FIGURE 1: Schematic diagram of research scope.

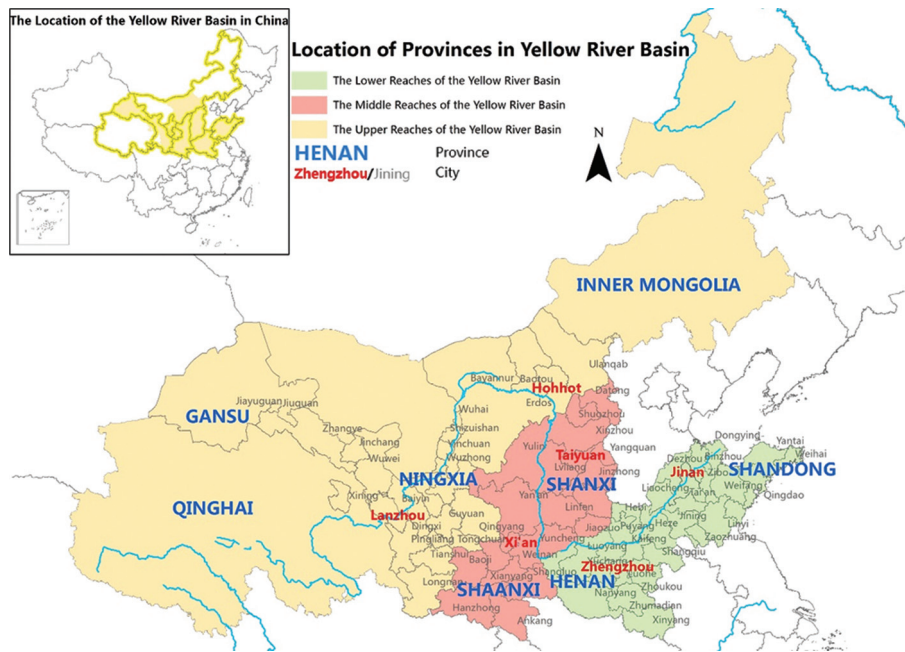


FIGURE 2: Location of provinces in Yellow River basin.

originally proposed by Tinbergen [36]. After improvement, it was introduced to calculate the intensity of two-way connections between cities [37, 38]. It can reflect both the radiation ability of central cities to surrounding cities and the acceptance degree of the radiation ability of central cities by surrounding cities. In this paper, the gravity model is modified by the two-way “time distance” between cities. According to the assumption of the gravity model, the connection intensity between two cities is directly proportional to the total “weight” of the two cities and inversely

proportional to the square of the shortest time distance between the two cities. The total “weight” of two cities can be expressed as the product of the GDP and the population of the two cities, while the time distance needs to be examined by comparing the shortest arrival time of various kinds of vehicles between cities.

Due to the limitation of hydrological conditions, the water transportation links between cities in the Yellow River basin are generally weak. In addition, they differ from the developed aviation in European and American countries.

TABLE 1: Main sources of data studied in this paper.

Data type	Data source	Value determination method and unit of data
(1) Highway commuting time	Consulting Google Maps, the point-to-point transportation time between city centers	Minutes
(2) Railway commuting time	Consulting the website of China Railway Customer Service Center	Minutes
(3) Economic, social, and demographic data of cities	The 2019 <i>China City Statistical Yearbook</i> and the statistical yearbooks of the provinces or autonomous regions	GDP: RMB 100 million yuan Population: 10 thousand people
(4) Transfer experience of mayors, deputy mayors, and secretaries of municipal CPC party committees among cities in the river basin*	Information of officials from the official websites of China's national and provincial governments, People's Daily Online, Xinhua Net, and local chronicles of cities	With transfer experience = 1 Without transfer experience = 0

Note: \*the data are used for later analysis. The secretary of a municipal CPC party committee is the highest leader of the Communist Party of China in the city, and his/her power is generally greater than that of the mayor.

TABLE 2: Comparison of the key indicators of city development in Yellow River basin, Yangtze River basin, and China's average level.

Key indicators	Research object cities in the Yellow River basin	Cities in the Yangtze River basin	Average level of prefecture-level cities in China
(1) 2015–2019 average annual GDP	81.7 billion RMB	96.7 billion RMB	77.01 billion RMB
(2) 2015–2019 average annual GDP growth rate (%)	6.74	7.97	6.93
(3) 2019 average urbanization rate (%)	61.86	71.05	60.60
(4) Average industrial output value	36.86 billion RMB	41.21 billion RMB	34.79 billion RMB
(5) 2019 proportion of agricultural output value (%)	9.77	12.36	10.98

First, the development level of high-speed railways and highways in China's Yellow River basin is much higher than that of aviation facilities, and the speed of high-speed railways reaches 200–300 km/h. Because the income level of residents in the basin is relatively low and the cost of air travel is relatively high, high-speed railways and highways are residents' first choices for traveling. Residents will choose air travel only when necessary and time is tight. Second, according to the data on urban passenger transport and freight transport from China's Ministry of Transport in 2019, the total passenger and freight transport by civil aviation in the Yellow River basin only accounted for 0.57% compared with highways (36.42%) and railways (59.91%). Third, airports are generally far away from urban areas. If the time for commuting to and from airports, security checks, and waiting for the flight is calculated, the actual commuting time is similar to or more than that of taking the high-speed rail. Fourth, the number of flights, flight times, and attendance rates is greatly affected by weather and route adjustment. Therefore, flight data are not calculated here on the premise they are not relevant to the research objectives.

Based on the above reasons, the time distance in this paper is mainly obtained by examination and comparing the shortest arrival time of cars, ordinary trains, bullet trains, and high-speed trains between cities in the Yellow River basin that can be obtained by consulting road and railway traffic information websites.

At the same time, this paper has considered the difference of two-way arrival time between cities. The two-way

time distance gravity model is shown in the following formula:

$$R_{ij} = \frac{\sqrt{G_i P_i} \sqrt{G_j P_j}}{T_{ij}^2}, \quad (1)$$

$$R_{ji} = \frac{\sqrt{G_j P_j} \sqrt{G_i P_i}}{T_{ji}^2}.$$

In formula (1),  $R_{ij}$  and  $R_{ji}$  are the connection intensity of city  $i$  to city  $j$  and city  $j$  to city  $i$ , respectively;  $P_i$  and  $P_j$  are the population sizes of city  $i$  and city  $j$ , respectively;  $G_i$  and  $G_j$  are the GDP values of city  $i$  and city  $j$ , respectively;  $T_{ij}$  and  $T_{ji}$  represent the shortest time distances from city  $i$  to city  $j$  and from city  $j$  to city  $i$ , respectively. After identification, there are 6006 links between every two of the 78 prefecture-level cities in the Yellow River basin, which can be divided into 3003 pairs:

$$F_1 = \max \frac{(R_{ij}, R_{ji})}{\sum_{i,j=1}^{n=78} R_{ij}}, \quad (2)$$

$$F_2 = \min \frac{(R_{ij}, R_{ji})}{\sum_{j,i=1}^{n=78} R_{ji}}.$$

In formula (2),  $n$  is the number of cities, and  $n = 1, 2, 3, \dots, 78$ ;  $\max(R_{ij}, R_{ji})$  and  $\min(R_{ij}, R_{ji})$  represent the larger

and smaller values in the connection intensity of each pair of cities, respectively;  $\sum_{i,j=1}^{n-1} R_{ij}$  represents the sum of the connection intensity values between every two cities;  $F_1$  and  $F_2$  can be used to compare the connection intensity among all cities and to analyze which cities have the strongest and weakest connections with each other in order to show the city network structure.

**3.4. Analysis Method for City Network Structure: Social Network Analysis Method.** Social network analysis (SNA) is applied to the study of network structure characteristics of social individuals (cities or people) and the importance of individuals in the network [39]. In recent years, SNA has been used to study connections among cities [40]. The main analysis indexes include three items: centrality, network density, and analysis of city cohesive subgroups.

**3.4.1. Centrality.** Centrality analysis mainly measures the leading force and attraction of cities in city networks [34], and the calculation indexes include degree centrality, closeness centrality, and betweenness centrality.

Degree centrality measures the number of connections between the city and other cities in the city network. It reflects the city's outward and inward connectivity (centrality), including "out-degree centrality"  $C'_{od}(i)$  and "in-degree centrality"  $C'_{id}(i)$ . The calculation method is shown in the following formula.  $X$  represents whether the city has a direct connection with other cities. If there is a direct connection, its value is 1; if there is no direct connection, its value is 0. The larger  $\sum_{j=1}^n X_{ij}$  and  $\sum_{i=1}^n X_{ji}$ , the more the connection of the city with other cities and the stronger its centrality:

$$\begin{aligned} C'_{od}(i) &= \frac{\sum_{j=1}^n X_{ij}}{(n-1)}, \\ C'_{id}(i) &= \frac{\sum_{i=1}^n X_{ji}}{(n-1)}. \end{aligned} \quad (3)$$

Closeness centrality measures the city's direct connection strength to other cities in the city network and is mainly based on the shortest time distance between cities in this paper. It reflects the level of the city's central function. The greater the closeness centrality, the stronger the direct connection between the city and other cities. Closeness centrality includes "out-closeness centrality"  $C'_{oc}$  and "in-closeness centrality"  $C'_{ic}$ , and the calculation method is as shown in the following formula.  $d(i,j)$  represents the shortest time distance between a city  $i$  and other cities  $j$ :

$$\begin{aligned} C'_{oc}(i) &= \frac{(n-1)}{\sum_{j=1}^n d(i,j)}, \\ C'_{ic}(i) &= \frac{(n-1)}{\sum_{j=1}^n d(j,i)}. \end{aligned} \quad (4)$$

Betweenness centrality  $C'_b(i)$  measures the importance of a city acting as an intermediary for other cities in the city

network, reflecting the intermediary function of the city. The calculation method is as shown in the following formula (5).  $g_{jk}$  is the total number of shortest paths by which city  $j$  connects city  $k$ ;  $g_{jk}(i)$  refers to the number of shortest paths between city  $j$  and city  $k$  passing through city  $i$

$$C'_b(i) = \frac{\sum_{j < k} g_{jk}(c_j) / g_{jk}}{(n-1)(n-2)}. \quad (5)$$

**3.4.2. Network Density.** Network density is an index reflecting the closeness and looseness of connection between cities in city networks.

The greater the network density, the closer the economic connection between cities. The calculation method is as shown in the following formula.  $D$  is the network density;  $m$  is the actual number of connections between cities in the city network, and  $n$  is the number of cities. " $n \times (n-1)$ " is the theoretical maximum number of connections between cities. The formula shows that the network density is expressed by comparing the actual connection between cities and the theoretical maximum network connection:

$$D = \frac{m}{[n \times (n-1)]}. \quad (6)$$

**3.4.3. Analysis of City Cohesive Subgroups.** In some parts of a city network, groups of cities are particularly, directly, and closely connected and are known as "internal cohesive subgroups." By analyzing the number, composition, and interaction of cohesive subgroups in a city network, the distribution structure of urban agglomeration networks can be investigated as a whole [41, 42]. In this paper, based on the data of city network connection intensity, UCINET6.5 software was applied to analyze the cohesive subgroups within the city networks in the Yellow River basin.

**3.5. Supplementary Research: Analysis on Intercity Transfer of Officials.** Under China's system, local officials have many social relations and resources and the right to make policies that can significantly impact regional development. At the provincial level, the fact that officials from developed provinces with outstanding achievements taking up posts in underdeveloped provinces is a common administrative means to strengthen interprovincial ties and cooperation and to support the development of underdeveloped regions. Officials holding office in different places develop important social relationships between cities.

The preliminary study in this paper showed that the number of officials transferred between cities in the Yellow River basin is small and concentrated locally, with excessively large regional differences, so this connection is not included in the overall study of city networks. However, in order to provide suggestions for strengthening connection between cities, this paper applied ArcGIS software for supplementary research on the transfer of mayoral-level officials among cities in the Yellow River basin and their office holding in different places. It also plotted a path

diagram for the transfer of officials among cities to reflect the political connection and potential cooperation among cities.

## 4. Results and Discussion

*4.1. Analysis of the Strength of Two-Way Economic Connection between Cities in the Yellow River Basin.* The intensity of the two-way economic connection between cities in the Yellow River basin is calculated by modifying gravity model formulas (1) and (2), and a diagram of two-way connection intensity is plotted using ArcGIS 10.2 software, as shown in Figures 3 and 4. Figure 3 represents the relatively high connection intensity  $F_1$  in each pair of economic connection, while Figure 4 represents the relatively low connection intensity  $F_2$  in each pair of economic connection. For clearer illustration, Figures 3 and 4 only show the connection with economic connection intensity values higher than RMB 100 million yuan \* 10,000 people/minutes<sup>2</sup>.

Table 3 shows 60 pairs of cities with the strongest connection intensity among the 3003 pairs of cities in the Yellow River basin and the reverse connection intensity between them. The calculation results show the following:

- (1) Generally, there is little difference in the intercity connection intensity in the two directions in the Yellow River basin, which indicates that most of the cities present a “reciprocal relationship” with similar mutual influence. A  $F_1/F_2$  greater than 1 is mainly seen in the four capital cities, namely, Xi’an, Zhengzhou, Jinan, and Taiyuan, indicating that the connection of these four cities with other cities is significantly stronger than the connection of other cities with these four cities, which reflects the characteristics of central cities.
- (2) The connection intensity values of cities in the Yellow River basin are quite different. The top five cities with the highest total connection intensity with other cities are Taiyuan (RMB 612.9 billion yuan \* people/minutes<sup>2</sup>), Zhengzhou (RMB 227 billion yuan \* 10,000 people/minutes<sup>2</sup>), Xi’an (RMB 196.9 billion yuan \* 10,000 people/minutes<sup>2</sup>), Jinan (RMB 165.6 billion yuan \* 10,000 people/minutes<sup>2</sup>), and Xianyang (RMB 137.3 billion yuan \* 10,000 people/minutes<sup>2</sup>). Longnan, Guyuan, Qingyang, Pingliang, and Baiyin are among other cities that have the lowest connection intensity values, which are all lower than RMB 1.5 billion yuan \* 10,000 people/minutes<sup>2</sup>. This shows that the cities in the Yellow River basin generally present a “center-edge” city structure and the development is unbalanced among cities.
- (3) There is a great regional difference in the connection intensity among cities in the Yellow River basin. In Figure 3, the intensity (relative thickness of lines) of city connection in the Yellow River basin shows significant trends of “south > north” and “lower reach > middle reach > upper reach.” The connection intensity and density among cities in Shandong Province in the lower reaches, as well as Shanxi

Province and Henan Province in the middle reaches, are higher than those in other regions. The connection among cities in the upper reaches of the river and cities on the north bank is significantly sparse, showing “voids,” and the development of cities there is relatively isolated.

### 4.2. An Analysis of City Network Structures in the Yellow River Basin

*4.2.1. An Analysis of the Centrality of Cities.* Various types of centrality of cities in the Yellow River basin were calculated according to formulas (3)–(5). The results suggest the following:

- (1) External connecting ability of cities: the distribution of out-degree centrality and out-closeness centrality of cities in the Yellow River basin is shown in Figures 5 and 6. The average value of cities’ out-degree centrality is only 20.8%, showing that the external connectivity of cities in the Yellow River basin is generally low, while the average value of cities’ out-closeness centrality is only 33.03%, showing that the city’s external connection ability is also relatively weak. The out-degree centrality values of Zhengzhou, Xi’an, and Jinan are only 18%, 9%, and 9%, respectively. Although the characteristics of the central cities are significant, their external connection and leading abilities are not strong. The weakness of central cities is a prominent problem in the backwardness of cities in the Yellow River basin.
- (2) Attraction of cities: The average value of cities’ in-degree centrality is only 31.32%, while the average value of cities’ in-closeness centrality is only 37.61%, which are relatively low levels. It can be seen from Figures 7 and 8 that the high-value areas of city in-degree centrality and in-closeness centrality are still located in the central cities in the lower and middle reaches such as Jinan and Zhengzhou, while those values are significantly lower in the middle and upper reaches.
- (3) Betweenness centrality: the average betweenness centrality of the city network basin is only 9.78%, which indicates that the intermediary hub function of cities in the region is generally weak. It can be seen from Figure 9 that some cities in the middle and upper reaches of the Yellow River have a relatively high betweenness centrality instead. For example, the intermediary function of Xi’an and Lanzhou is higher than that of Zhengzhou and Taiyuan in the middle reaches as well as Jinan in the lower reaches. This shows that there are few connection nodes, simple contact ways, and single channels among cities in these regions. Most cities can only rely on a small number of cities to establish contact with other cities, and the city network structure remains in the traditional “single center-marginal cities” state. Guyuan, Ulanqab, Weihai, Jiayuguan, and other

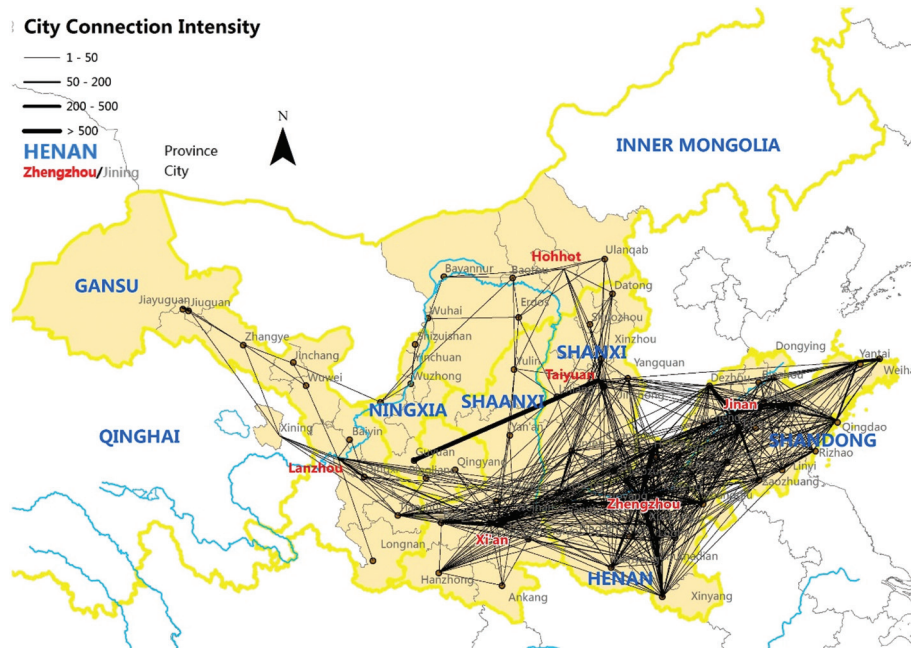


FIGURE 3: Distribution map of the maximum value of city connection intensity in the Yellow River basin.

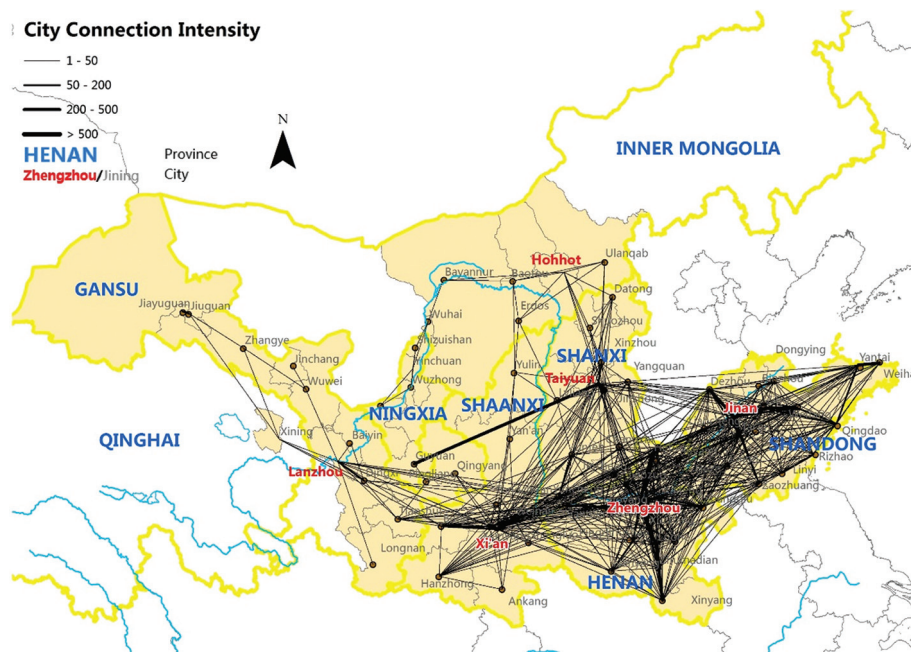


FIGURE 4: Distribution map of the minimum value of city connection intensity in the Yellow River basin.

cities have a betweenness centrality of 0, indicating that they are marginalized in the city network.

4.2.2. *Analysis on City Network Density.* According to formula (6), the average density of city networks in the Yellow River basin is only 13.2%. The levels of connection and integration in the development of urban agglomerations in the Yellow River basin are not high, and differentiation is significant. There are five areas with relatively high city

networks density values, namely, the Shandong Peninsula urban agglomeration in the lower reaches of the Yellow River, the Zhengzhou and Taiyuan urban agglomerations in the middle reaches of the Yellow River, and the Gansu-Shaanxi and Inner Mongolia-Ningxia urban areas in the upper reaches of the Yellow River. The total value of connection intensity among cities in these five regions accounts for 90% of that of the 3003 pairs of cities in the Yellow River basin. It can be seen from Table 4 that the city network density values of Shandong Peninsula, Zhengzhou, and



TABLE 3: Top 60 (ranked by  $F_1$ ) of intercity connection intensity in the Yellow River basin.

Direction of connection	$F_1$ (%)	$F_1$ ranking	Direction of connection	$F_2$ (%)	$F_2$ ranking	$F_1/F_2$
Taiyuan-Guyuan	7.462	1	Guyuan-Taiyuan	1.237	8	6.03
Xi'an-Xianyang	3.559	2	Xianyang-Xi'an	3.559	1	1.00
Taiyuan-Jinzhong	2.370	3	Jinzhong-Taiyuan	2.370	2	1.00
Jinan-Tai'an	1.931	4	Tai'an-Jinan	1.931	3	1.00
Xi'an-Weinan	1.733	5	Weinan-Xi'an	1.545	4	1.12
Zhengzhou-Xinxiang	1.472	6	Xinxiang-Zhengzhou	1.472	5	1.00
Zhengzhou-Xuchang	1.416	7	Xuchang-Zhengzhou	1.416	6	1.00
Zhengzhou-Kaifeng	1.309	8	Kaifeng-Zhengzhou	1.309	7	1.00
Jinan-Zibo	1.196	9	Zibo-Jinan	1.196	9	1.00
Hebi-Anyang	0.987	10	Anyang-Hebi	0.987	10	1.00
Jinan-Dezhou	0.957	11	Dezhou-Jinan	0.957	11	1.00
Xuchang-Luohe	0.939	12	Luohe-Xuchang	0.939	12	1.00
Yantai-Weihai	0.878	13	Weihai-Yantai	0.878	13	1.00
Zhumadian-Luohe	0.796	14	Luohe-Zhumadian	0.796	14	1.00
Zibo-Weifang	0.793	15	Weifang-Zibo	0.793	15	1.00
Xinxiang-Hebi	0.788	16	Hebi-Xinxiang	0.788	16	1.00
Zhengzhou-Jiaozuo	0.611	17	Jiaozuo-Zhengzhou	0.445	23	1.37
Zhumadian-Xinyang	0.608	18	Xinyang-Zhumadian	0.608	17	1.00
Nanyang-Pingdingshan	0.576	19	Pingdingshan-Nanyang	0.536	19	1.08
Zhengzhou-Luoyang	0.552	20	Luoyang-Zhengzhou	0.552	18	1.00
Zhoukou-Xuchang	0.539	21	Xuchang-Zhoukou	0.501	21	1.08
Xinxiang-Anyang	0.514	22	Anyang-Xinxiang	0.514	20	1.00
Rizhao-Linyi	0.447	23	Linyi-Rizhao	0.447	22	1.00
Xuchang-Zhumadian	0.429	24	Zhumadian-Xuchang	0.352	27	1.22
Qingdao-Weifang	0.420	25	Weifang-Qingdao	0.420	24	1.00
Jinan-Weifang	0.373	26	Weifang-Jinan	0.373	25	1.00
Zhengzhou-Luohe	0.372	27	Luohe-Zhengzhou	0.372	26	1.00
Luoyang-Sanmenxia	0.342	28	Sanmenxia-Luoyang	0.320	30	1.07
Zhengzhou-Hebi	0.341	29	Hebi-Zhengzhou	0.341	28	1.00
Zhengzhou-Anyang	0.336	30	Anyang-Zhengzhou	0.336	29	1.00
Tai'an-Dezhou	0.328	31	Dezhou-Tai'an	0.230	38	1.43
Jiuquan-Jiayuguan	0.312	32	Jiayuguan-Jiuquan	0.312	31	1.00
Shangqiu-Kaifeng	0.292	33	Kaifeng-Shangqiu	0.292	32	1.00
Zhengzhou-Zhumadian	0.287	34	Zhumadian-Zhengzhou	0.287	33	1.00
Zhengzhou-Pingdingshan	0.281	35	Pingdingshan-Zhengzhou	0.269	34	1.05
Zhoukou-Zhengzhou	0.265	36	Zhengzhou-Zhoukou	0.265	35	1.00
Zhengzhou-Shangqiu	0.256	37	Shangqiu-Zhengzhou	0.256	36	1.00
Zaozhuang-Tai'an	0.245	38	Tai'an-Zaozhuang	0.245	37	1.00
Xi'an-Baoji	0.241	39	Baoji-Xi'an	0.203	41	1.19
Xianyang-Weinan	0.218	40	Weinan-Xianyang	0.218	39	1.00
Taiyuan-Xinzhou	0.210	41	Xinzhou-Taiyuan	0.210	40	1.00
Lanzhou-Dingxi	0.202	42	Dingxi-Lanzhou	0.202	42	1.00
Baoji-Xianyang	0.202	43	Xianyang-Baoji	0.182	46	1.11
Linfen-Yuncheng	0.197	44	Yuncheng-Linfen	0.197	43	1.00
Binzhou-Dongying	0.194	45	Dongying-Binzhou	0.167	48	1.16
Qingdao-Yantai	0.188	46	Yantai-Qingdao	0.188	44	1.00
Jinan-Zaozhuang	0.184	47	Zaozhuang-Jinan	0.184	45	1.00
Xinyang-Luohe	0.175	48	Luohe-Xinyang	0.175	47	1.00
Xinxiang-Xuchang	0.165	49	Xuchang-Xinxiang	0.158	49	1.04
Zhengzhou-Nanyang	0.164	50	Nanyang-Zhengzhou	0.146	53	1.12
Shangqiu-Heze	0.157	51	Heze-Shangqiu	0.136	55	1.16
Qingdao-Zibo	0.152	52	Zibo-Qingdao	0.152	50	1.00
Xi'an-Sanmenxia	0.150	53	Sanmenxia-Xi'an	0.150	51	1.00
Weinan-Sanmenxia	0.149	54	Sanmenxia-Weinan	0.142	54	1.05
Xuchang-Xinyang	0.147	55	Xinyang-Xuchang	0.147	52	1.00
Kaifeng-Xuchang	0.147	56	Xuchang-Kaifeng	0.026	163	5.72
Hohhot-Ulanqab	0.135	57	Ulanqab-Hohhot	0.135	56	1.00
Zhengzhou-Xinyang	0.133	58	Xinyang-Zhengzhou	0.129	58	1.03
Weinan-Yuncheng	0.133	59	Yuncheng-Weinan	0.133	57	1.00
Dezhou-Zibo	0.127	60	Zibo-Dezhou	0.127	59	1.00

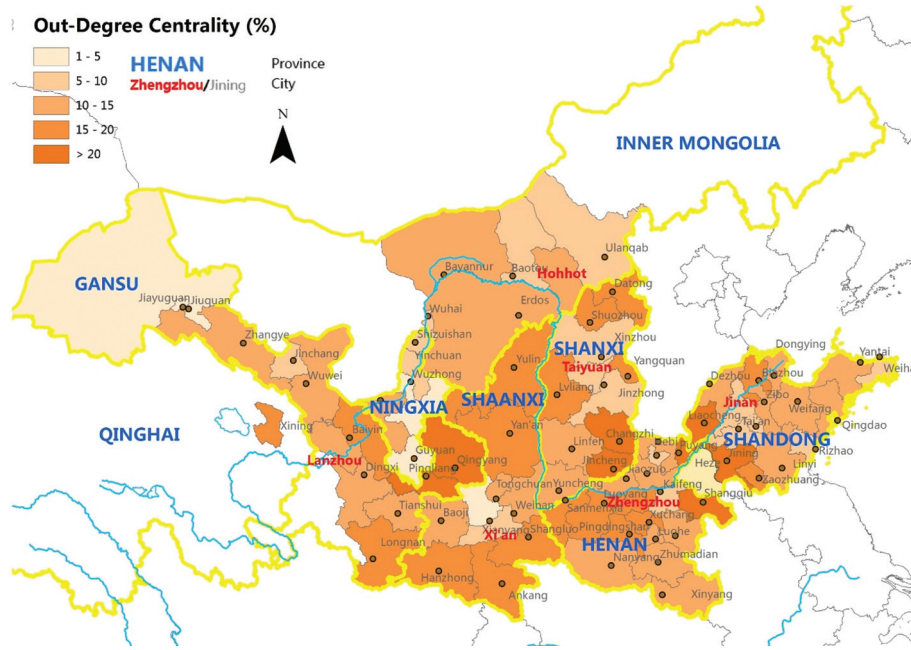


FIGURE 5: Distribution map of out-degree centrality of cities in the Yellow River basin.

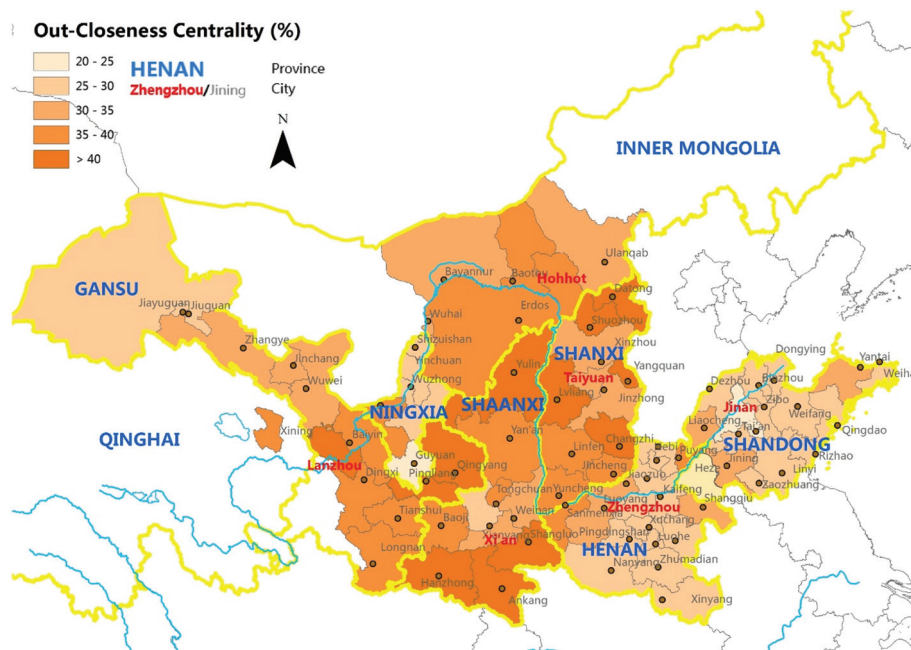


FIGURE 6: Distribution map of out-closeness centrality of cities in the Yellow River basin.

Taiyuan urban agglomerations are all over 50%, while that of Gansu-Shaanxi urban agglomerations is only 26.5%.

It can be seen from Figure 10 that the network density of urban agglomerations in the Yellow River basin is also dense in the east and sparse in the west and dense in the south and sparse in the north and that the development gap between the cities on the north and south banks is larger than that between the east and west regions. Although the central cities

are not strong, the overall “central-edge” city structure in the basin is clear. Zhengzhou City, Henan Province, has the highest network density in the middle reaches of the Yellow River, followed by Xi’an City, Lanzhou City, and Jinan City, which are arranged on an east-west axis. Taiyuan City, Linfen City, and Changzhi City of Shanxi Province in the middle reaches of the Yellow River are the main channels connecting the south and north banks, but their ability of

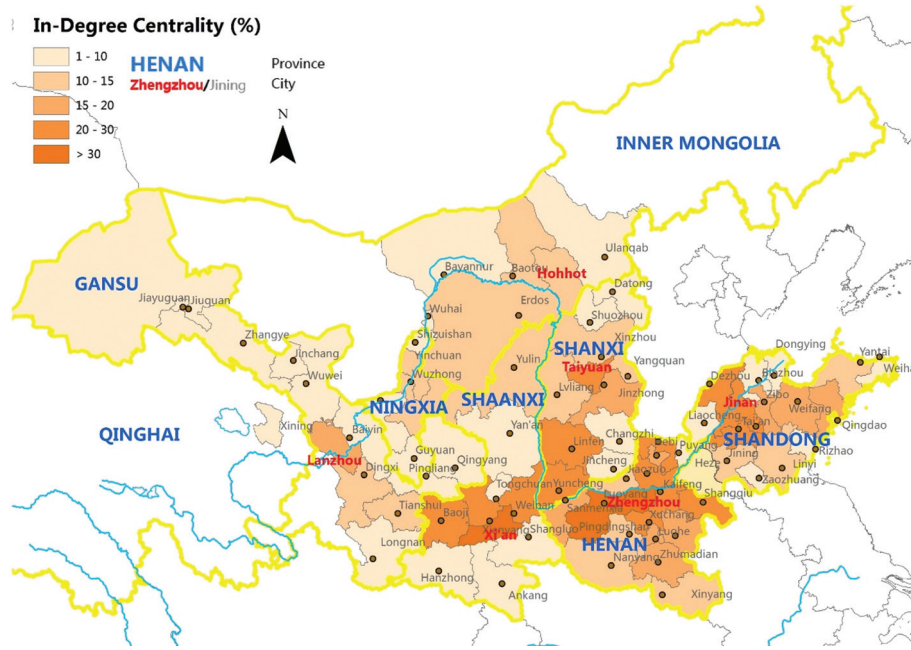


FIGURE 7: Distribution map of in-degree centrality of cities in the Yellow River basin.

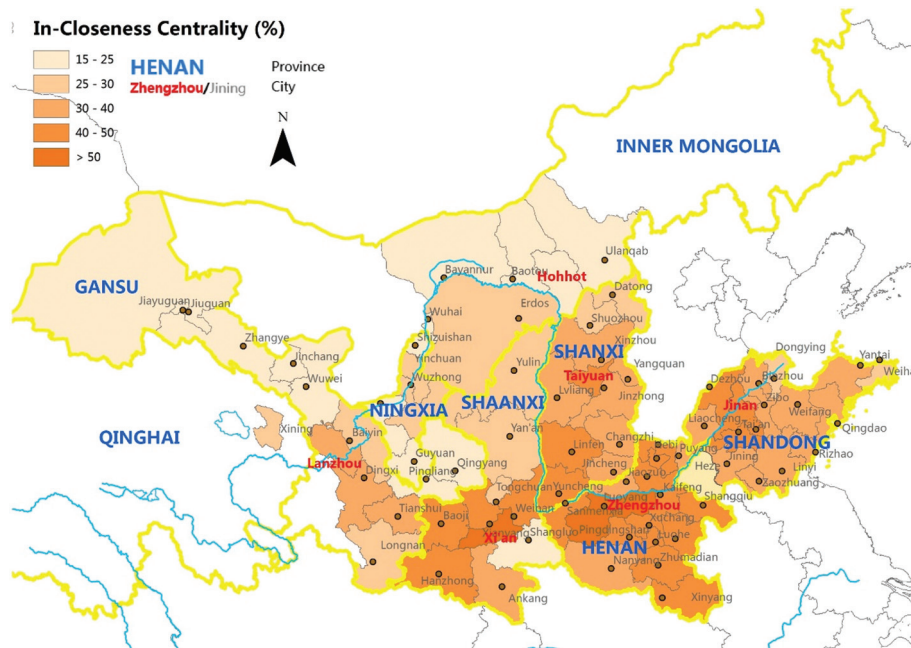


FIGURE 8: Distribution map of in-closeness centrality of cities in the Yellow River basin.

connecting the upper and lower reaches in the east-west direction is significantly weaker than that of Zhengzhou. The central cities present an inverted T-shaped distribution.

The spatial distribution of cities with high network density is shown in Figure 11, showing an inverted T-shaped structure.

4.3. *Analysis on Cohesive Subgroups of Cities.* The gravity matrix data between cities in the Yellow River basin was put into UCINET6.5 software to analyze the subgroups of the urban network in the Yellow River basin. The result shows

that cities in the Yellow River basin can be divided into eight third-class subgroups and four second-class subgroups (Table 5). However, most of the subgroups are links among cities within a province, while the links among subgroups are weak (Figures 12 and 13).

The spatial distribution of cohesive subgroups of city networks in the Yellow River basin is shown in Figure 13.

4.4. *Analysis of the Transfer Network of Officials among Cities in the Yellow River Basin.* For this study, ArcGIS software was applied to draw the path diagram of officials' transfer

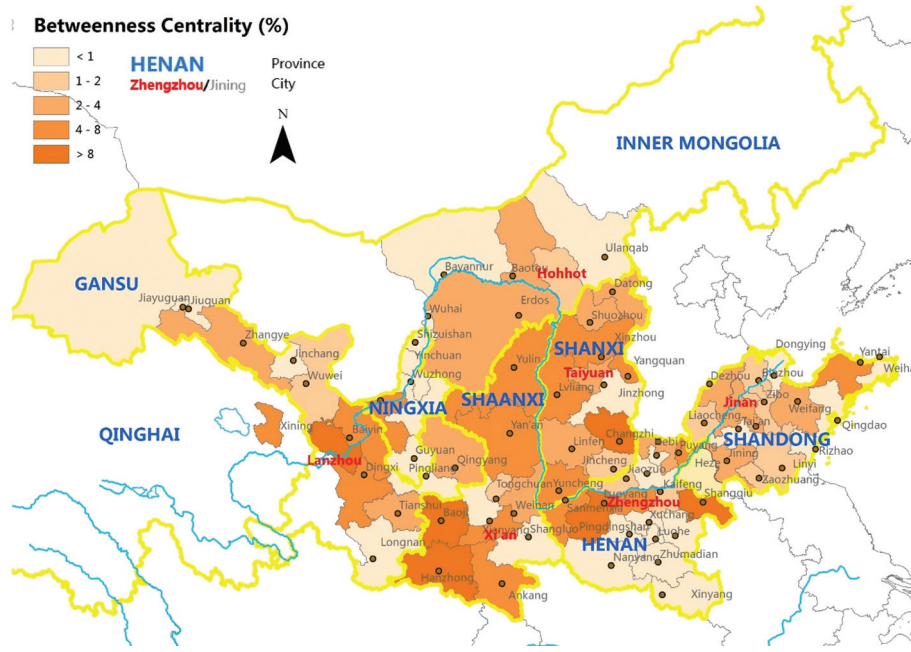


FIGURE 9: Distribution map of betweenness centrality of cities in the Yellow River basin.

TABLE 4: Average value of city network indexes and city development indexes of the Yellow River basin and its subregions.

Subregion	Network density (%)	Out-degree centrality (%)	In-degree centrality (%)	Out-closeness centrality (%)	In-closeness centrality (%)	Betweenness centrality (%)	2019 urbanization rate (%)	2019 average GDP (100 million yuan RMB)
Overall Yellow River basin	13.2	20.8	31.3	36.7	37.6	9.78	56.78	2262.27
Zhengzhou urban agglomeration	55.5	33.9	47.3	64.7	70.7	9.79	50.73	3904.85
Shandong Peninsula urban agglomeration	52.5	22.2	50.6	63.0	68.1	9.08	60.99	4102.32
Taiyuan urban agglomeration	50.0	33.0	55.0	60.8	65.0	12.63	57.79	2755.56
Gansu-Shaanxi urban agglomeration	26.5	19.8	79.1	27.9	39.1	29.28	47.35	1301.30
Inner Mongolia-Ningxia urban agglomeration	43.6	29.0	18.0	54.9	56.2	11.75	44.10	1990.63

among cities in the Yellow River basin, which reflects the political connection and cooperation possibilities among cities. Figures 14 and 15 show the following:

- (1) The transfer of officials among cities in the Yellow River basin is mainly within the province and less between provinces. The “network” has not been formed yet, and the regional differences are too great. Therefore, this study does not include this connection in the city network analysis.
- (2) “Horizontal transfer” (with the same level of position before and after transfer) is common among cities in the basin, with two-way transfer, but it mainly occurs among noncentral cities. Officials transferred across

provinces are usually transferred horizontally. “Promoted transfer” (the official position was promoted after the transfer) is characterized by a one-way transfer from noncentral cities to central cities.

- (3) There are far more officials transferred into than transferred out of the regional central cities, and the five cities with the greatest numbers of officials transferred into are Zhengzhou (47), Jinan (40), [[parms resize(1),pos(50,50),size(200,200),bgcol(156)]] central cities provide little support to noncentral cities by transferring excellent officials out, but they have absorbed a great number.

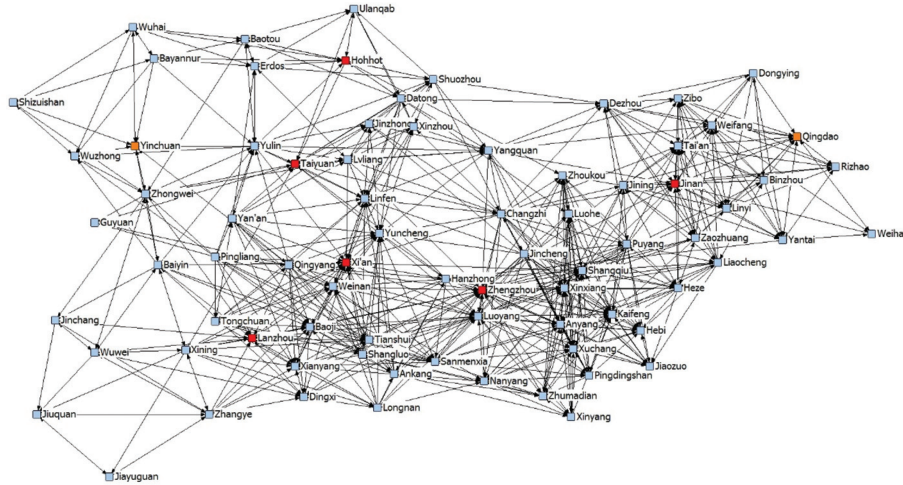


FIGURE 10: Structure of city networks in the Yellow River basin.

TABLE 5: Analysis table of cohesive subgroups of city networks in the Yellow River basin.

Region	The 2 <sup>nd</sup> class subgroup		The 3 <sup>rd</sup> class subgroup	
	Location	Provinces involved	Area name	Cities involved
The Yellow River basin	North bank of middle reaches	Inner Mongolia, Ningxia, and Shanxi	Taiyuan urban agglomeration Inner Mongolia, Ningxia area	Taiyuan, Datong, Jinzhong, Ulanqab, Yizhou, Shuozhou, Lvliang, Linfen, and Yangquan Hohhot, Yinchuan, Wuhai, Baotou, Bayannur, Wuzhong, Ordos, Zhongwei, Yulin, and Shizuishan
	Middle and upper reaches	Shaanxi, Gansu, Qinghai, and part of Henan, and Shanxi	Xi'an urban agglomeration	Xi'an, Shangluo, Qingyang, Yuncheng, Ankang, Longnan, Baoji, Tianshui, Hanzhong, Pingliang, Guyuan, Yan'an, Sanmenxia, Weinan, and Xianyang
	Lower reaches	Shandong	Lanzhou-Xining urban agglomeration	Lanzhou, Xining, Jiayuguan, Jinchang, Jiuquan, Zhangye, Dingxi, Wuwei, and Baiyin
			Southern Shandong urban agglomeration	Dongying, Zaozhuang, Liaocheng, Heze, and Jining
	South bank of middle reaches	Henan and part of Shanxi	Shandong urban agglomeration (e.g., southern Shandong) Henan-Shanxi Adjoining area Zhengzhou urban agglomeration	Jinan, Qingdao, Zibo, Yantai, Rizhao, Linyi, Dezhou, Weihai, Tai'an, and Weifang Changzhi, Jincheng, Jiaozuo, and Puyang Xinxiang, Nanyang, Xinyang, Zhoukou, Zhumadian, Shangqiu, Xuchang, Luohe, Zhengzhou, Kaifeng, Luoyang, Pingdingshan, Anyang, and Hebi

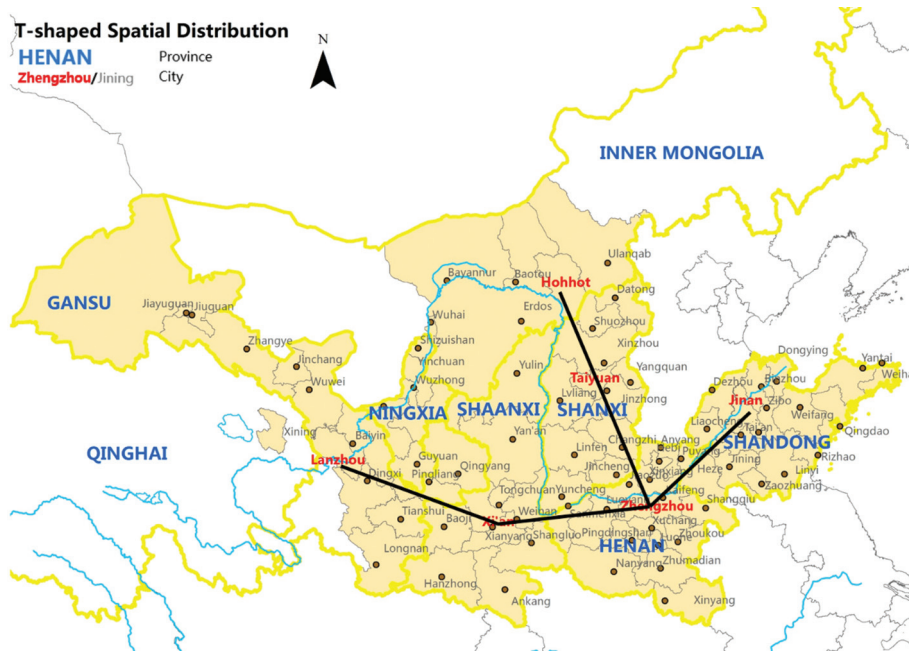


FIGURE 11: Spatial distribution of cities with high network density.

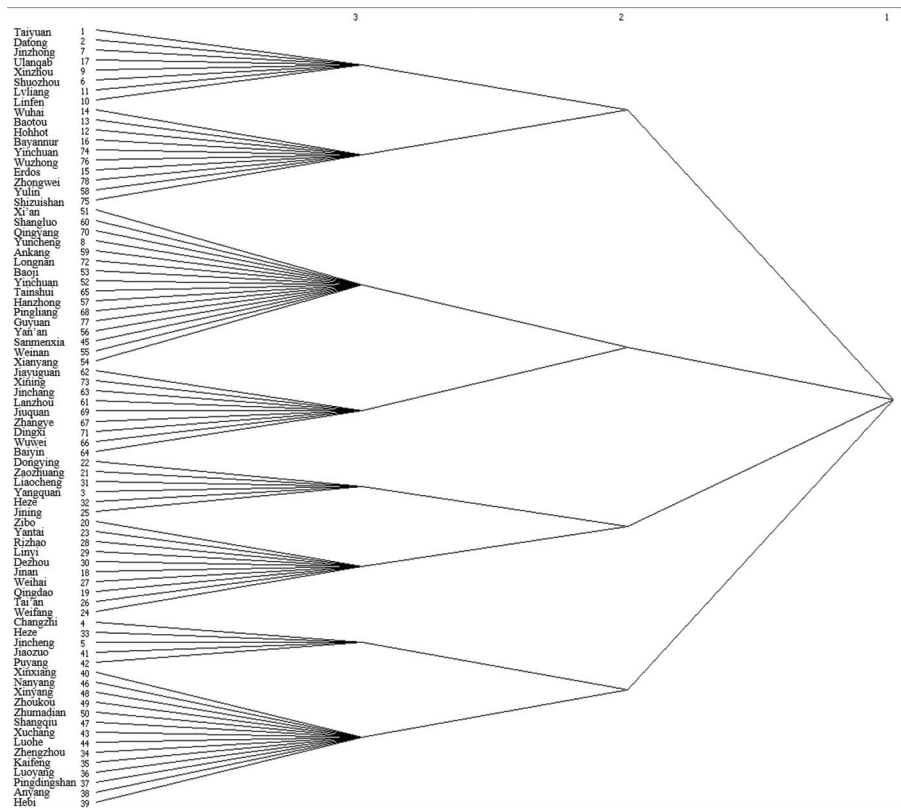


FIGURE 12: Cohesive subgroups of city networks in the Yellow River basin.

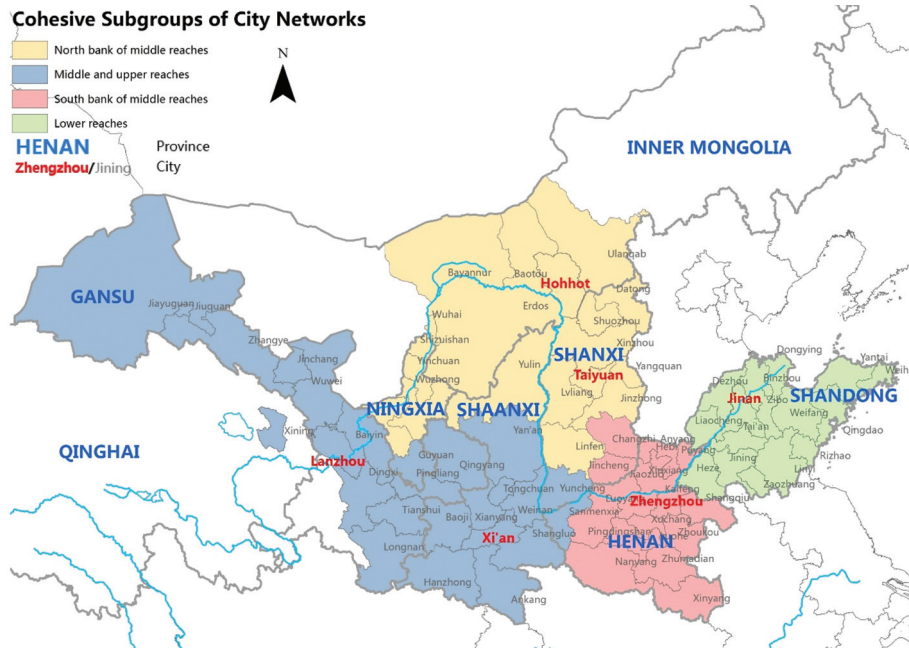


FIGURE 13: Spatial distribution of cohesive subgroups of city networks in the Yellow River basin.

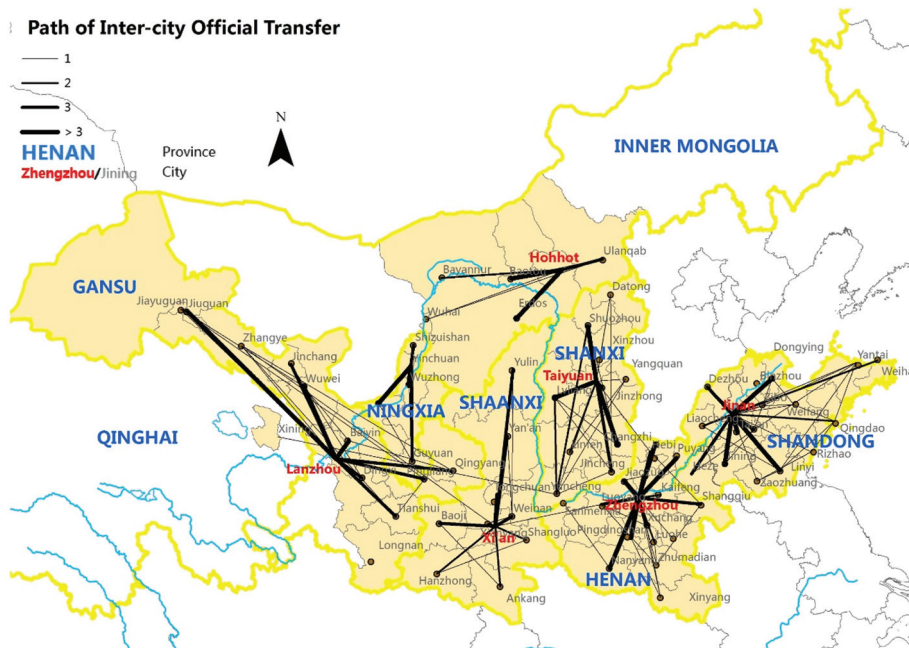


FIGURE 14: Path diagram of intercity official transfer in the Yellow River basin.

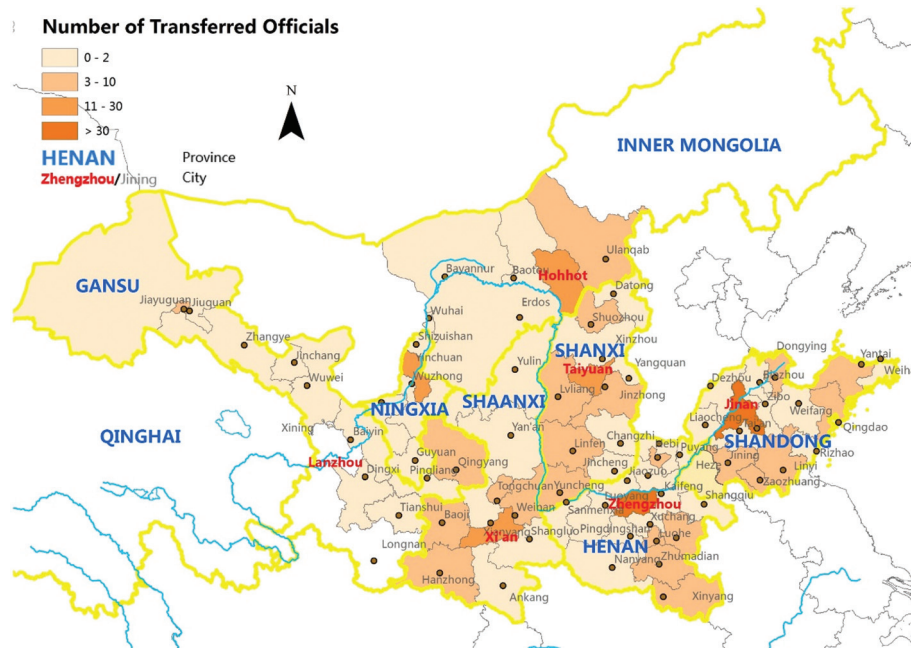


FIGURE 15: Distribution of the numbers of officials transferred into cities in the Yellow River basin.

## 5. Conclusion

In this paper, the entire city network in the Yellow River basin was examined empirically for the first time. The city social network analysis method was improved based on the empirical research on characteristics of the city network in a typical large river basin. It can provide guidance for the planning and policies of urban system construction in the Yellow River basin, as well as reference for other similar regions in the world with lagging development and fragile ecology. The main conclusions are as follows:

(1) The values of intercity connection intensity, out-degree centrality, in-degree centrality, and betweenness centrality of cities in the Yellow River basin are generally low and highly differentiated. The values of intercity connection intensity and city network density are characterized by “south bank > north bank” and “lower reach areas > middle reach areas,” and the development gap between cities on the north and south banks is larger than that between cities on the upper reaches and the lower reaches. The connection intensity of cities in the Yellow River’s upper reaches in western China and cities in Inner Mongolia on the north bank is generally lower by more than 80% than the average value of the overall river basin city network. These cities appear as voids in the city network, are marginalized in the network, and have the problem of isolated development.

(2) The intensity of the two-way connection among cities in the basin is generally similar, but the connection intensity of a few central cities to other cities is significantly higher than that of other cities to themselves. Although the central cities are not

strong, the overall “central-edge” city structure in the Yellow River basin is clear. The central cities present an inverted T-shaped spatial distribution as a whole. However, the central cities’ weak connecting and leading abilities are a prominent problem in the development of cities in the Yellow River basin.

(3) There is a positive correlation between the city network development level and regional economic development level. Cities in the lower and middle reaches with higher economic development levels have roughly similar two-way intercity connection intensities, while cities in economically underdeveloped areas in the middle and upper reaches have greater numerical differences in interconnection and more unbalanced urban development; the status of the central city and the marginal city is more disparity.

(4) The cohesive subgroups of city networks in the Yellow River basin have a high internal connection intensity, while the connection among cohesive subgroups is weak. Although city networks in different river basin sections exist, the connection among urban agglomerations in the river basin is still weak, and the problem of fragmented development in the river basin is prominent. It is difficult for industries and people to be transferred along the river basin in cascades and to form linkages among the upper, middle, and lower reaches.

Based on the empirical analysis, this paper puts forward some suggestions for development policies and planning of the Yellow River basin:

(1) The Yellow River basin lacks central cities with strong economic and leading capabilities, so it is



difficult to copy the strategy of taking developed cities in Shanghai, Jiangsu, and Zhejiang as economic cores and leading inland development as in the Yangtze River basin of China. The strategy of “beaded chain” can be adopted; i.e., the four secondary cohesive subgroups of city networks (Figures 12 and 13) in the Yellow River basin can be taken as the zoning scheme of the recent basin planning, giving priority to strengthening the internal connections within the secondary subgroups, such as vigorously developing intercity railways. Connection of various infrastructures among adjacent subgroups should be gradually strengthened, and the interconnected development among the upper, middle, and lower reaches should be promoted.

- (2) In the near future, the layout of central city construction can be based on the current inverted T-shaped structure, and the economy and industrial policies, resources, and infrastructure construction in the basin should be given key support. First, an east-west core city chain of Lanzhou–Xi’an–Zhengzhou–Jinan connecting the upper and lower reaches should be built. Then, taking Zhengzhou as the intersection of the T-shape, the core cities in the Yellow River basin that connect the upper and lower reaches could be cultivated, and Xi’an City, Lanzhou City, and Jinan City developed as subcenters. Changzhi City and Taiyuan City of Shanxi Province in the central region could be used as the main connecting channels between the north and south banks of the Yellow River basin, to gradually strengthen the construction of traffic channels between the north and south banks of the Yellow River in other regions. This would resolve the imbalanced development between the north and south banks.
- (3) Regional node cities should be cultivated in regions with relatively few connection nodes in the middle and upper reaches of the Yellow River. Yinchuan City of Ningxia Autonomous Region, Yulin City and Yan’an City of Shaanxi Province, Dingxi City and Zhangye of Gansu Province, and Xining City of Qinghai Province could be priority choices.
- (4) Because the ecological environment in the Yellow River basin is fragile and the economic power of the cities is weak, resources for the construction of intercity connection facilities must be jointly provided by relevant cities to improve the efficiency of resource utilization and to avoid duplicated construction. Therefore, it is important to learn from the systems of the Tennessee Valley Authority (TVA) in the USA, the International Commission for the Protection of the Rhine (ICPR) in Europe, and the Yangtze River Economic Belt Committee in China, improve the functions and jurisdiction of the Yellow River Committee, and strengthen the pace of

coordination in policies, funds, and infrastructure construction in the region.

- (5) The current situation whereby most of the outstanding city officials in the Yellow River basin transfer into the central cities is unfavorable for the development of backward areas. The outstanding officials from lower reach provinces should be transferred to upper reach provinces and from central cities to other cities to enhance connection and cooperation among cities and to support the development of underdeveloped cities.

Problems for the future on this topic include the following:

- (1) Due to the lack of data on the point-to-point traffic volume of various modes of transportation through roads and railways among cities in the basin, this study has not considered the traffic volume of various modes of transportation in calculating time distance, in order to simplify the analysis for the time being. This is also a common problem in all studies at present [20–22, 24, 28]. In the future, the method for this study could be further improved after the establishment of big data resources for intercity transportation in the Yellow River basin.
- (2) The idea of gravity modelling in this paper shows that the connection intensity between two cities is directly proportional to the economy and population size of the two cities and inversely proportional to commuting time, while the ecological connection has not been considered. With the strengthening of ecological protection of the Yellow River in China, we can introduce an ecological aspect into the study after more detailed data on ecological functional areas in the basin are established.
- (3) In subgroups of city networks in the Yellow River basin, the impact of the cross-city transfer of city officials on the formation of city network may be further studied in the future to reveal the effects of official ties between cities.

## Data Availability

The economic, social, and demographic data supporting this study are from the 2019 China City Statistical Yearbook and the statistical yearbooks of the provinces or autonomous regions in China. The processed data are available in the Supplementary Files (1). The original data are available from the corresponding author upon request. Intercity highway and railway commuting time data used to support the findings of this study are available from Google Maps and the China Railway Customer Service Center. The original data are available in the Supplementary Files (2). The transfer experience of mayors, deputy mayors, and secretaries of municipal CPC party committees’ data supporting this study comes from the official websites of China’s national and provincial governments, People’s Daily Online,

Xinhua Net, and local chronicles of cities. The data are the phased achievement of the National Natural Science Foundation of China. The processed data are available in the Supplementary Files (3). Requests for original data, six months after publication of this article, will be considered by the corresponding author. The GIS data supporting spatial analysis and mapping in this study are from the National Geomatics Center of China <http://www.ngcc.cn/ngcc/html/1/391/392/16114.html>. The processed 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 study was supported by the National Natural Science Foundation of China (Grant no. 71974220). The authors also gratefully acknowledge many useful insights by Professor Yan Song of the University of North Carolina at Chapel Hill.

## Supplementary Materials

Supplemental file (1): these include the economic, social, and demographic data. Supplemental file (2): these contain matrix of the shortest time distance between cities. Supplemental file (3): these are the transfer experience of city officials in the Yellow River basin. Supplemental file (4): these carry gravity model results. Supplemental file (5): these are the results of cities' degree centrality, closeness centrality, and betweenness centrality in the city networks of the Yellow River basin. (*Supplementary Materials*)

## References

- [1] S. Hennemann, B. Derudder, and P. J. Taylor, "Cutting the Gordian knot of visualizing dense spatial networks: the case of the world city network, 2013," *Environment & Planning A*, vol. 47, no. 6, pp. 1332–1340, 2015.
- [2] B. Derudder and P. J. Taylor, "Change in the world city network, 2000–2012," *The Professional Geographer*, vol. 68, no. 4, pp. 624–637, 2016.
- [3] J. V. Beaverstock, M. A. Doel, P. J. Hubbard, and P. J. Taylor, "Attending to the world: competition, cooperation and connectivity in the world city network," *Global Networks*, vol. 2, no. 2, pp. 111–132, 2002.
- [4] X. Ma and G. Li, "Study on world city network theory within global space of flow," *Economic Geography*, vol. 31, no. 10, pp. 1630–1637, 2011.
- [5] J. Jiao, J. Wang, F. Jin, and H. Wang, "Impact of high-speed rail on inter-city network based on the passenger train network in China, 2003–2013," *Acta Geographica Sinica*, vol. 71, no. 2, pp. 256–280, 2016.
- [6] M. Li and L. Y. Wang, "Discussing how to extract city characteristic with the examples of ancient capital cities in the Yellow River Basin," *China Ancient City*, no. 5, pp. 25–30, 2017.
- [7] H. Liu, L. Wang, Q. Li et al., "Spatial and temporal patterns of city connection networks in the Yellow River Basin based on Tencent's big data of population migration," *Economic Geography*, vol. 40, no. 4, pp. 28–37, 2020.
- [8] The Political Bureau of the CPC Central Committee, "Outline of ecological protection and high quality development plan of the Yellow River Basin," *China Water Conservancy*, vol. 70, no. 6, p. 6, 2020.
- [9] J. Xi, "Speech at the symposium on ecological protection and high quality development in the Yellow River River Basin," *China Water Resources*, no. 20, pp. 1–3, 2019.
- [10] S. Ardekani and R. Herman, "Urban network-wide traffic variables and their relations," *Transportation Science*, vol. 21, no. 1, pp. 1–16, 1987.
- [11] D. A. Smith and M. F. Timberlake, "World city networks and hierarchies, 1977–1997: an empirical analysis of global air travel links," *American Behavioral Scientist*, vol. 44, no. 10, pp. 1656–1678, 2001.
- [12] P. J. Taylor, G. Catalano, and D. R. F. Walker, "Exploratory analysis of the world city network," *Urban Studies*, vol. 39, no. 413, pp. 266–279, 2002.
- [13] J. H. Choi, G. A. Barnett, and B.-S. Chon, "Comparing world city networks: a network analysis of internet backbone and air transport intercity linkages," *Global Networks*, vol. 6, no. 1, pp. 81–99, 2010.
- [14] C. A. Jarmon and J. M. Vanderleeuw, "City leaders and economic development networks: the all-channel star network," *Journal of Political Science*, vol. 19, pp. 1–32, 2011.
- [15] Z. Neal, "Validity in world city network measurements," *Tijdschrift Voor Economische en Sociale Geografie*, vol. 105, no. 4, pp. 427–443, 2014.
- [16] X. Jin, G. Hu, H. Ding, S. Ye, Y. Lu, and J. Lin, "Evolution of spatial structure patterns of city networks in the Yangtze River economic belt from the perspective of corporate pledge linkage," *Growth and Change*, vol. 51, no. 8, pp. 833–851, 2020.
- [17] C. Varol, O. Y. Ercoskun, and N. Gurer, "Local participatory mechanisms and collective actions for sustainable urban development in Turkey," *Habitat International*, vol. 35, no. 1, pp. 9–16, 2011.
- [18] E. C. Jennifer and A. R. Angel, "Mississippi gulf coast advances sustainability plan," *Biocycle*, vol. 53, no. 1, pp. 52–55, 2012.
- [19] D. Wang and R. Zhuang, "The preliminary probe into the quantitative analysis of regional economic links," *Scientia Geographica Sinica*, vol. 16, no. 1, pp. 51–57, 1996.
- [20] G. Li, L. Wang, and K. Yang, "The measurement and analysis of economic relationship between Shenzhen and Zhujiang delta," *Economic Geography*, vol. 21, no. 1, pp. 33–37, 2001.
- [21] H. Wu, Y. Shen, and Q. Hu, "Spatial correlation characteristics of economic growth in Beijing-Tianjin-Hebei Region and its explanation: based on spatial autocorrelation and network analysis," *Jiangxi Social Science*, vol. 36, no. 3, pp. 75–80, 2016.
- [22] S. Wang, Y. Song, Y. Zhang, and J. Li, "Spatial organization and driving mechanism of connected network of urban agglomeration in the middle reaches of the Yangtze River," *Economic Geography*, vol. 40, no. 6, pp. 87–97, 2020.
- [23] C. Li, "City network structure of Shandong province based on information flow," *Journal of Arid Land Resources and Environment*, vol. 29, no. 12, pp. 51–56, 2015.
- [24] C. Wang and M. Wang, "Evolution of associated network characteristics between cities in Shandong Province: comparing central place and flow space theories," *Geographical Research*, vol. 36, no. 11, pp. 2197–2212, 2017.

- [25] Y. Cheng and L. Wang, "Reinterpretation of the theory of central place in a context of space of flow: a study on provincial city network in Shandong Province," *Economic Geography*, vol. 37, no. 12, pp. 25–33, 2017.
- [26] S. Zhao, H.-Y. Zhou, and L.-J. Zhu, "Applying the Voronoi map to the research on the city network of Henan," *Journal of Institute of Surveying and Mapping*, vol. 20, no. 3, pp. 206–209, 2003.
- [27] C. Miao and H. Wang, "On the direction and intensity of urban economic contacts in Henan Province," *Geographical Research*, vol. 25, no. 2, pp. 222–232, 2006.
- [28] D. Meng and Y. Lu, "Impact of high-speed railway on accessibility and economic linkage of cities along the railway in Henan Province, China," *Scientia Geographica Sinica*, vol. 31, no. 5, pp. 537–543, 2011.
- [29] Q. Song and J. Liu, "Research on structure of urban network among Zhongyuan urban agglomeration based on the industrial perspective," *Henan Science*, vol. 34, no. 3, pp. 424–430, 2016.
- [30] T. Geng and J. Li, "The research of Shaanxi Province city network connection based on Baidu index," *Henan Science*, vol. 34, no. 7, pp. 1166–1172, 2016.
- [31] H. He and B. Lyu, "The measurement on economic connection in Changzhutan urban agglomeration," *Economic Geography*, vol. 34, no. 7, pp. 67–74, 2014.
- [32] P. Chong, S. Chen, and B. Wang, "Analyzing city network's structural resilience under disruption scenarios: a case study of passenger transport network in the middle reaches of Yangtze River," *Economic Geography*, vol. 39, no. 8, pp. 68–76, 2019.
- [33] F. Peng, "Economic spatial connection and spatial structure of Guangdong-Hong Kong- Macao Greater Bay and the surrounding area cities: an empirical analysis based on improved gravity model and social network analysis," *Economic Geography*, vol. 37, no. 12, pp. 57–64, 2017.
- [34] H. Meng, X. Huang, and J. Yang, "Network structure and development concept in Huaihai economic zone," *Economic Geography*, vol. 39, no. 12, pp. 1–10, 2019.
- [35] J. Guo, S. Wang, B. Li, and J. Guo, "The spatial effect of Harbin-Dalian high-speed rail to the northeast city tourism economic link," *Scientia Geographica Sinica*, vol. 36, no. 4, pp. 521–529, 2016.
- [36] J. Tinbergen, *Shaping the World Economy: Suggestions for an International Economic Policy*, Twentieth Century Fund, New York, NY, USA, 1962.
- [37] K. Zhao and Y. Wu, "Study on evolution characteristics and synergy between high-speed railway network and economic network in China," *East China Economic Management*, vol. 34, no. 2, pp. 77–85, 2020.
- [38] J. Wang, J. Jiao, and F. Jin, "Spatial effects of high-speed rails on interurban economic linkages in China," *Acta Geographica Sinica*, vol. 69, no. 12, pp. 1833–1846, 2014.
- [39] A. W. Wolfe, "Social network analysis: methods and applications," *Contemporary Sociology*, vol. 91, no. 435, pp. 219–220, 1995.
- [40] S. Wang, Y. Song, H. Wen et al., "Network structure analysis of urban agglomeration in the Yangtze River economic belt under the perspective of bidirectional economic connection: based on time distance and social network analysis method," *Economic Geography*, vol. 39, no. 2, pp. 73–81, 2019.
- [41] D. Nepelski and G. De Prato, "Corporate control, location and complexity of ICT R&D: a network analysis at the city level," *Urban Studies*, vol. 52, no. 4, pp. 721–737, 2015.
- [42] X. Li, "Research on the Yangtze River delta urban agglomeration network structure based on social network analysis," *Urban Studies*, vol. 18, no. 12, pp. 80–85, 2011.

## Research Article

# The Spread of Information in Virtual Communities

Zhen Zhang <sup>1</sup>, Jin Du,<sup>1</sup> Qingchun Meng,<sup>2</sup> Xiaoxia Rong <sup>3</sup>, and Xiaodan Fan <sup>1</sup>

<sup>1</sup>Department of Statistics, The Chinese University of Hong Kong, Hong Kong SAR, China

<sup>2</sup>School of Management, Shandong University, Jinan, China

<sup>3</sup>School of Mathematics, Shandong University, Jinan, China

Correspondence should be addressed to Xiaoxia Rong; [rongxiaoxia@sdu.edu.cn](mailto:rongxiaoxia@sdu.edu.cn) and Xiaodan Fan; [xfan@cuhk.edu.hk](mailto:xfan@cuhk.edu.hk)

Received 7 October 2020; Revised 6 November 2020; Accepted 10 November 2020; Published 26 November 2020

Academic Editor: Liang Wang

Copyright © 2020 Zhen 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.

With the growth of online commerce, companies have created virtual communities (VCs) where users can create posts and reply to posts about the company's products. VCs can be represented as networks, with users as nodes and relationships between users as edges. Information propagates through edges. In VC studies, it is important to know how the number of topics concerning the product grows over time and what network features make a user more influential than others in the information-spreading process. The existing literature has not provided a quantitative method with which to determine key points during the topic emergence process. Also, few researchers have considered the link between multilayer physical features and the nodes' spreading influence. In this paper, we present two new ideas to enrich network theory as applied to VCs: a novel application of an adjusted coefficient of determination to topic growth and an adjustment to the Jaccard coefficient to measure the connection between two users. A two-layer network model was first used to study the spread of topics through a VC. A random forest method was then applied to rank various factors that might determine an individual user's importance in topic spreading through a VC. Our research provides insightful ways for enterprises to mine information from VCs.

## 1. Introduction

Virtual communities (VCs) provide an interactive experience that, if positive, may instil customer loyalty [1]. They enable consumers to learn the functions of products and follow up conveniently with buying online, as well as provide a channel for receiving customer feedback, which plays an important role in product innovation [2, 3]. Mining information provided by consumers in VCs enables companies to adjust the next generation of their products to improve customer satisfaction [4].

Complex network theory has been a major tool in the study of the physical structure and dynamic processes of social, biological, and technological networks [5]. In analyses of information spreading, users are represented as nodes [6]; these nodes may reside in multiple possible states, depending on whether they have learned information and whether they can transmit it to a neighbour [7–9]. Among real-world VCs, social networks such as Weibo, WeChat,

Twitter, and Facebook have different physical structures, leading to different patterns of topic transmission.

Consumer VCs are different from traditional social networks in that they centre around products. They provide a real-time look into customers' experiences with a product from the date of release. Users in VCs may post their feedback for other users to view, which, in the best-case scenario, may encourage loyalty among existing consumers while encouraging new consumers to buy in. Thus, understanding consumer VCs can provide insight into public opinion trends, helping the company maintain existing markets or develop new markets [10].

In consumer VCs, information is transmitted via posts, which are about particular topics [11]. Networks change over time; thus, time dimensions, i.e., temporal networks, have been incorporated into network analyses [12]. For consumer VCs, the identification of key time points during topic emergence can provide an angle for addressing the consulting service appropriately. Additionally, there are

multiple ways for users in VCs to interact with one another, of which “replying” and “searching for interest” are common ways to get information. Mining the influential users and specifying the important network features are also critical for improving the efficiency of multilayer network information transformation [13].

Thus, this paper tries to answer the following questions:

- (i) Following the introduction of a new product, how does the number of topics concerning the product grow over time?
- (ii) How do topics spread through a VC? What features characterise the influence of individual users during the information-spreading process?

Here, we focused specifically on the Huawei P10/P10 Plus for our case study. Two new concepts are introduced. First, the coefficient of determination was applied to the growth of topics after a new product was introduced to yield the node sequential emergence coefficient of determination (NSECD), which was used to identify the moment when most of the growth had finished. Second, the Jaccard coefficient was adjusted to gain a new measure of similarity between two users in a VC. Subsequently, a two-layered network model, representing two ways of spreading information, was introduced to study the spread of a topic through a VC and identify the most influential users. The random forest method was then applied to rank the importance of various factors affecting a user’s impact on spreading topics through a VC. Finally, some suggestions are provided for future research.

The remainder of this paper is organised as follows. Section 2 provides a literature review on existing methods. Section 3 describes the Huawei P10/P10 Plus dataset and its pre-processing. Section 4 introduces the NSECD statistic for studying the growth of new topics after the introduction of a new product. Section 5 proposes the adjusted Jaccard coefficient and the two-layered network model. Later in the section, simulations carried out to identify key users in the network are described, with the random forest method used to find features important to users’ information-spreading performance. Section 6 recaps suggestions for enterprise management of VCs and proposes directions for further research.

## 2. Literature Review

Online social networks exert a major influence on life today [14]. Sange et al. found that online social networks provide a platform for spreading both objective facts and fake news [15]. Park et al. noted the rapid growth of mobile devices such as smartphones and examined rapid information propagation in mobile social networks [16]. Up to now, the spreading of information has been investigated intensively in interdisciplinary fields [17].

Research into information propagation in VCs has two main foci: influence factor analysis and propagation path analysis. Influence factor analysis attempts to identify which factors make a node influential in a network; such factors may include gender, age, beliefs, etc. On the other hand, propagation path analysis studies the way in which

information is transmitted through the network by, for instance, assigning weights to edges and setting transmission probabilities based on these weights.

We first provide examples of influence factor analysis. Li et al. developed a multinomial naive Bayes classifier, categorised microblog posts based on content, and found that the information type has a significant influence on propagation patterns in terms of scale and topological features [18]. Zeng and Zhu proposed an emotional model for information propagation based on the emotional states of network users [19]. Hsu developed an integrated conceptual model and explored the effects of brand-evangelism-related behavioural decisions of enterprises on VC members [20]. Wan et al. used a least squares support vector machine to study consumer electronics supply chains [21]. However, all of the above studies were mainly analyses of the influencing factors but did not consider the differences among VCs.

With regard to propagation form, Huo and Cheng established a modified ignorant-wiseman-spreader-stifler model to analyse the spread of rumours through a network [22]. Xu et al. proposed a new iterative algorithm called SpectralRank, which assumes that a node’s propagation capability is proportional to the number of neighbouring nodes after adding a ground node to the network [23]. Shao et al. introduced the NL centrality algorithm to identify influential nodes in a network; the algorithm considers both the semilocal structure of a node and its topological position. [24]. Wang et al. proposed a method based on an integral k-shell to identify the influential nodes in a command-and-control network [25]. Escalante and Odehnal proposed a deterministic SIRS-type model for rumour propagation and applied it in simulations with two types of rumours: an original rumour, followed by a second counteracting rumour based on a complex network [26]. Li et al. presented the Potential Concentration Label method to help locate multiple sources of contagion under a susceptible-infected-recovered model [27]. Zhang et al. introduced a susceptible-infected-true-removed model of rumour spreading to account for members in a network who know or can discern the truth [28]. Xiong et al. introduced the location concept to a local social network model [29] and further extended it to the recommendation system via information spreading in a local-based social network [30, 31]. In their recent research, Xiong et al. combined the location and temporal effects of a social network, proposing constructive advice on dynamic management [32]. Zhang et al. studied networks that can be subdivided into smaller groups called communities and proposed a node-ranking algorithm called AI Rank using two factors: attractive power (which measures the number of followers a node has compared to its neighbours) and initiating power (which accounts for the communities that a node’s neighbours belong to) [33].

Although these studies considered node importance ranking together with the topology of the network, they generally treated networks as single-layered. This does not account for the fact that there may be more than one way for information to propagate through a network. Hence, a multilayer network model may be more appropriate. We propose a two-layered model in this paper.

### 3. Data Preprocessing and Description

The Pollen Club is the official VC for Huawei’s products, including smartphones, laptops, and other electronic devices. Each user is assigned a unique identifier with which they can express their opinions about products freely, look through other users’ posts, and reply to posts [34].

For our initial dataset, we selected 2000 web pages about the Huawei P10/P10 Plus, containing 2,392,035 posts about the product. After removing duplicate and spam posts, we retained 57,560 original posts and 826,328 reply posts by 129,362 users as our dataset. The data were acquired directly from club.huawei.com to avoid interview effects [35].

Next, core topics were extracted from the posts. In a previous study [36], 100 topics about phones were selected initially (see Table 1). After sorting out those with higher frequency, the remaining topics were grouped into three categories: system, software, and hardware, according to their features, as shown in Table 2.

The remainder of this paper focuses on 61 of these topics. These topics did not all appear on the first day of our dataset, but emerged over the course of the study as users bought and used the product and formulated questions about the product. In the next section, we will discuss the emergence of topics.

### 4. Dynamic Analysis of Topic Emergence

According to the classic product lifecycle, a new product goes through three stages: emergence, growth, and maturation. Emergence refers to the period before the product’s launch [37]. Growth refers to a period of high consumer activity after the product launch, when consumers who had been eagerly awaiting the launch are ready to buy the product. Maturity refers to a period of low consumer activity afterwards, when consumers may continue to buy the product but do so at a slower pace because enthusiastic consumers have already done so.

For VCs, the transition point from growth to maturation is of interest to the company, because it signals the point after which fewer resources should be needed to monitor and respond to VC posts concerning the product. The purpose of this section is to introduce a statistic that will identify such a transition point.

For this purpose, we used the growth in the cumulative number of topics from the 61 topics up to a given date to measure the VC’s interest in the product. We could, for instance, select the date at which the cumulative number of topics reached 90% or 95% of 61. However, any such choice would involve some uncertainty regarding which threshold to use. Instead, we propose a transition point that avoids such an arbitrary threshold choice.

Based on the adjusted coefficient of determination in statistics [38], we define the NSECD as follows:

$$r_t = 1 - \frac{T}{T-t+1} \left( 1 - \frac{n_t}{n+1} \right) (1 \leq t \leq T), \quad (1)$$

where  $r_t$  is the NSECD value on the  $t^{\text{th}}$  day,  $n_t$  is the cumulative number of new topics on the  $t^{\text{th}}$  day,  $n$  is the total number of topics, and  $T$  is the total number of days.

As  $t$  increases,  $T/(T-t+1)$  increases while  $1 - (n_t/n+1)$  decreases. We expect that in a typical dataset, strong growth in new topics early in the study period will lead  $r_t$  to increase initially while faster growth in the factor  $T/(T-t+1)$  together with a saturation of new topics will lead to  $r_t$  decreasing later on. Our key moment will be when this function reaches its maximum; i.e.,

$$t^* = \arg \max_t (r_t). \quad (2)$$

Figure 1 shows the NSECD on each day in our dataset. In this dataset, it was calculated that  $t^* = 33$ . Accordingly,  $r_{t^*} = 0.8457$ .

To summarise the lifecycle of topic growth in our case study, the period from the start of the study to the product launch date (Day 9) was considered the “emergence” stage. Twelve new topics were added during this stage, reflecting early interest in the Huawei P10/P10 Plus. The period from the product launch date to our key moment (Day 33) was the “growth” stage. The period from our key moment to the end of the study was considered the “maturity” stage.

In Figure 2, 12 new topics were added during the “emergence” stage, all on Day 9, the final day of the stage. This reflected early interest in the Huawei P10/P10 Plus.

The “growth” stage continued the trend of new topics from the last day of the emergence stage. New topics appeared rapidly early in this stage, as users shared their opinions towards the product from different aspects, but the rate at which new topics appeared slowed towards the end. By the end of this stage, 58 topics had emerged, 95.08% of the total number of topics considered.

The “maturity” stage witnessed even slower growth in new topics compared to the earlier stages, as most topics had already appeared earlier.

### 5. Network Modelling

In this section, we first introduce a two-layer network model and then perform information-propagation simulations to determine which users are the most effective at spreading information in a VC.

*5.1. Structure of the Network.* Here, we establish a two-layered network model for VCs. Each user is represented by a node that occurs on both layers. The two layers correspond to two ways in which a user in a VC may interact with another user: by replying to their posts or by searching for their posts.

The first layer is the “flow of information by replies (FIR)” network (denoted as  $\text{FIR}(V_1, E_1, W_1)$ ). Given two users  $U_i$  and  $U_j$ , let  $w_{ij}^1$  be the number of times  $U_j$  replied to a post by  $U_i$  within the dataset. If  $w_{ij}^1 > 0$ , then an arrow  $a_{ij}^1$  from  $i$  to  $j$  is drawn. The FIR network consists of  $V_1$ , the set of all users,  $W_1$ , the set of all  $w_{ij}^1$  weights, and  $E_1$ , the set of all  $a_{ij}^1$  arrows.

The second layer is called the “flow of information by interest (FII)” network (denoted as  $\text{FII}(V_2, E_2, W_2)$ ). It is inspired by the idea that two users are likely to search for each other’s posts, only when they share the same interests.

TABLE 1: Topics and frequency.

Serial number	Topics	Frequency	Serial number	Topics	Frequency	Serial number	Topics	Frequency	Serial number	Topics	Frequency
1	System	6569	26	Game	2810	51	Withdraw money	22	76	Motherboard	183
2	Life	1408	27	Light	692	52	Defrayment	1218	77	Touch	97
3	Taking pictures	3162	28	Flash back	695	53	King Glory	1471	78	Translation	42
4	Hibernation	204	29	Record	216	54	Beta	425	79	Clipboard	3
5	Power consumption	2905	30	Noise	123	55	Jingdong	601	80	Power saving mode	318
6	Data	1780	31	Font	499	56	Vmall	468	81	Wall paper	469
7	Wifi	1992	32	Loudspeaker	130	57	Theme	850	82	3D paper	15
8	Upgrade	4133	33	Program	1012	58	Antifingerprint oleophobic coating	622	83	Firmware bag	54
9	Charging	2459	34	Code	1007	59	Ring	472	84	Desktop	1066
10	Power off	934	35	Cloud service	191	60	Battery	2094	85	Radio	75
11	Unlock	1795	36	Lock screen	1968	61	Navigation key	236	86	NFC	622
12	WeChat	5569	37	Memory	1946	62	System B172	228	87	Developer options	129
13	Video	2574	38	Pattern	2684	63	System B213	231	88	Automatic rotation	56
14	Clarity	83	39	Backups	421	64	System B167	269	89	Vague	975
15	Telephone	2806	40	Resolution	398	65	Home key	495	90	Root	432
16	Net	1946	41	Touch off	2	66	Android8.0	95	91	Log version	6
17	Pixel	349	42	Pick-up hand	1719	67	Calendar	117	92	Kugou	54
18	Fingerprint	3241	43	Update	6422	68	Voice assistant	454	93	Compatibility	59
19	Payment	219	44	Camera	1778	69	Type	217	94	GPS	330
20	Block	172	45	Lightness	781	70	Infrared remote control	46	95	Weibo	709
21	Bank card	139	46	Player	137	71	Gesture operation	17	96	QQ	2253
22	Transportation card	345	47	Music	841	72	Application treasure	34	97	Position	803
23	Consumption	790	48	Screen capture	618	73	Split screen	216	98	Internet speed	443
24	Message	1075	49	Color	774	74	Heat	2700	99	4G	2277
25	Black screen	591	50	Location	600	75	Chip	279	100	Dominant frequency	33

TABLE 2: Classification of topics.

Type	Topics
System	System, upgrade, lock screen, unlock, black screen, font, resolution, update, lightness, color, screen capture, location, telephone, net, mode, vague, data, Wifi, power off, beta, theme, ring, voice assistant, heat, wall paper, desktop, NFC, root, 4G, internet speed, GPS, position
Software	WeChat, fingerprint, transposition card, consumption, message, game, flash back, program, code, backups, music, defrayment, video, King Glory, Jingdong, Vmall, Weibo, QQ
Hardware	Life, taking pictures, power consumption, charging, memory, pick-up hand, camera, light, antifingerprint oleophobic coating, battery, home key

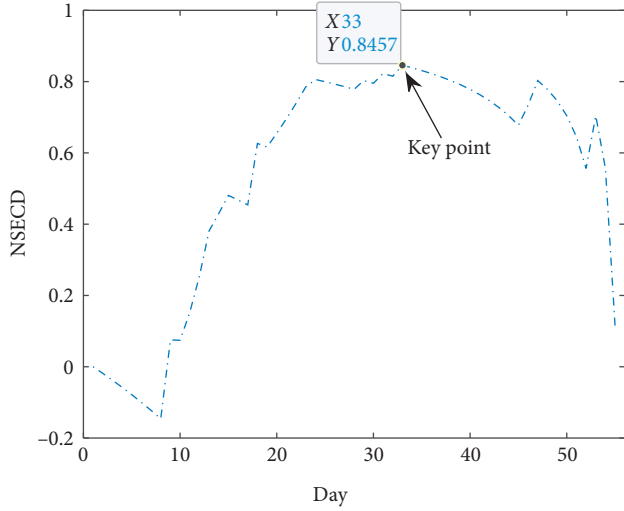


FIGURE 1: Graph of the node sequential emergence coefficient of determination.

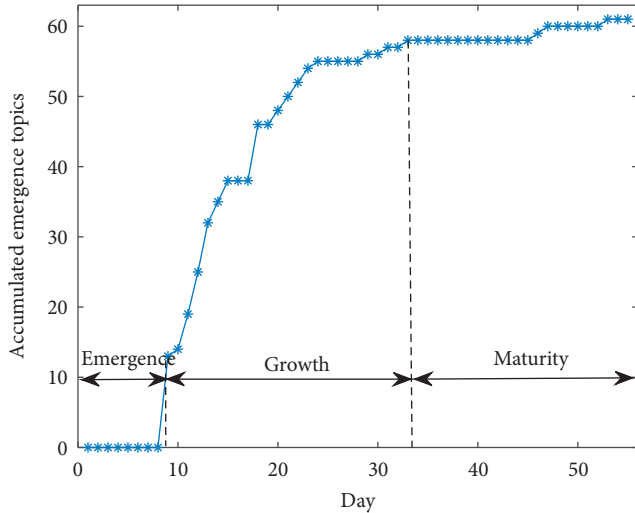


FIGURE 2: The topic lifecycle of Huawei P10/P10 Plus.

Given two users  $U_i$  and  $U_j$ , we can construct a measure  $w_{ij}^2$  representing the commonality of interests and draw an edge  $a_{ij}^2$  between them when  $w_{ij}^2$  is above some preset threshold  $\varepsilon$ . The FII network consists of  $V_2$ , the set of all users,  $W_2$ , the set of all  $w_{ij}^2$  weights, and  $E_2$ , the set of all  $a_{ij}^1$  arrows. It remains to define each  $w_{ij}^2$  weight.

Let  $N$  be the total number of topics and let the topics have a fixed order. Then, let  $\vec{F}^i = (F_1^i, \dots, F_k^i, \dots, F_N^i)^T$  be the vector comprising the numbers of the main  $U_i$  posts concerning each of the topics. Because some users may focus on only a portion of the topics, most entries in  $\vec{F}^i$  are zero. As such, a Jaccard coefficient is taken into consideration. Let  $R_i = \{k: F_k^i > 0\}$ , i.e., the set of all topics for which  $U_i$  has at least one main post. By definition, the Jaccard coefficient of  $\vec{F}^i$  and  $\vec{F}^j$  can be calculated as follows:

$$\rho_{ij} = \frac{\#(R_i \cap R_j)}{\#(R_i \cup R_j)}, \quad (3)$$

where  $\#(R_i \cap R_j)$  and  $\#(R_i \cup R_j)$  denote the number of elements in the intersection and union of topics mentioned by users  $U_i$  and  $U_j$ , respectively.

The disadvantage of the Jaccard coefficient is that it does not distinguish a pair of casual users who post about only one topic, which happens to be the same, and a pair of enthusiastic users who post about many shared topics. For example, consider the following two situations:

Situation 1: users  $U_i$  and  $U_j$  post only on “system” and “updates.”

Situation 2: users  $U_i$  and  $U_j$  post only on “system.”

Based on equation (3), the Jaccard coefficient assigns a weight of 1 to both situations. However, this may not be appropriate, because the users in Situation 1 may be more active in the VC and, hence, more likely to exchange information by searching. Thus, we propose the following adjusted Jaccard coefficient:

$$\rho_{ij}^a = \frac{\#(R_i \cap R_j)}{\#(R_i \cup R_j)} \frac{\#(R_i \cup R_j)}{N} = \frac{(\#(R_i \cap R_j))^2}{\#(R_i \cup R_j) \times N}. \quad (4)$$

In our case study,  $N = 61$ . Based on equation (3),  $\rho_{ij}^a = 0.0328$  in Situation 1 and  $\rho_{ij}^a = 0.0164$  in Situation 2. Like the Jaccard coefficient, the adjusted Jaccard coefficient is always between 0 and 1, and is 0 when both users share no common topics. However, unlike the classical Jaccard coefficient, it is 1 only when both users share all topics.

In Figure 3, we plot the distribution of values of  $\rho_{ij}^a$  over all pairs of users in our case study. Figures 3(a) and 3(b) show that the overwhelming majority of pairs of users were associated with a small value of  $\rho_{ij}^a$ . A small  $\rho_{ij}^a$  means that there is little possibility for information spreading. In our



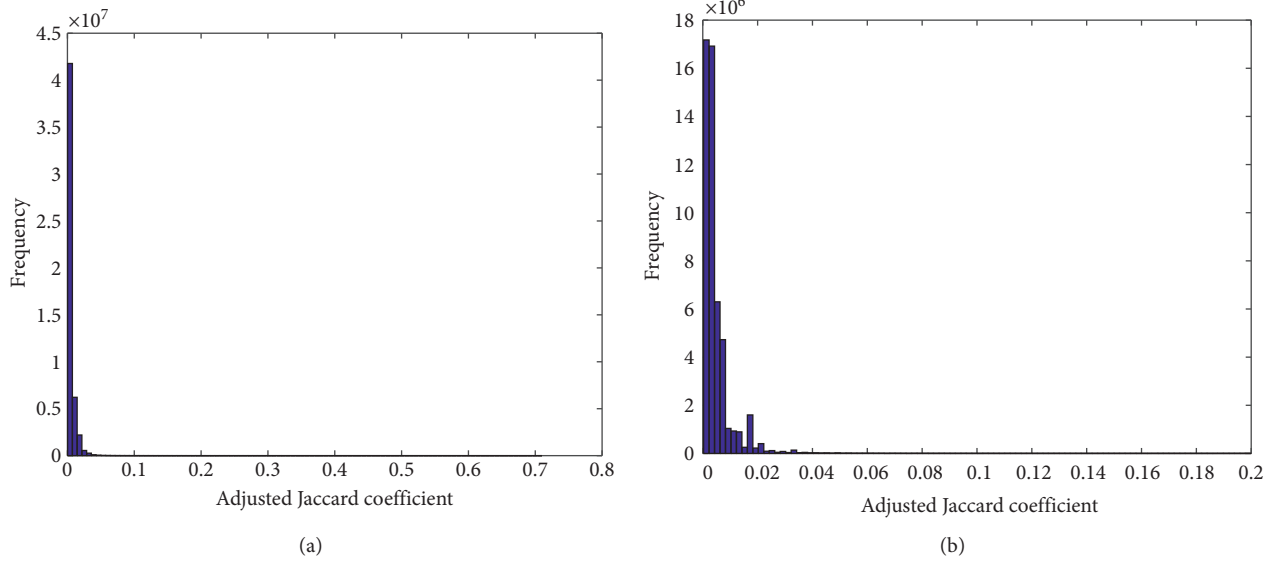


FIGURE 3: Distribution of  $\rho_{ij}^a$ : (a) distribution of  $\rho_{ij}^a$ ; (b) distribution of  $\rho_{ij}^a$  ( $\rho_{ij}^a < 0.2$ ).

analysis, we used a threshold value of  $\varepsilon = 0.2$ . That is, an edge  $a_{ij}^2$  was drawn between User  $U_i$  and User  $U_j$  only when  $\rho_{ij}^a > 0.2$ . This made the FII network a sparse network. Notably, the difference in the range longitudinal axis between Figures 3(a) and 3(b) is caused by the difference in the bar intervals.

Gephi software was used to draw illustrations of both layers for our case study. The results are shown in Figure 4.

To illustrate how our two-layer network model works, consider the following simple example with seven users, as shown in Figure 5.

Here,  $a-g$  represent the seven users. Users may send or receive information from the FIR or FII networks simultaneously. For example, User  $a$  can receive information from User  $b$  or  $c$  in the FIR network, or from User  $b$  or  $d$  in the FII network. User  $a$  can send information to User  $b$  or  $d$  in the FII network. Information can spread through both channels simultaneously. The larger the weight of an edge, the more likely information is to flow through it at any given step in both layers.

**5.2. Information Propagation and User Importance.** Next, the two-layered network model will be expanded with information transmission mechanisms, to simulate the flow of information through the VC. Suppose we are interested in the propagation of a specific piece of information through the network. The information-propagation model will be based on a simple two-state framework; that is, at any given time, a node is in one of two possible states:

- (1) A susceptible state, representing the states of users who do not receive the information
- (2) An infected state, indicating the states of users who receive the information

Time is treated as discrete. At each time point  $t$ , infected nodes have a probability of transmitting the information to

their susceptible out-neighbours. The spread of the information is thus a random process. The mathematical model of this process is described below.

For convenience, susceptible and infected states are denoted by  $\alpha$  and  $\beta$ , respectively. We denote the FIR layer and the FII layer as  $l = 1$  and  $l = 2$ , respectively. The key notations are listed as follows:

- (i)  $H_i^{(t)}$  is the state of User  $i$  at time  $t$
- (ii)  $r_i^{1,t}$  and  $r_i^{2,t}$  are the indicator variables for User  $i$  being in states  $\alpha$  or  $\beta$ , respectively, at time  $t$
- (iii)  $N^1(i)$  is the set of out-neighbours of User  $i$  in the FIR network
- (iv)  $N^2(i)$  is the set of neighbours of User  $i$  in the FII network
- (v)  $N(i) = N^1(i) \cup N^2(i)$  is the group of target users that User  $i$  may spread information to
- (vi)  $I(\cdot)$  is the indicator function. It takes the value of 1 when the assertion inside is true and 0 when the assertion inside is false

The key mechanisms for this model are as follows:

- (1) For User  $i$  ( $i = 1, 2, 3, \dots, n$ ), each day  $t - 1$  except for the last step and each layer  $l$   $p_i^{l,t-1}$  is a draw from  $U(0, 1)$ . If User  $i$  is infected at  $t - 1$ , the value of this draw will decide which of the  $i$ 's out-neighbours on the layer  $l$  become infected at time  $t$ . More probable infections are always prioritised over less probable infections. A low value of  $p_i^{l,t-1}$  means that the node will be easily infected, whereas a high value means that it will be difficult to infect.
- (2) For each User  $i$ , each layer  $l$ , and each in-neighbor  $j$  of  $i$ ,  $p_{ji}^l$  will measure the infectiousness of the  $j \rightarrow i$  channel. Large values indicate that  $j$  easily infects  $i$ , whereas small values indicate that  $j$  has difficulty

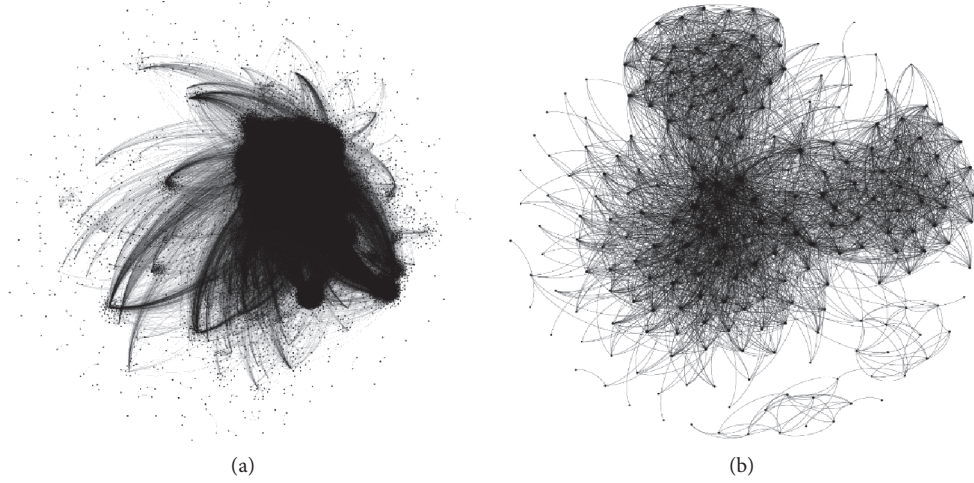


FIGURE 4: The two-layered network model applied to Huawei P10/P10 Plus: (a) FIR network; (b) FII network.

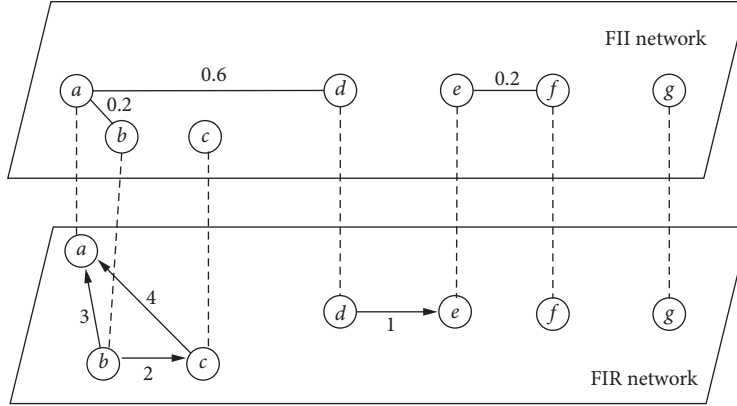


FIGURE 5: An example of FII and FIR networks.

infecting  $i$ . The values will depend only on the network structure from the previous section.

Note that out-neighbours and in-neighbours of node  $i$  are different only when the layer's edges are directed, i.e., the FIR network in this case. If the layer's edges are not directed, out-neighbours and in-neighbours are the same and are simply called neighbours.

At the starting time  $t = 1$ , a single node  $i_0$  is infected while all others are susceptible; i.e.,  $H_{i_0}^1 = \beta$  while  $H_i^1 = \alpha$  for

$i \neq i_0$ . Inductively, assume we know  $H_i^{t-1}$  for all  $i$ . Then, the node states of the next day are given as

$$H_i^t = \begin{cases} \alpha, & \text{if } r_i^{1,t} = 1, \\ \beta, & \text{if } r_i^{2,t} = 1, \end{cases} \quad (5)$$

where  $r_i^{1,t}$  and  $r_i^{2,t}$  can be determined by

$$\begin{cases} r_i^{1,t} = 1 - r_i^{2,t}, \\ r_i^{2,t} = I(H_i^{t-1} = \alpha) I \left( \left[ \sum_{l=1}^2 \sum_{\substack{j: i \in N^l(j), \\ H_j^{t-1} = \alpha}} I(p_{ji}^l > p_j^{l,t-1}) \right] > 0 \right) + I(H_i^{t-1} = \beta), \end{cases} \quad (6)$$

and where values of  $p_{ji}^1$  are calculated as follows:

$$\begin{cases} p_{ji}^1 = \frac{w_{ji}^1}{\sum_{k \in N^1(j)} w_{jk}^1}, \\ p_{ji}^2 = p_{ji}^a. \end{cases} \quad (7)$$

This means that infected nodes stay infected, whereas noninfected nodes become infected if, on some layers, one of its infected in-neighbours has a strong enough connection to overcome the given node's threshold at that time and layer. The information-spreading process begins with a single node  $i_0$  infected, while all other nodes are susceptible. Any node can be used as the starting node. The process ends when either no infected node has any noninfected out-neighbours, or  $t$  has reached some specified time limit. In our simulations, we used a time limit of  $t = 10$ . Note that because the spreading process depends on random draws ( $p_{ji}^{l,t-1}$ ), the process itself will be random.

A standard way to measure node  $i_0$ 's importance in a network is by the extent of the infection that begins at  $i_0$  [39]. To this end, we define the information-spreading rate of a node  $i_0$  as  $C(i_0) = n_{i_0}/n$ , where  $n_{i_0}$  is the number of infected nodes at the end of the spreading process and  $n$  is the total number of nodes. Because the spread of information depends on random variables ( $p_{ji}^{l,t-1}$ ),  $C(i)$  will itself be a random variable. Its expectation is too complicated to calculate explicitly; however, it can be estimated by repeated sampling as follows:

$$E(C(i)) \approx \frac{1}{N_0} \sum_{k=1}^{N_0} C_k(i), \quad (8)$$

where  $N_0$  is the number of trials used, i.e., the number of simulations with  $i$  as the starting node, and  $C_k(i)$  is the value obtained in the  $k^{\text{th}}$  trial. Due to the large size of our dataset and the resulting high computation time,  $N_0$  was set as 30. Users were then ranked according to their mean information-spreading rate across these 30 trials.

The simulation procedure implemented in MATLAB (MathWorks, Natick, MA, USA) is as follows. Let  $i_0$  be the node where the information starts. We track the spread of information throughout the network through the set of infected nodes, denoted as  $I$ , and the set of noninfected nodes, denoted as  $S$ :

Step 1: begin with the network nodes  $V = \{1, 2, \dots, n\}$  and edges  $E = \{a_{ij}^1, a_{ij}^2\}$  with weights  $W = \{w_{ij}^1, w_{ij}^2\}$  and starting node  $i_0$ .

Step 2: if  $i_0$  has no out-neighbours, halt and output  $t = 1$  and  $C(i_0) = 1/n$ . Otherwise, initialise  $t = 1$ ,  $I = \{i_0\}$ , and  $S = \{i \leq n: i \neq i_0\}$ .

Step 3: while  $t \leq 10$ ,

Step 3a: let  $\partial I = \{i \in I: N(i) \cap S \neq \emptyset\}$ . If this is empty, exit the while loop, otherwise continue.

Step 3b: for  $i$  in  $\partial I$ , (1) draw  $p_i^{1,t}$  and  $p_i^{2,t}$  from  $U(0, 1)$ , and (2) for each  $j \in N(i) \cap S$ , calculate  $p_{ij}^l$  as above. If  $p_{ij}^l > p_i^{l,t}$ , insert  $j$  into  $I$  and remove  $j$  from  $S$ .

Step 3c: update  $C(i_0) = \#(I)/n$  and  $t = t + 1$ .

Step 4: output  $C(i_0)$ .

An example of a two-layered network and an information-spreading history with node  $a$  as the starting node is provided in Figure 6.

Notice that node  $a$  cannot spread information to node  $b$  in the FIR layer but may do so via the FII layer. From there, the information can spread from node  $b$  to  $c$  via the FIR layer. For a different spreading path, notice that  $a$  cannot spread to  $d$  in the FIR layer but may do so through the FII layer. From there, the information can spread to  $e$  through the FIR layer.

The algorithm was run on this simple example once for each of the nodes as the starting node. Table 3 shows the average spreading rate with examples in Figure 5 across 30 trials.

In their corresponding simulations Nodes  $a$ ,  $b$ ,  $c$ , and  $d$  had the highest mean spreading rates, with a possibility of spreading information to all nodes, except for the isolated node  $g$ . Of these four, node  $b$  performed best in terms of average spreading rate. Also, Nodes  $c$  and  $d$  were more efficient by 1% compared to node  $a$ , although node  $a$  had more links than those two. This suggests that  $c$  and  $d$  have special positions in this network, which matches the graph.

The algorithm was then applied to the real Pollen Club dataset described in Section 3. The results for the top 20 of 129,362 users, as ranked by mean spreading rate across 30 trials in MATLAB, are summarised in Table 4. The standard deviations were not too large for only 30 trials and could likely be shrunk further by running more trials as needed.

In Table 4, IUG and OUG stand for ‘‘intermediary user group’’ and ‘‘ordinary user group,’’ respectively [36]. The OUG refers to customers who bought Huawei products and registered for the Pollen Club. The IUG refers to customers who received official training from Huawei and are willing to answer questions from other customers.

Specially, one of the IUG members is on the top of the list, indicating that the Pollen Club is organised by the customers themselves, saving the company and the effort of doing so. The number next to each OUG member is the user level. There are 12 levels in total, with higher levels indicating greater user experience. Users can advance their levels by joining activities in the Pollen Club. As can be seen, except for User 4948, the OUG members in our top 20 generally have high user levels. This shows that our method of ranking a user's spreading rate aligns well with Huawei's own method of ranking users.

To prove our model's effectiveness, comparisons were conducted using the probability 1 transformation model shown in Appendix. The results using our model were more consistent with trends observed in the real data.

**5.3. Feature Selection of the Spreading Process.** In this section, we investigate the relationships between spreading rate and network features that can be computed simply, without the

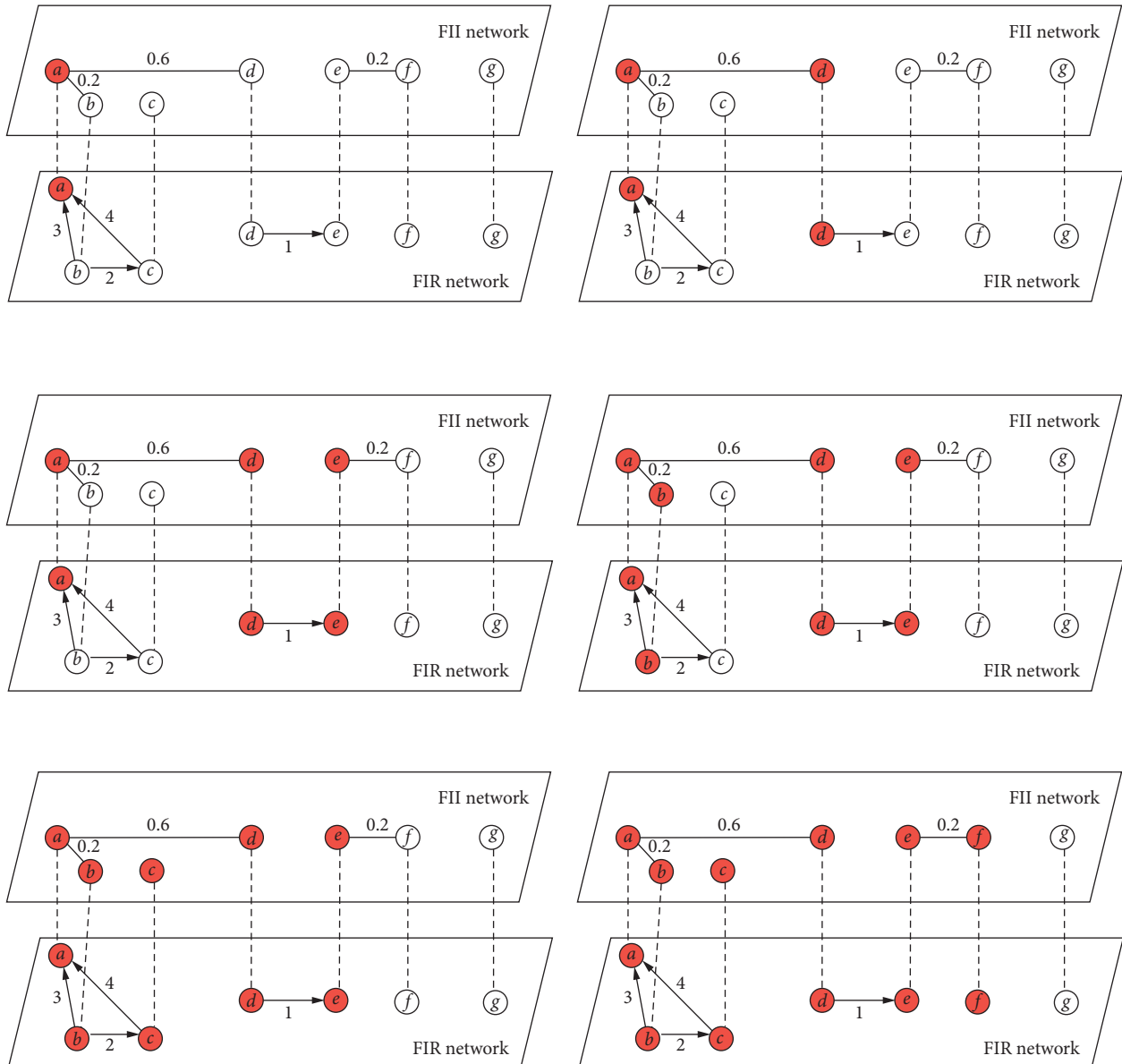


FIGURE 6: One possible information-spreading history in a 7-node network: (a)  $t = 1$ , (b)  $t = 2$ , (c)  $t = 3$ , (d)  $t = 4$ , (e)  $t = 5$ , and (f)  $t = 1$ .

TABLE 3: Mean and standard deviations of spreading rate of each node in Figure 5.

Source node	Mean spreading rate	Standard deviations of spreading rate
a	0.8048	0.1093
b	0.8333	0.0541
c	0.8143	0.0764
d	0.8095	0.0781
e	0.2762	0.0362
f	0.2810	0.0261
g	0.1429	0.0000

need for repeated simulations. Twenty-two network features [40] (denoted as  $x_1$  through  $x_{22}$ , respectively) of the two-layer model for a node  $i$  are considered. Ten are for the FII

layer, and 12 are associated with the FIR layer. The 22 features are listed in Table 5.

Each feature  $x_1$  through  $x_{22}$  is normalised to have a mean of 0 and a variance of 1. We then ask which of these features can predict the spreading rate. This provides insight into which features may cause a node to spread information more efficiently through the networks. We ran a random forest algorithm [41] by Breiman with the code referenced therein using scikit-learn [42]. Random forest is an ensemble method used to model regression in a nonlinear way.

Recall that the random forest algorithm consists of randomly selecting a subset of the users and a subset of the network features, forming a tree by selecting at each node the network feature and a boundary value for that feature. This splits the users into two branches at each such node, and continuing until in all branches, the number of users is

TABLE 4: Pollen Club Top 20 P10/P10 Plus users.

Rank	User ID	Group (level)	$C(i_0)$ (mean (std))
1	6050	IUG	0.0823 (0.0156)
2	65412	OUG (7)	0.0804 (0.0180)
3	78838	IUG	0.0800 (0.0157)
4	73370	IUG	0.0798 (0.0137)
5	62489	OUG (5)	0.0772 (0.0150)
6	9014	IUG	0.0762 (0.0206)
7	78835	OUG (6)	0.0761 (0.0205)
8	59364	OUG (10)	0.0757 (0.0217)
9	1008	OUG (5)	0.0753 (0.0167)
10	126016	OUG (6)	0.0743 (0.0191)
11	50128	OUG (9)	0.0742 (0.0199)
12	9381	OUG (7)	0.0729 (0.0204)
13	71961	OUG (6)	0.0725 (0.0201)
14	4948	OUG (2)	0.0716 (0.0150)
15	93168	OUG (7)	0.0710 (0.0210)
16	89378	OUG (7)	0.0708 (0.0216)
17	91874	OUG (7)	0.0704 (0.0154)
18	79875	OUG (6)	0.0700 (0.0188)
19	85851	OUG (7)	0.0697 (0.0299)
20	84435	OUG (5)	0.0695 (0.0141)

TABLE 5: Meaning of features.

Network layer	Feature name	Symbol	Meaning
	Degree	$x_1$	The number of neighbours of node $i$
	Weighted degree	$x_2$	The sum of the weights of all edges containing node $i$
	Eccentricity	$x_3$	The maximum distance from node $i$ to any other node
	Closeness centrality	$x_4$	The reciprocal of the sum of the shortest distances from node $i$ to all other nodes
	Harmonic closeness centrality	$x_5$	The number of pairs of distinct nodes $j, k$ both different from node $i$ for which the shortest path between $j$ and $k$ passes through node $i$
FII	Betweenness centrality	$x_6$	The number of pairs of distinct nodes $j, k$ both different from node $i$ for which the shortest path between $j$ and $k$ passes through node $i$
	Hub score	$x_7$	Let $A$ be the adjacency matrix of the network. The hub score of node $i$ is the $i$ th component of the eigenvector corresponding to the maximum eigenvalue of $A\tau A$
	Authority score	$x_8$	It equals to the weighted sum of its neighbours' hub score. That is, $x_8(i) = \sum w_{ij} x_7(j) j \in N(i)$
	Local clustering coefficient	$x_9$	It is given by the proportion of links between the vertices within node $i$ 's neighbourhood divided by the number of links that could possibly exist between them
	Eigen centrality1	$x_{10}$	The same as hub score except for we use the matrix adjacency matrix $a$ instead of $A\tau A$ here
	Indegree	$x_{11}$	The number of number of edges ending at node $i$
	Outdegree	$x_{12}$	The number of number of edges starting at node $i$
	Weighted indegree	$x_{13}$	The sum of the weight of all edges ending at node $i$
	Weighted outdegree	$x_{14}$	The number of number of edges ending at node $i$
	Eccentricity	$x_{15}$	Same as $x_3$
	Closeness centrality	$x_{16}$	Same as $x_4$
FRI	Harmonic closeness centrality	$x_{17}$	Same as $x_5$
	Betweenness centrality	$x_{18}$	Same as $x_6$
	Hub score	$x_{19}$	Same as $x_7$
	Authority score	$x_{20}$	Same as $x_8$
	Page rank coefficient	$x_{21}$	An algorithm introduced by Google to rank web pages. It does so by estimating the probability that a person randomly clicking on links will arrive at the given page
	Local clustering coefficient	$x_{22}$	It is given by the proportion of links between the vertices within node's neighbourhood divided by the number of links that could possibly exist between them

TABLE 6:  $R^2$  from random forest regression.

Round	Training dataset	Testing dataset
1	0.9703	0.7818
2	0.9706	0.7755
3	0.9704	0.7758
4	0.9697	0.7970
5	0.9699	0.7913

TABLE 7: Rank of the indexes for features in 10 rounds.

Rank	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Round 8	Round 9	Round 10
1	16	16	16	16	16	16	16	16	16	16
2	17	17	17	17	17	17	17	17	17	17
3	15	15	15	15	15	15	15	15	15	15
4	19	19	19	19	19	19	19	19	19	19
5	12	22	12	22	22	22	22	22	22	22
6	22	12	22	12	12	12	12	12	12	12
7	14	14	14	14	14	14	14	14	14	14
8	18	18	18	18	18	18	18	18	18	18
9	20	20	20	20	20	20	20	20	20	20
10	13	13	13	13	10	10	13	13	13	13
11	11	11	11	11	5	7	11	11	11	10
12	10	21	21	10	13	8	21	21	21	11
13	7	7	10	21	7	11	10	10	10	7
14	21	5	8	7	11	5	9	7	7	1
15	8	10	5	8	8	13	7	5	5	9
16	1	4	7	5	21	4	5	8	8	2
17	2	8	4	4	4	21	4	4	4	21
18	9	3	1	2	1	2	8	2	1	8
19	5	1	2	3	2	1	2	1	2	5
20	4	2	9	9	6	9	3	6	9	3
21	3	6	3	1	3	3	1	9	3	4
22	6	9	6	6	9	6	6	3	6	6
23	16	16	16	16	16	16	16	16	16	16
24	17	17	17	17	17	17	17	17	17	17
25	15	15	15	15	15	15	15	15	15	15

at most some threshold value. Given a new vector of values for the features, each tree is used to predict a value for the spreading rate. The forest as a whole then makes a prediction by averaging the values of each of the individual trees' predictions. The key parameters in this process are follows:

- (i)  $n_{tree}$ : the number of trees used. Increasing this value should decrease variance without leading to overfitting [43]. As suggested by reference [44], the number of trees was set to 500.
- (ii)  $m'$ : the number of features selected in each tree. We followed the suggestion in [44] to take  $m' = \log_2 m \approx 4$ , where  $m$  is the total number of features.
- (iii)  $n'$ : the number of users to sample in each tree. We used all the users, i.e.,  $n' = n = 129362$ .
- (iv)  $leaf'$ : the maximum leaf size, which controls when the construction of each tree halts. We used  $leaf' = 1$ , i.e., we continued splitting until there was only one user in each branch.

Recall that after running the random forest algorithm on a training set, it outputs a regression function. Given a new point  $x = (x_1, \dots, x_{22})$ , the regression function outputs a predicted adjusted spreading rate  $f(x)$ , which can then be compared with the spreading rates from the simulations. We performed five-fold cross-validation on the Pollen Club dataset together with the spreading rates. Table 6 shows the resulting  $R^2$  values when these regression functions were tested against the training set and the testing set.

The random forest algorithm can also be used to rank the importance of network features. Each tree ranks its  $m'$  selected features according to the decrease in variance at the corresponding node. That is, the variance in spreading rates at the parent node is compared to the sum of variances at the two child nodes; the greater the decrease, the more important that feature according to that tree. During the same process, the decreasing variance can be calculated for each tree. Finally, the decreasing variance for each of the  $m$  features is summed as the corresponding feature importance.

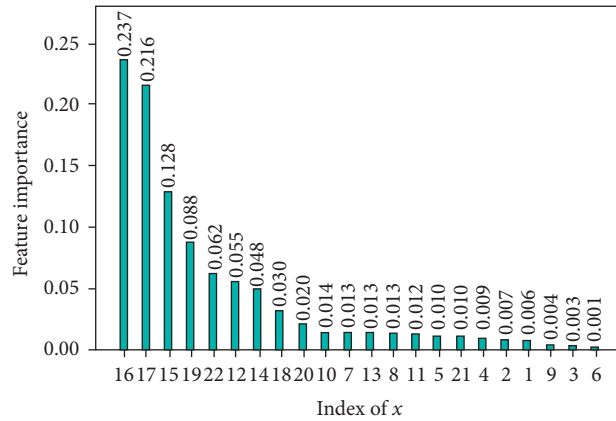


FIGURE 7: The importance ranking of features.

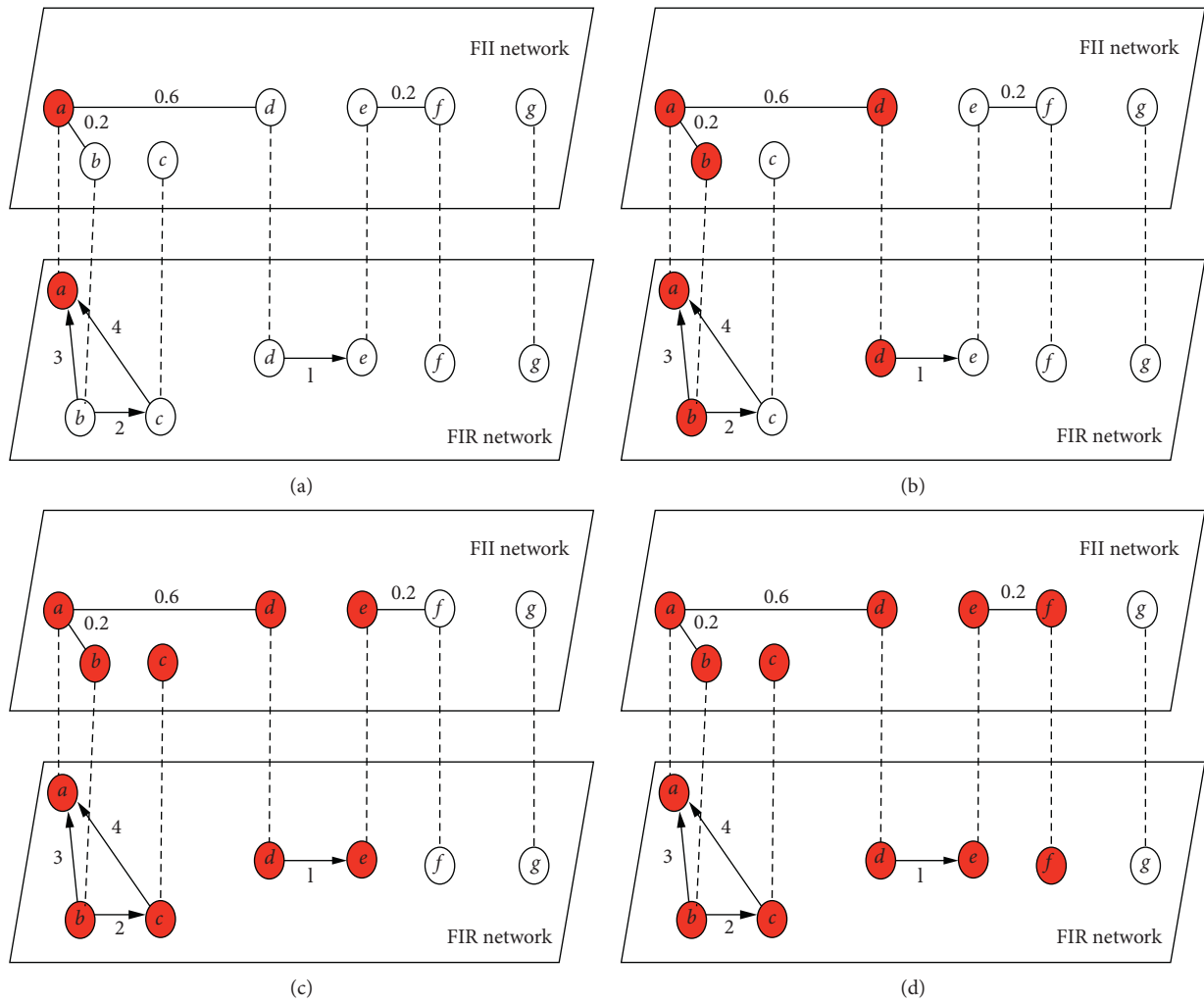


FIGURE 8: Probability 1 information-spreading history in a 7-node network: (a)  $t=1$ , (b)  $t=2$ , (c)  $t=3$ , and (d)  $t=4$ .

To check the stability, we again divided the dataset into 10 pieces, to check whether each piece yielded a similar ranking of the network features. The result is shown in

Table 7. We used Kendall's W-test [45] to assess the degree of agreement in the rankings across the ten runs and obtained a W value of 0.7570. According to the  $\chi^2$  test, this corresponds

TABLE 8: Top 20 P10/P10 Plus users.

Rank	User ID	Group (level)	$C(i_0)$
1	119794	OUG(2)	0.904757
2	119744	OUG(1)	0.904749
2	119781	OUG(1)	0.904749
2	119803	OUG(1)	0.904749
5	119741	OUG(2)	0.904741
5	119802	OUG(1)	0.904741
5	119823	OUG(1)	0.904741
8	119772	OUG(4)	0.904733
8	119826	OUG(1)	0.904733
10	119771	OUG(3)	0.904725
10	50128	OUG(2)	0.904725
12	9381	OUG(3)	0.904701
13	71961	OUG(5)	0.904693
13	4948	OUG(3)	0.904693
15	93168	OUG(1)	0.904686
16	89378	OUG(6)	0.904670
16	91874	OUG(3)	0.904670
16	79875	OUG(2)	0.904670
16	85851	OUG(1)	0.904670
20	84435	OUG(1)	0.904662

to a  $p$  value of  $3.8064 \times 10^{-11}$ , indicating very high confidence in a genuine agreement in the rankings.

Lastly, we obtained a single overall ranking based on the whole dataset. The results are shown in Figure 7. The top nine features all belong to the FIR layer. The most important features are closeness centrality ( $x_{16}$ ) and harmonic closeness centrality ( $x_{17}$ ) in the FIR layer. In the FII layer, the most important feature is eigencentrality; it indicates that the user in the central position in the FIR network can affect the information spreading. Also, the quality of the neighbours in the FII network plays an important role in transformation.

## 6. Conclusions

In this paper, we proposed the NSECD to identify the key moment in topic growth in a VC after a new product is introduced. A two-layer model was developed for assessing information propagation in a VC, where information can flow among users either by replies to posts (the FIR layer) or searching for topics of common interest (the FII layer). We applied this model to our case study, which focused on the P10/P10 Plus device in Huawei's Pollen Club, to identify which users were most effective at spreading information through the network. Lastly, we compared these results with commonly used network features using a random forest algorithm and found that spreading effectiveness correlated best with closeness centrality and harmonic closeness centrality in the FIR layer and eigencentrality in the FII layer.

We have two suggestions for how our model may be improved in future research. First, the infectiousness formula in the FIR layer can be modified to consider not just the quantity of postreplies but also their quality. For instance, natural language analysis [46] could be used to score the quality of posts. Second, the network model can be extended to have more than two layers. For example, many VCs enable

users to follow other users; thus, a third layer could be used to capture these follower relations.

In conclusion, this research introduces new concepts for network theory and provides suggestions for how companies could manage their VCs.

## Appendix

Comparisons with a probability 1 transformation model: the probability 1 transformation model is defined as follows. The information-propagation model will also be based on a simple two-state framework; that is, at any given time, a node is in one of two possible states. In contrast to the model described in Section 5.2, if node  $i$ 's state is infected at time  $t$ , the states of neighbours (*i.e.*  $N(i) = \cup N^l(i)$ ) in all layers of node  $i$  will turn into infected states.<sup>1</sup> This means that once infected, it will transform the information to all neighbours with probability 1 (Figure 8).

To illustrate this clearly, the same example used in Figure 5 in Section 5.2 is demonstrated. If node  $a$  is chosen as the initial source of the information, the information will spread to all other nodes in both layers, except for the isolated node, in four steps. Because the probability 1 model does not involve uncertainty, it can be implemented with one time step. To coordinate with our model as described in the main text, the maximum spreading time was set as 10. Where the meaning of the headline of the table is the same as that of Table 4. The results shown in Table 8 suggest that the lower-grade OUG members are more influential, which contradicts the meaning of the grade. This is due to the probability 1 model, which considers only the degree of the nodes and ignores the weight preference and uncertainty. Additionally, our model was more effective in predicting how information spreads among the nodes, compared with the probability 1 model.



## Data Availability

All data, the models used during the study that appear in the submitted article, and the original 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 study was supported by the National Natural Science Foundation of China (no. 71974115), the Innovation Method Fund of China (2018IM020200), and a theme-based research grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (project no. T12-710/16-R).

## References

- [1] D. Ariely, *Predictably Irrational: The Hidden Forces that Shape Our Decisions*, HarperCollins, New York, NY, USA, 2008.
- [2] R. J. Brodie, A. Ilic, B. Juric, and L. Hollebeek, "Consumer engagement in a virtual brand community: an exploratory analysis," *Journal of Business Research*, vol. 66, no. 1, pp. 105–114, 2013.
- [3] S.-C. Chu and Y. Kim, "Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites," *International Journal of Advertising*, vol. 30, no. 1, pp. 47–75, 2011.
- [4] M. E. Zaglia, "Brand communities embedded in social networks," *Journal of Business Research*, vol. 66, no. 2, pp. 216–223, 2013.
- [5] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, no. 5916, pp. 892–895, 2009.
- [6] S. G. B. Roberts, R. I. M. Dunbar, T. V. Pollet, and T. Kuppens, "Exploring variation in active network size: constraints and ego characteristics," *Social Networks*, vol. 31, no. 2, pp. 138–146, 2009.
- [7] S. Boccaletti, V. Latora, and Y. Moreno, "Complex networks: structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.
- [8] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
- [9] M. E. J. Newman, *Networks: An Introduction*, Oxford University Press, Oxford, UK, 2010.
- [10] X.-W. Wang, Y.-M. Cao, and C. Park, "The relationships among community experience, community commitment, brand attitude, and purchase intention in social media," *International Journal of Information Management*, vol. 49, pp. 475–488, 2019.
- [11] D. Baum, M. Spann, J. Füller, and C. Thürridl, "The impact of social media campaigns on the success of new product introductions," *Journal of Retailing and Consumer Services*, vol. 50, pp. 289–297, 2019.
- [12] P. Holme and J. Saramäki, "Temporal networks," *Physics Reports*, vol. 519, no. 3, pp. 97–125, 2012.
- [13] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks," *Journal of Complex Networks*, vol. 2, no. 3, pp. 203–271, 2014.
- [14] C.-W. Tsai and S.-J. Liu, "SEIM: search economics for influence maximization in online social networks," *Future Generation Computer Systems*, vol. 93, pp. 1055–1064, 2019.
- [15] C. Sang, T. Li, S. Tian, Y. Xiao, and G. Xu, "SFTRD: a novel information propagation model in heterogeneous networks: modeling and restraining strategy," *Physica A: Statistical Mechanics and Its Applications*, vol. 524, pp. 475–490, 2019.
- [16] J. Park, J. Ryu, and S.-B. Yang, "ActiveDBC: learning knowledge-based information propagation in mobile social networks," *Wireless Networks*, vol. 25, no. 4, pp. 1519–1531, 2019.
- [17] P. Lu, L. Deng, and H. Liao, "Conditional effects of individual judgment heterogeneity in information dissemination," *Physica A: Statistical Mechanics and Its Applications*, vol. 523, pp. 335–344, 2019.
- [18] L. Li, Q. Zhang, J. Tian, and H. Wang, "Characterizing information propagation patterns in emergencies: a case study with Yiliang earthquake," *International Journal of Information Management*, vol. 38, no. 1, pp. 34–41, 2018.
- [19] R. Zeng and D. Zhu, "A model and simulation of the emotional contagion of netizens in the process of rumor refutation," *Scientific Reports*, vol. 9, no. 1, pp. 1–15, 2019.
- [20] L. C. Hsu, "Investigating the brand evangelism effect of community fans on social networking sites," *Online Information Review*, vol. 43, no. 5, 2019.
- [21] X. L. Wan, Z. Zhang, and X. X. Rong, "Exploring an interactive value-adding data-driven model of consumer electronics supply chain based on least squares support vector machine," *Scientific Programming*, vol. 2016, Article ID 3717650, , 2016.
- [22] L. A. Huo and Y. Y. Cheng, "Dynamical analysis of a IWSR rumor spreading model with considering the self-growth mechanism and indiscernible degree," *Physica A: Statistical Mechanics and Its Applications*, vol. 536, Article ID 120940, 2019.
- [23] S. Xu, P. Wang, and C. X. Zhang, "Spectral learning algorithm reveals propagation capability of complex networks," *IEEE Transactions on Cybernetics*, vol. 49, no. 12, pp. 4253–4261, 2018.
- [24] Z. Shao, S. Liu, and Y. Zhao, "Identifying influential nodes in complex networks based on Neighbours and edges," *Peer-to-Peer Networking and Applications*, vol. 12, no. 6, pp. 1528–1537, 2019.
- [25] Y. Wang, B. Chen, and W. Li, "Influential node identification in command and control networks based on integral k-shell," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 6528431, , 2019.
- [26] R. Escalante and M. Odehnal, "A deterministic mathematical model for the spread of two rumors," *Afrika Matematika*, vol. 31, no. 2, pp. 315–331, 2020.
- [27] X. Li, Y. Liu, and C. Zhao, "Locating multiple sources of contagion in complex networks under the SIR model," *Applied Sciences*, vol. 9, no. 20, p. 4472, 2019.
- [28] J. P. Zhang, H. M. Guo, and W. J. Jing, "Dynamic analysis of rumor propagation model based on true information spreader," *Acta Physica Sinica*, vol. 68, no. 15, Article ID 150501, 2019.
- [29] X. Xiong, S. Qiao, and Y. Li, "Affective impression: sentiment-awareness POI suggestion via embedding in heterogeneous LBSNs," *IEEE Transactions on Affective Computing*, 2019.

- [30] X. Xiong, S. Qiao, and N. Han, "Where to go: an effective point-of-interest recommendation framework for heterogeneous social networks," *Neurocomputing*, vol. 373, pp. 56–69, 2020.
- [31] X. Xiong, S. Qiao, Y. Li, N. Han, G. Yuan, and Y. Zhang, "A point-of-interest suggestion algorithm in multi-source geo-social networks," *Engineering Applications of Artificial Intelligence*, vol. 88, Article ID 103374, 2020.
- [32] X. Xiong, F. Xiong, and J. Zhao, "Dynamic discovery of favorite locations in spatio-temporal social networks," *Information Processing & Management*, vol. 57, no. 6, Article ID 102337, 2020.
- [33] W. Zhang, J. Yang, X.-Y. Ding, X.-M. Zou, H.-Y. Han, and Q.-C. Zhao, "Groups make nodes powerful: identifying influential nodes in social networks based on social conformity theory and community features," *Expert Systems with Applications*, vol. 125, pp. 249–258, 2019.
- [34] Huawei Pollen Club, "Huawei pollen club," 2020, <https://club.huawei.com/forum.html>.
- [35] P. V. Marsden, "Interviewer effects in measuring network size using a single name generator," *Social Networks*, vol. 25, no. 3, pp. 1–16, 2003.
- [36] Q. C. Meng, Z. Zhang, and X. L. Wan, "Properties exploring and information mining in consumer community network: a case of Huawei pollen club," *Complexity*, vol. 2018, Article ID 9470580, , 2018.
- [37] R. Vernon, "International investment and international trade in the product cycle," *International Economic Policies and Their Theoretical Foundations*, Academic Press, vol. 80, no. 2, pp. 415–435, 1992.
- [38] X. Q. He and W. Q. Liu, *Applied Regression Analysis*, China Renmin University Press, Beijing, China, 2nd edition, 2014.
- [39] X. L. Wan, Z. Zhang, and C. Zhang, "Stock market temporal complex networks construction, robustness analysis, and systematic risk identification: a case of CSI 300 index," *Complexity*, vol. 2020, Article ID 7195494, , 2020.
- [40] G. R. Chen, X. F. Wang, and L. I. Xiang, *Introduction to Complex Networks: Models Structures and Dynamics*, Higher Educational Press, Beijing China, 2012.
- [41] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [42] F. Pedregosa, G. Varoquaux, and A. Gramfort, "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [43] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016.
- [44] L. Raynal, J.-M. Marin, P. Pudlo, M. Ribatet, C. P. Robert, and A. Estoup, "ABC random forests for Bayesian parameter inference," *Bioinformatics*, vol. 35, no. 10, pp. 1720–1728, 2019.
- [45] A. P. Field, *Kendall's Coefficient of Concordance*, Wiley StatsRef: Statistics Reference Online, Hoboken, NJ, USA, 2014.
- [46] X. Xiong, Y. Li, R. Zhang, Z. Bu, G. Li, and S. Ju, "DGI: recognition of textual entailment via dynamic gate matching," *Knowledge-Based Systems*, vol. 194, no. 22, Article ID 105544, 2020.

## Research Article

# Athlete Social Support Network Modeling Based on Modern Valence Bond Theory

Ningshe Zhao 

*School of Information Engineering, Xi'an University, Xi'an 710065, China*

Correspondence should be addressed to Ningshe Zhao; [zhaoningshe@foxmail.com](mailto:zhaoningshe@foxmail.com)

Received 25 June 2020; Revised 9 October 2020; Accepted 1 November 2020; Published 24 November 2020

Academic Editor: Fei Xiong

Copyright © 2020 Ningshe 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.

Based on the Valence Bond theory, an attempt is proposed to the complex network. The principle of chemical bonding of the basic particles that make up the substance creates a metaphor between the formation of social networks. By analyzing the integration of atoms by relying on the chemical bonds between particles, then the social basis for the connection between social network nodes should depend on the tangible or intangible attribute resources that characterize social capital around the main node. Based on the above analysis, the social node is divided into active nodes and passive nodes, and a dynamic model of social network formation is proposed, the Valence Bond model of social network. Through this model, the actual athlete group nodes are depicted, and the representation of the model and the evolution of network structure are given with the actual data.

## 1. Introduction

Big data and social network have become a powerful tool to study social phenomena, and the research on social network is still very popular [1–6]. The research of network dynamics is that, under the impetus of “external stimulus” or the trigger of “internal message,” the state of node itself or the information contained in the network changes (the propagation of message), or the change of connection relationship between nodes [7–10]. This process is similar to the formation and decomposition of matter. How do the molecules that make up matter come together into stable states? Atoms are combined with a common electric charge. Because of the simultaneous existence of positive and negative charge, it is inevitable that the effects of both gravity and repulsion between molecules and molecules exist at the same time, and the distance shows different forces at different times. From a different point of view, each matter is a network of particles, connected by gravity and repulsion. The formation of the network should also have a mechanism, that is, there are two kinds of links at the same time, and the social network of the widespread competition and cooperative relations are similar. The connection between nodes is just a superficial phenomenon, and the real connection is the

attraction of node negative resources to other nodes. From the point of view of social networks, it is similar to social capital [11–15]. From the macropoint of view, it may affect the decision of the main node to establish the connection, or the recommendation of the link tendency [10, 12, 16–18]. On the microlevel, from a node point of view, the capital itself has the source of gravity. In fact, there are two kinds of nodes in the network, one is the principal node, which is the active node and is the demand, and the other is the resource node, which is the passive node and is to meet the demand. The nature and number of resource nodes owned by each principal node are different, resulting in different sizes of their influence. The connection of active nodes to each other is essentially driven by the demand for resources. The more the active node's perception is stronger, the greater the probability of establishing a connection with other principal nodes, such as the principal node in the position of the structure hole. If the network is depicted in this way, it is more clearly described and more predictable, for example, to know whether there may be a connection between node  $i$  and node  $j$  in the future, as long as the analysis the resources of node  $i$  and node  $j$ . If they share resource nodes, there is already a connection, if they do not have a similar resource, they have a higher probability of a connection to each other,

and if they have the same resource, the probability of a connection to each other is very small. Considering network evolution, a portion of the principal node will multiply the required resources through the adjacent principal node, thus disconnecting the connection swayed through the resource. The reality is that, after the main node demand changes, it will eliminate old resources with a high probability and create new connections to meet new demands. In this way, the network has become a new pattern. Then, there is a theoretical basis.

Any network consists of two nodes, the principal node and the resource node on the dependency. The principal node provides the demand and generates the power of evolution, and the resource node realizes the transfer or function and completes the evolution. There is a demand (whether subjective or objective) between the principal node and the node with related resources to produce a force, resulting in a connection. The different states of the resource cause the network to have different states. Resource state has parasitic state, and free state and free state resources will be adsorbed by the main node, if the parasitic resources are discarded by the main node, it becomes a free state or undefined state. The two states are different, free state is not bound by the principal node, and undefined resources are parasitic to the principal node.

In practice, some physical connections in the social network may be very low, and physical connections with relatives between two relatives exist, but there is little interaction, and that connection can actually be ignored. What is the reason? If the principal node has no requirements, the role of the fact will not be produced. According to the above phenomenon, the network can be simplified, some weak contacts removed, making the network more sparse, and the advantage is more convenient for storage and relationship delivery. The key is to record the needs of the main node, and there is a demand of the main node may have a role to occur, such as competitive resources, attract talent, request assistance. To study this network dynamics, the demand of the principal node is the driving force of the whole network. Social nodes and life nodes have certain independent intentions, which belong to the main nodes.

From the point of view of social capital theory, Pierre Bourdieu pointed out that social capital is a collection of all kinds of social resources, which is formed by the accumulation of the quantity and quality of capital owned by individuals as actors in various social network relations [10, 12, 17]. This accumulation mainly depends on the size and initiative of the network relationship in which the individual is located [18–20]. The above theory provides microsupport for this view.

## 2. Social Network Representation Based on Valence Bond Theory

In modern Valence Bond theory [21–23], there are two basic points of view. (1) When two atoms are close, unpaired valence electrons in the opposite direction of spin can be paired to form a covalent bond. (2) The more the atomic orbits of the bonded atoms overlap with each other, the more

stable the covalent bonds are formed. Therefore, the covalent key should be formed as far as possible along the direction of the largest overlap of the atomic orbit, which is the principle of maximum overlap of atomic orbits. When the composition is stable, it is in the lowest energy state. Communities and even social networks can be compared to the formation of compounds. The formation of covalent compounds reflects the sharing of resources, while the ion key reflects the transfer of resources, such as trade relations and exchange relations.

The principal node is not an isolated existence, but has a certain companionship. From the phenomenon point of view, this accompaniment is determined by the performance of the principal nodes interacting with each other, but from the microanalysis, it is determined by the attributes. For a long time, we do not pay enough attention to the analysis of node properties [23–28]. However, the behavior of the principal node depends on its properties. The principal node is active and operates within the network at a certain probability. For social nodes, the principal nodes have a certain degree of free, while the properties are parasitic.

*2.1. Model Thought Description.* For the principal node in the network, the principal node  $i$  is expressed in the vector  $\eta_i = (p_{i1}, p_{i2}, \dots, p_{in})$ , where  $p_{in}$  is an integer, and  $p_{ij}$  ( $i, j$  as integer) is an integer. When  $p_{ij} > 0$ , it indicates the strength of the attribute  $k$  professional support capability. When  $p_{ik} < 0$ , it indicates the strength of the professional demand for the attribute  $k$ . Vectors  $\lambda_i = (h_{i1}, h_{i2}, \dots, h_{in})$ , and  $h_{ij} \in [0, 1]$  is the attribute weights.

NA (node activity): within the network  $W$ , for any principal node  $i$ , let

$$S_i = [\eta_i, \lambda_i]. \quad (1)$$

It indicates the activity of the node, represents the measurement of the activity capacity of the node, and the two nodes with the same NA have the same activity ability.

NSA (network system activity): for network  $W$ , the system activity is weighted by the activity of all nodes, namely,

$$A_w = \sum_{i=1}^N k_i S_i, \quad (2)$$

where  $k_i$  denotes the influence weight of the main node; obviously, the more influential main nodes, the more stable the network performance.

Network activity characterizes the stability of the network; it is an integer, and the activity is positive network. The greater the activity is, the more the network tends to diverge, with a greater probability of node free. For networks with negative activity, the smaller the activity is, the more the network tends to converge and the less likely the nodes to escape.  $W$  is usually a small world network that can grow by node aggregation connecting to form a larger network or by connections break which split into multiple small world networks.

When the node is aggregated, the key value of the passive node is consumed by the active node, the attribute value of

the active node is partially satisfied with the key bit, and the attribute value of the node changes in real time. When the internode bond breaks, the free node recovers the original activity. Network system activity also changes.

**2.2. Node Aggregation.** The social principal node activity has a certain suddenness, where the probability obeys the power law distribution, that is,  $P(r) \sim r^{-\alpha}$ , where  $r$  is arranged in the descending order of activity.

The principal node with negative activity is due to the inertia, forming the adsorption effect, when there are multiple active positive body nodes will expand competition and when there are multiple negative nodes at the same time, and the number of positive nodes is small. There are also negative nodes in order to compete with the negative node swell, which can be called the key position competition. The smaller the negative node activity, the greater the adsorption force, the greater the positive node NA, and the stronger the overall competitiveness. The competitively successful positive node produces a new connection, and the node state that does not compete to the key bit remains unchanged.

For any principal node, the smaller the  $S_i$ , the smaller the probability of movement; when there is negative infinity, the probability of movement tends to be 0. Starting with probability  $p_i$ , if node  $i$  meets  $j$ , check the attribute, check whether the attribute meets the conditions of integration (one of which is negative), and form a composite node (node group) at the meeting point to form a temporary steady state. If multiple nodes meet at the same time, then the active sort, the active high node mobility is strong, the first integration, the integration from the most inert nodes. As shown in Figure 1, when there are  $a$ ,  $b$ , and  $c$  three positive nodes, node activity in turn decreases, and they are combined with negative nodes, and finally form a node group, as shown in the right part of Figure 1, and the node group external performance is like that of a large node.

**2.3. Node Group Fission.** The properties of the principal node are dynamic indicators, and each attribute has a continuous change process according to a certain rate.

There are two aspects of change. On the one hand, the adsorption of the inert node is attenuated, and when the gravitational force between the active node is insufficient, the active node will escape and create a new movement. On the other hand, a node activity is enhanced, and the current key position and its activity no longer match with the activity overflow and natural escape. Typically, of course, both aspects of the situation exist at the same time over a historical period of time. The probability of escape is related to the amount of active spillage,  $P(e) = k\Delta S^\beta$ . The resulting escape causes the connection to be disconnected, as shown in Figure 2.

**2.4. Connection Prediction Algorithm for the Evolution of Network Structure.** For a complex network  $W$  with a certain scale,  $W = (W_c, W_s)$ , where  $W_c$  is network composition, while  $W_{S0}$  is initial network structure. The trend of network



FIGURE 1: Node aggregation.



FIGURE 2: Node group fission.

evolution is spending system activity; the smaller the system activity, the more stable the network.

For all nodes, a power-based law produces a period of random evolution.

- (1) For all nodes, compute node activity, the production of network activity descending sequence  $S$ .

Connection establishment prediction: (2) to (5) steps.

- (2) To sequence  $S$ , take a node  $i$  with probability  $p_i$  to start a random walk, and in a unit time, it is only allowed to walk once.
- (3) If the node wanders away, start with the least probability of negative node  $j$  and execute (5), otherwise, remove one node and execute (2).
- (4) If the current node is the last positive node, the algorithm ends.
- (5) If node  $i$  and  $j$  combination conditions are true, the probability  $p_j$  is the competition key; if the competition is successful, then turn to (2), otherwise  $j$  removes the position of a negative node until the competition is successful or the negative node traversal ends and then turns to (2).

Connection fracture prediction: (6) to (7).

- (6) Check the connection that has been formed (node group), take one of its edges, and if the bonding activity overflow  $\Delta S = 0$ , then it remains unchanged. If the bonding activity overflow  $\Delta S > 0$ , the probability of escape is  $p$ . If the bonding activity overflow  $\Delta S < 0$ , there is a probability  $p$  for disconnection (passive node rejection).
- (7) If this is the last edge, continue, else get the next edge, execute (6).
- (8) If this is the last node, the algorithm ends, else get the next node.

**2.5. Probability Function Construct.** The constructed probability function is  $p_i = (a \tan(S/A)/\pi) + 0.5$ , where the value of the shape adjustment parameter  $A$  is related to the value field of  $S_i$ . The function image is shown in Figure 3.

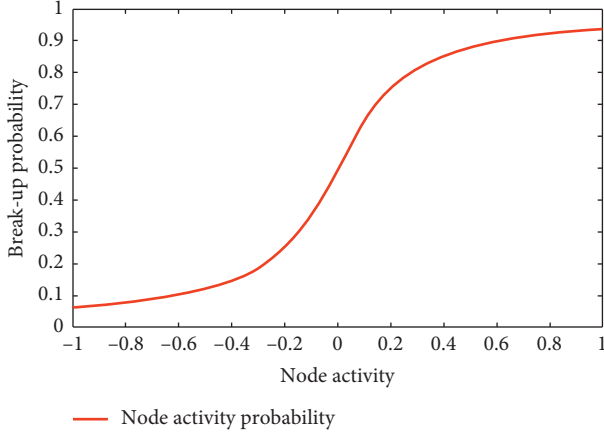


FIGURE 3: Node activity probability function construct.

**2.6. Node Attribute Field Classification.** For different applications, there are different weight configurations. The weight of the properties in this model gives the priority of the probe when the node establishes the connection.

The attribute field of the node can be divided into polar field and nonpolar field, the nonpolar field means that the content of the attribute is only different, there is no comparison on the quantity, and the description field belongs to the nonpolar data, and therefore does not consider.

As for the storage of the network, because the connection of the network in this paper is based on the application of social capital, it is necessary to use more special expressions, such as the element value of the adjacency matrix in two cases. (1) There are links. Storing elements as an attribute subscript indicates that the current connection is linked through that attribute. (2) In the case of no connection, a subscript of 0 is saved as  $-1$ , indicating no connection.

### 3. Modelling Athletes' Social Support Networks

In sports applications, the model is suitable for the description of athletes' networks. In the study of athletes' social support networks, we get a static link between athletes and answers on multiple topics. After the analysis of the answers of the theme questionnaire, we can get a description of some attributes of athletes. From the athlete's social support network questionnaire, the principal node activity usually has a certain motivation, which comes from individual needs, such as daily help, economic help needs and economic ability inversely proportional, emotional help and gender differences, and personality-related, employment support needs, and employment information holding degree, symn. Achievement requirements are related to achievement support capabilities.

**3.1. Attribute Analysis of Athlete Nodes.** The descriptive information about the subject node is gender, age, sports program, and education level. These descriptive attributes are related to certain support attributes, such as education, which may be related to social support skills, income, etc.

The issue of the relevance of research properties is beyond the scope of this article.

Social support attribute, economic help contact can use income as a capacity indicator. When income is less than 0, only expenditure exists. The more the spending is, the greater the degree of economic help needed is. Achievement needs are positive values that are related to achievement support capabilities. A higher value indicates a stronger achievement support ability. If the achievement support ability takes a negative value, the smaller the value is, the stronger the achievement demand is.

**3.2. Athlete Social Support Attributes.** By refining, the social support attribute vector for any athlete can be expressed (actual support ability, emotional support ability, employment support ability, achievement support ability, social interaction ability, marriage topic support ability, and management topic support ability) as depicted in Table 1. The actual support capacity refers to the degree of access and support for daily help. Table 2 is the normalization of Table 1 data to weaken the influence of dimension, where the last column is the calculated node activity.

**3.3. Athlete's Social Support Network Model.** If there are  $n$  athletes in this study and each athlete has  $m$  attributes, then the principal node vector is  $\beta_i = (p_{i1}, p_{i2}, \dots, p_{im})$ ,  $i \in [1, n]$ . Then, athlete network  $W$  is shown as follows:

$$W = (W_c, W_s), \quad (3)$$

where  $W_{c_{n \times m}} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nm} \end{bmatrix}$  and

$$W_{s_{n \times n}} = \begin{bmatrix} \theta & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \theta & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \theta \end{bmatrix}. W_c \text{ is an } n \times m \text{ network com-}$$

position matrix, each row represents a node, each column represent node attribute.  $W_s$  is an  $n$ -step network structure matrix, rows and columns represent network nodes, and it is a super-adjacent matrix. In  $W_s$ ,  $\alpha_{ij} = (b_1, b_2, \dots, b_m)$  is a Boolean vector that represents the vector of node  $N_i$  connected to node  $N_j$ , and the component  $b_k = 1$  indicates that node  $i$  and node  $j$  are connected at property  $k$ .

### 4. Representation and Prediction of Athlete Support Network

A social network observation model can be represented by a simple formula:

$$R = UV^T + e, \quad (4)$$

where  $U$  and  $V$  are  $m \times d$  and  $n \times d$  matrices, respectively.  $e$  is error matrix, and target matrix  $R$  can be obtained by  $U$  and  $V$  approximation. As the same is true of the athlete network, when the evolution simulation is carried out, the  $e$  is taken as a disturbance term to continuously inject subtle

TABLE 1: Athlete social support attribute stake table.

Serial number	Actual support	Emotional support	Employment support	Achievement support	Income	Social interaction	Marriage support	Management support
1	72	-42	-29	52	1676	-97	31	57
2	70	2	-86	94	1078	-54	49	54
3	-69	54	27	-4	5013	85	-45	-12
4	-11	-59	-95	25	-6525	-26	-3	20
5	36	7	-15	33	3989	89	11	-64
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
91	-68	87	20	30	3236	-78	-23	-21

TABLE 2: Normalization of social support attributes of athletes.

Serial number	Actual support	Emotional support	Employment support	Achievement support	Income	Social interaction	Marriage support	Management support	Activity
1	1.25	-0.80	-0.37	0.68	0.29	-1.75	0.98	1.03	0.00
2	1.21	-0.05	-1.36	1.36	0.18	-1.00	1.56	0.98	0.22
3	-1.40	0.84	0.61	-0.21	0.89	1.42	-1.45	-0.26	0.24
4	-0.31	-1.09	-1.52	0.25	-1.21	-0.51	-0.11	0.34	-0.47
5	0.57	0.03	-0.12	0.38	0.71	1.49	0.34	-1.23	0.17
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
91	-1.38	1.40	0.49	0.33	0.57	-1.42	-0.75	-0.43	-0.21

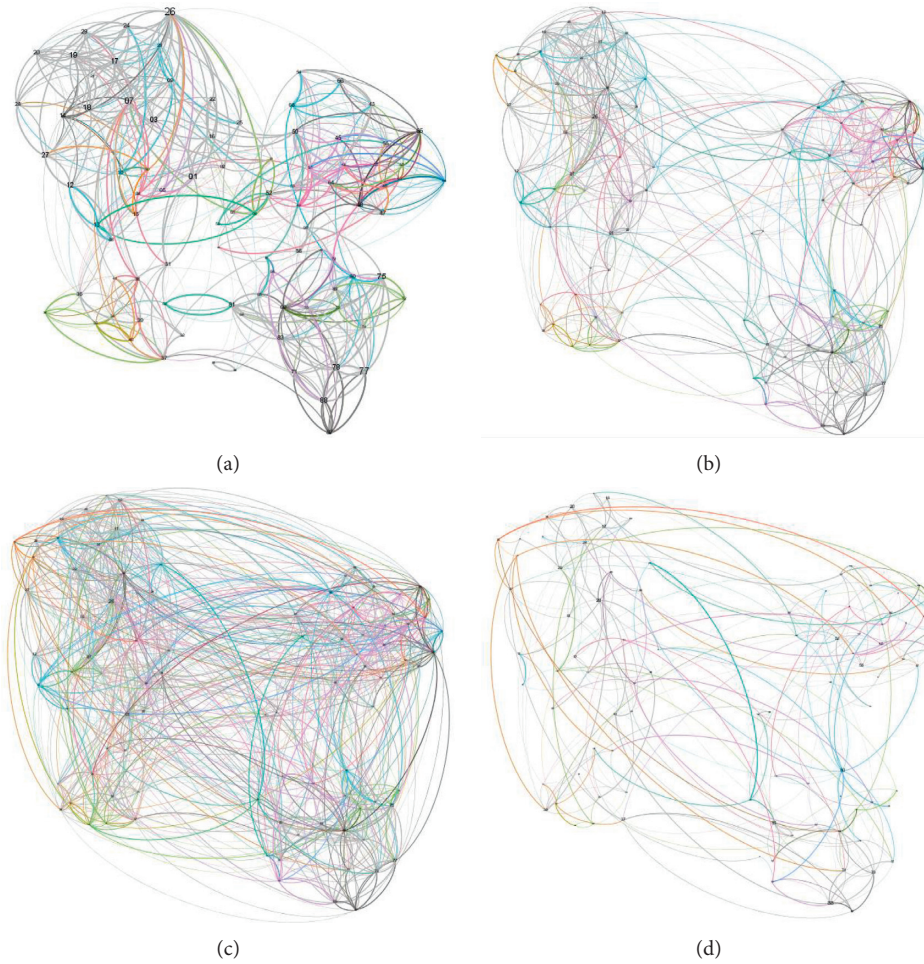


FIGURE 4: The evolution of the network structure of athletes. (a) The original. (b) After 100 iterations. (c) After 500 iterations. (d) After 1000 iterations.

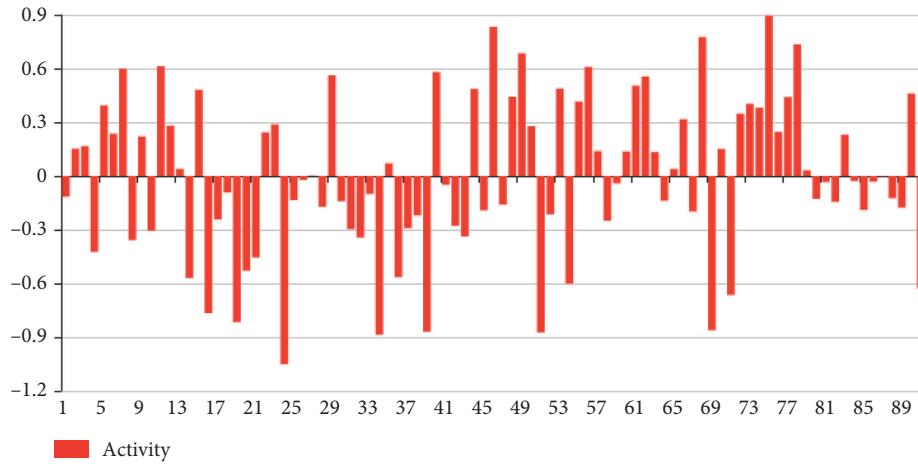


FIGURE 5: Node activity distribution.

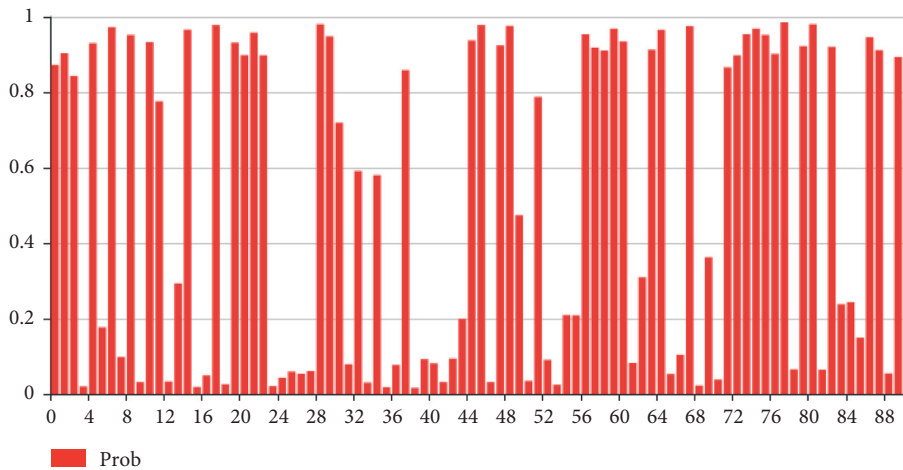


FIGURE 6: Probability distribution of node activity.

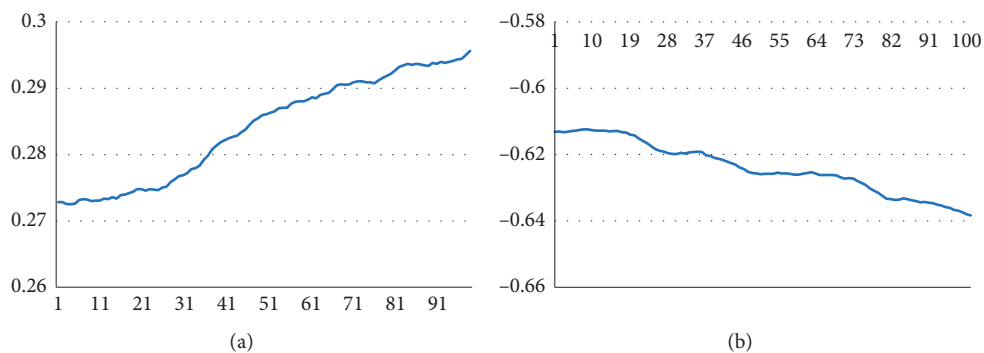


FIGURE 7: Time series of node activity evolution. (a) Active node ( $x$ ). (b) Passive node ( $y$ ).

disturbances into the network, which can stimulate the variation evolution of the original network.

The social support network of 91 athlete nodes is represented and predicted [26, 29–32], and the effect is shown in Figure 4. the different subnets of each athlete are represented by the edges of different colors, the network is simulated according to the evolution process of the network structure described in the article, and the subplot 1 shows the structure

of the original social support network of the athlete group. Subgraphs 1, 3, and 4 represent the overall structure of the network after 100, 500, and 1000 iterations according to the prediction algorithm, respectively.

By defining the activity of all nodes, the node set is divided into two subsets according to the positive and negative attributes of node activity, that is, passive node set and active node set. Passive nodes have low probability of



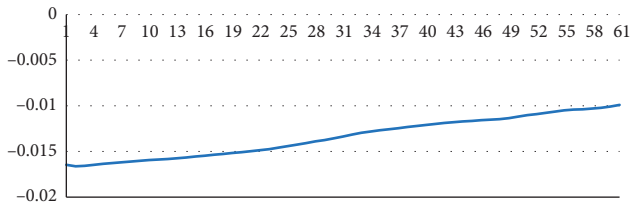


FIGURE 8: Time series of network system activity evolution.

walking and strong inertia, but play the role of resource adsorption, while active nodes play the nature of active nodes moving closer to passive nodes, as shown in Figure 5; this is the initial activity distribution for all nodes. Figure 6 shows the activity probability distribution calculated according to the activity of each node.

As shown in Figure 7, the left is the active sequence of an active node, whose activity is increasing under disturbance, and the right is the active sequence of a passive node, which is more and more inert under disturbance injection.

As shown in Figure 8, it is a network system activity sequence. In this example, although the network system activity is less than 0, the system has adsorption effect, but the trend is gradually approaching to 0, indicating that the whole network evolves towards stability.

## 5. Conclusion

A presentation model of complex network is put forward, which is expounded experimentally from the aspects of representation and structural evolution.

**5.1. Advantages and Disadvantages.** The model has the following advantages. (1) Incorporating node attributes into the modeling process and making node analysis more specific. (2) Explaining the connection mechanism between the principal nodes and giving the connection reasons. (3) The model is not only suitable for static network ingress but also more suitable for the prediction and evolution of network structure analysis. (4) Can guide the design questionnaire from the establishment of network links from two aspects of the design problem. (5) Suitable for the representation of multiple networks.

The model also has some shortcomings; the network represented by the model cannot be well supported by the current network mainstream visualization software, so according to the model, to draw a network map requires researchers to write their own programs.

**5.2. Future Research.** In the following research, the model will be improved through verification feedback in practice. The focus of the work will be on two aspects; on the one hand, it continues to improve the theory, combined with the existing node influence algorithm [33–37], node activity, and change probability index optimization; on the other hand, it hopes to write appropriate multinetwork visualization software to better model presentation.

## Data Availability

The data used to support the finding of 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 Shaanxi Provincial Social Science Fund Project (under Grant nos. 2018R13 and 2015P002) and Xi'an University 2020-2021 Education and Teaching Achievements Award Cultivation Project (JY2020CGPY01).

## References

- [1] Z. Li and J. Luo, "The social behavior and relational network characteristics of Chinese—viewpoint of a social network," *social science Front*, vol. 199, no. 1, pp. 159–164, 2012.
- [2] J. Luo and F. Zeng, "Study on organization based on complex system perspective," *Foreign Economics and Management*, vol. 41, no. 12, pp. 112–134, 2019.
- [3] J. Liu, *Practical Guide to Integrated Network Analysis UCINET Software*, GE Zhi Publishing House, Shanghai, China, 2014.
- [4] K. Xu, S. Zhang, H. Chen et al., "Measurement and analysis of online social network," *Chinese Journal of Computer*, vol. 37, no. 1, pp. 165–188, 2014.
- [5] X. Wang, "Data science and social networks: big data, small world," *Science and Society*, vol. 1, no. 1, pp. 27–35, 2014.
- [6] M. Jia, H. Xu, J. Wang, and Y. Bai, "Handling big data of online social networks on a small machine," *Computational Social Networks*, vol. 2, no. 1, 2015.
- [7] J. Ding, L. Xu, H. Qian et al., "A study on the multi-level relationship of the overall social network of Chinese professional athletes," *Journal of Xi'an Institute of Physical Education Nol*, vol. 4, pp. 405–409, 2016.
- [8] Y. Qian, *Online Social Networking Analysis*, Electronic Industry Press, Shenzhen, China, 2014.
- [9] L. Wang, Y. Tian, and J. M. Du, "Concept dynamics in social networks," *China Science: Information Science*, vol. 48, no. 1, pp. 3–23, 2018.
- [10] X. Xu, Q. Huang, and Z. Mao, "Research on social capital in the urban water ecological management network," *Research on the Modernization of National Governance*, vol. 1, pp. 250–268, 2019.
- [11] C. Hu, D. Yang, and L. Yang, "Government-social capital cooperation (PPP) Project stakeholder value network research," *Journal of Engineering Management*, vol. 2, no. 6, pp. 72–77, 2019.
- [12] F. J. Chen, Z. J. Wang, and F. Meng, "A study of interactive relationship between virtual social capital, internet word of mouth and purchase intention in social electronic commerce," *Journal of Xinjiang University of Finance and Economics*, vol. 68, no. 3, pp. 55–61, 2016.
- [13] B. Lai, *Research on the Effect of Social Capital and Electronic Word of Mouth on Brand Trust in Virtual Community*, South China University of Technology, Guangzhou, China, 2017.
- [14] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation

- for social recommendation,” *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4186–4199, 2020.
- [15] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, “Topological influence-aware recommendation on social networks,” *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [16] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, “Social recommendation with evolutionary opinion dynamics,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 3804–3816, 2020.
- [17] Y. Zhao and J. Luo, “How to measure social capital: a review of empirical research: how to measure social capital,” *foreign social sciences*, vol. 1, no. 2, pp. 18–24, 2005.
- [18] Z. Liu and W. Luo, “Research on dynamic cooperation evolution of platform enterprises and partners based on social capital theory—longitudinal case study of cainiao network,” *Business Economics and Management*, vol. 341, no. 3, pp. 15–27, 2020.
- [19] B. Lai, *A Study on the Influence of Social Capital and Network Word-Of-Mouth on Brand Trust in Virtual Community*, South China University of Technology, Guangdong, China, 2017.
- [20] X. Zhang and H. Zhang, “The application of the price key theory in the determining of the composition of the compound is used in the case of an example analysis,” *Mathematical Resolution Research*, vol. 440, no. 19, pp. 96–97, 2019.
- [21] H. Hou, Z. Huang, B. Wang, and B.-S. Wang, “Revisit on the teaching of valence bond theory in the curriculum of structural chemistry,” *University Chemistry*, vol. 31, no. 9, pp. 24–28, 2016.
- [22] W. Wu, L. Song, and Y. Mo, “The development of the theory of modern Valence Bond theory,” *Journal of Xiamen University (Natural Science Edition)*, vol. 40, no. 2, pp. 338–343, 2001.
- [23] M. Xie, *The Study of Node Attribute Inference Based on Network Representation Learning*, University of Electronic Science and Technology, Chengdu, China, 2019.
- [24] X. Zhang, H. Cheng, and Y. Fang, “Predict the co-relationship based on metapath and node attributes,” *Computer Engineering and Applications*, pp. 1–9, 2020.
- [25] N. Zhang, H. Liu, and L. Jiang, “V-UKIR propagation model research and simulation based on node attributes and information value,” *Computer Applications Research*, vol. 37, no. 10, pp. 1–6, 2019.
- [26] S. Hussain, L.-J. Muhammad, and A. Yakubu, “Mining social media and DBpedia data using gephi and R,” *Journal of Applied Computer Science & Mathematics*, vol. 12, no. 1, pp. 14–20, 2018.
- [27] Y. Guan, Y. Xiang, and C. Kang, “Research and application of visual analysis method based on gephi,” *Telecommunications Science*, vol. 29, no. S1, pp. 112–119, 2013.
- [28] Q. huang, B. Liu, N. Zhao et al., “Modeling for professional athletes’ social networks based on statistical machine learning,” *IEEE Access*, vol. 8, no. 1, pp. 4301–4310, 2020.
- [29] B. Yu and X. Hu, “Toward training and assessing reproducible data analysis in data science education,” *Data Intelligence*, vol. 1, no. 4, pp. 333–344, 2019.
- [30] Y. Wang, Y. Tian, T.-s. Li, D. Peng, and Y. Zhou, “Visual analysis of shipping recruitment information based on gephi,” *Big Data*, vol. 4, no. 3, pp. 81–91, 2018.
- [31] N. E. huang and F. Qiao, “A data driven time-dependent transmission rate for tracking an epidemic: a case study of 2019-nCoV,” *Science Bulletin*, vol. 65, no. 6, pp. 425–427, 2020.
- [32] Y. Pan, H. Yu, and Y. Wu, “Link prediction method based on complex network dynamics model,” *Journal of Network and Information Security*, vol. 5, no. 6, pp. 67–74, 2019.
- [33] W. Liu and W. Chen, “Link prediction in complex networks,” *Information and Control*, vol. 49, no. 1, pp. 1–23, 2020.
- [34] S. Guan, “Explosive synchronization on complex networks,” *Chinese Science: Physical Mechanics Astronomy*, vol. 50, no. 1, pp. 51–68, 2020.
- [35] Z. Ren, “Advances in the research of node influence in dynamic complex networks,” *Journal of Physics*, vol. 69, no. 4, pp. 24–32, 2020.
- [36] H. Wang, Z. Le, X. Gong et al., “A review of link prediction methods based on feature classification,” *Computer Science*, vol. 47, no. 8, pp. 302–312, 2020.
- [37] W. Zhao, F. Zhang, and J. Liu, “Advances in discovery of complex network communities,” *Computer Science*, vol. 47, no. 2, pp. 10–20, 2019.

## Research Article

# Characterization of 2-Path Signed Network

Deepa Sinha <sup>1</sup> and Deepakshi Sharma <sup>2</sup>

<sup>1</sup>Department of Mathematics, South Asian University, Akbar Bhawan Chanakyapuri, New Delhi-110021, India

<sup>2</sup>Department of Mathematics, Ramanujan College, University of Delhi, New Delhi-110021, India

Correspondence should be addressed to Deepa Sinha; [deepasinha2001@gmail.com](mailto:deepasinha2001@gmail.com)

Received 21 April 2020; Revised 15 August 2020; Accepted 1 October 2020; Published 31 October 2020

Academic Editor: Fei Xiong

Copyright © 2020 Deepa Sinha and Deepakshi Sharma. 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 *signed network* is a network where each edge receives a sign: positive or negative. In this paper, we report our investigation on *2-path signed network* of a given signed network  $\Sigma$ , which is defined as the signed network whose vertex set is that of  $\Sigma$  and two vertices in  $(\Sigma)_2$  are adjacent if there exist a path of length two between them in  $\Sigma$ . An edge  $ab$  in  $(\Sigma)_2$  receives a negative sign if all the paths of length two between them are negative, otherwise it receives a positive sign. A signed network is said to be *clusterable* if its vertex set can be partitioned into pairwise disjoint subsets, called *clusters*, such that every negative edge joins vertices in different clusters and every positive edge joins vertices in the same clusters. A signed network is *balanced* if it is clusterable with exactly two clusters. A signed network is *sign-regular* if the number of positive (negative) edges incident to each vertex is the same for all the vertices. We characterize the 2-path signed graphs as balanced, clusterable, and sign-regular along with their respective algorithms. The 2-path network along with these characterizations is used to develop a theoretic model for the study and control of interference of frequency in wireless communication networks.

## 1. Introduction

The *intersection graphs* or *networks* [1, 2] form a large family of structures which include many important network such as interval [3, 4], permutation [5, 6], chordal [7, 8], circular-arc [9, 10], circle [11], string [12], line [13, 14], and path [15, 16]. Most of these networks are of great significance not only theoretically but because of their applicability in the fields such as transportation [17], wireless networking [18], scheduling problem [19], molecular biology [20], circuit routing [21], and sociology. The *2-path network* is an intersection graph of *open neighborhoods* [22]. Formally, a 2-path network for a given network is obtained by joining the pair of vertices which form a path of length two in the original network.

*Signed networks* are the network where each edge receives a sign: positive or negative (ref. Figure 1, the edges 13, 32, 24, and 25 are negative in (b)). The concept stemmed from psychologist Heider [23, 24] who used the concept of balanced theory to model relations between individuals and

society using *triads*. Harary formulated and restructured the signed networks by introducing structural balance theory [25] for social balance and was used for portfolio management [26], where they used signed graphs to analyse the extent of hedging in a portfolio. These networks are widely used in data clustering [27–29]. Signed networks of some intersection networks have been already studied [30–33]. Also path graphs of signed graphs are discussed in [34–38]. In this paper, we try and establish the results for signed networks defined on 2-path networks.

The study of networks has come a long way with its applicability in many fields [39–43]. One such application is in wireless networks and frequency allocation problem. The frequency allocation [44] in a wireless network [45] or radio frequency allocation [46] is one of the typical problems that we still face. The interference takes place in a network when the transmission from one station interacts with the transmission from another station. There are three dimensions to the frequency spectrum first being the space in which the emission is radiated, second is the frequency

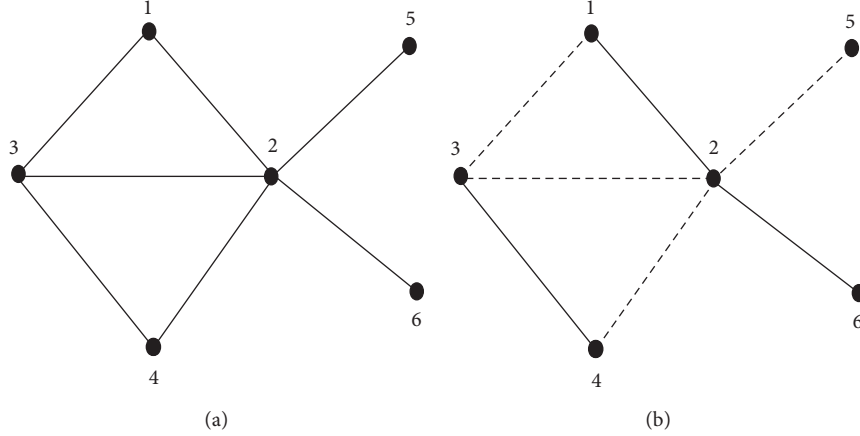


FIGURE 1: (a) Given network. (b) Its signed network.

bandwidth, and third is the time. If these three dimensions occur simultaneously for a channel receiving transmission, then interference takes place. Many methods and models such as dipole-moment models [47], Kurtosis detection algorithm and its versions [48], and decomposition method based on reciprocity [49] have been studied for this problem. We, on the other hand, bring a model with a very different approach using the theoretic aspects of signed networks and 2-path networks.

We aim at detecting and reducing the interference by assuming the stations/channels as vertices and transmission between them as edges. Furthermore, we assume that two channels (vertices) are joined by a positive edge if and only if they are in different time or frequency bandwidth (assuming that space is always same) and by a negative edge if both are the same in time and frequency bandwidth. If  $\Sigma$  is the given signed network representing a channel network, then its 2-path signed network  $(\Sigma)_2$  provides with the interference pattern of different channels at a specific channel. In the 2-path signed network the edge  $uw$  is given a negative sign if and only if there exist all-negative two paths in  $\Sigma$  say  $uv$  and  $vw$  between them, that is, they are the same in time as well as frequency and thus the interference takes place at  $v$  because of the vertices  $u$  and  $w$ .

*1.1. Definitions and Preliminaries.* For standard terminology and notation in network theory, one can refer Harary [50] and West [51], and for signed networks literature, one can refer [52]. Throughout the text, we consider finite, undirected network with no loops or multiple edges. A *signed network* is an ordered pair  $\Sigma = (\Sigma^u, \sigma)$ , where  $\Sigma^u$  is a network  $G = (V, E)$ , called the underlying network of  $\Sigma$ , and  $\sigma: E \rightarrow \{+, -\}$  is a function from the edge set  $E$  of  $\Sigma^u$  into the set  $\{+, -\}$ , called the *signature* (or *sign* in short) of  $\Sigma$ . A signed network is *all-positive* (or *all-negative*) if all its edges are positive (negative). Furthermore, it is said to be *homogeneous* if it is either all-positive or all-negative and *heterogeneous* otherwise. A signed network  $\Sigma$  is said to be *sign-regular* if the number of positive edges (negative edges),  $d^+(v)$  ( $d^-(v)$ ) incident at a vertex  $v$  in  $\Sigma$ , is independent of

the choice of  $v$  in  $\Sigma$ , that is,  $\Sigma$  is  $(i, j)$ -sign-regular, where  $i = d^+(v)$  is the positive degree of every vertex  $v$  in  $\Sigma$  and  $j = d^-(v)$  is the negative degree of every vertex  $v$  in  $\Sigma$ .

A *negative section* of a cycle or a path  $Z$  is a maximal set  $D$  of vertices of  $Z$  such that the subsigned network consisting of the edges of  $Z$  joining vertices in  $D$  is all-negative and connected. A *marked signed network* is an ordered pair  $\Sigma_\mu = (\Sigma, \mu)$ , where  $\Sigma = (\Sigma^u, \sigma)$  is a signed network and  $\mu: V(\Sigma^u) \rightarrow \{+, -\}$  is a function from the vertex set  $V(\Sigma^u)$  of  $\Sigma^u$  into the set  $\{+, -\}$ , called a *marking* of  $\Sigma$ . We define  $(V(\Sigma))_\mu = \{v_i^+, v_i^-, \forall v_i \in V(\Sigma)\}$ . Let  $v$  be an arbitrary vertex of a network  $\Sigma$ . We denote the set consisting of all the vertices of  $\Sigma$  adjacent with  $v$  by  $N(v)$ . This set is called the *neighborhood set* of  $v$  and sometimes we call it as *neighborhood* of  $v$ . Next, we define two marked neighborhoods as  $N_*(t) = \{v^\mu \in (V(\Sigma))_\mu: tv \text{ is an edge with sign } \mu\}$  and  $N_*^-(t) = \{v^- \in (V(\Sigma))_\mu: tv \text{ is an edge}\}$ . For each  $N(t)$  of  $\Sigma^u$ , there exist  $N_*(t)$  in  $\Sigma$  and vice versa. It is important to note that the vertex marking for each vertex is neighborhood dependent, i.e., if a vertex  $v_i$  forms a negative edge with  $v_j$  and a positive edge with  $v_k$  in  $\Sigma$ , then  $v_i^-$  belongs to the marked neighborhood of  $v_j$ ,  $N_*(v_j)$  and  $v_i^+ \in N_*(v_k)$ . Also, each vertex in  $N(t)$  appears with exactly one mark in  $N_*(t)$ .

The *clique* of a network  $\Sigma^u$  is a subset of vertices such that every two vertices in the subset are connected by an edge.  $\delta(N(t))$  is a clique generated by vertices in  $N(t)$ . A *cycle* in a signed network  $\Sigma$  is said to be *positive* if the product of the signs of its edges is positive or, equivalently, if the number of negative edges in it is even. A cycle which is not positive is said to be *negative*. A signed network is *balanced* if all its cycles are positive. A signed network is said to be *clusterable* if its vertex set can be partitioned into pairwise disjoint subsets, called *clusters*, such that every negative edge joins vertices in different clusters and every positive edge joins vertices in the same clusters. A *common neighbor* of vertices  $v_i$  and  $v_j$  is a vertex  $v$  such that  $v \in N(v_i) \cap N(v_j)$ .

*Property 1.* If  $v_i, v_j, v_k \in V(\Sigma)$ , then by property  $P(v_i, v_j; v_k)$ , we mean  $\{v_i^-, v_j^-\} \subseteq N_*(v_k)$ .

*Property 2.* By property  $\mathbf{P} := P(v_i, v_j)$ , we mean that if property  $P(v_i, v_j; v)$  holds for one common neighbor  $v$  of  $v_i$  and  $v_j$ , then it holds for every common neighbor  $v$  of  $v_i$  and  $v_j$ .

The *2-path signed network* [34]  $(\Sigma)_2 = (V, E', \sigma')$  of a signed network  $\Sigma = (V, E, \sigma)$  is defined as follows. The vertex set is the same as the original signed network  $\Sigma$ , and two vertices  $u, v \in V((\Sigma)_2)$  are adjacent if and only if there exist a path of length two in  $\Sigma$ . The edge  $uv \in V((\Sigma)_2)$  is negative if and only if all the edges in all the two paths in  $\Sigma$  between them are negative otherwise the edge is positive. The definition of 2-path signed network was used [53], to bring out some basic results, which was further extended and shaped in the present paper. The 2 subsets  $\{v_i, v_j\}$  having property  $\mathbf{P}$  are named as  $\mathbf{P}$  pairs and the set of all  $\mathbf{P}$  pairs is denoted by  $\mathbf{P}_*$ . A negative section  $v_1, v_2, \dots, v_k$  in cycle or path is said to be  $\mathbf{P}$  section if every alternate vertices forms  $\mathbf{P}$  pair. If  $v_k = v_1$  in the  $\mathbf{P}$  section, then such a  $\mathbf{P}$  section is called  $\mathbf{P}$  cycle.

*Remark 1.* An edge  $v_i v_j$  is negative in 2-path signed network  $(\Sigma)_2$  if and only if  $\{v_i, v_j\}$  a neighborhood of a vertex  $v_k$  has property  $\mathbf{P}$ .

Consider the signed network (a) of Figure 2; clearly, by definition of 2-path signed network, 14 and 23 are the only negative edges as there exist an all-negative path between 1 and 4 and 2 and 3 of length two in  $\Sigma$ . Let us consider the marked neighborhood for each of the vertices 1 to 4:

$$\begin{aligned} N_*(1) &= \{2^-\}, \\ N_*(2) &= \{1^-, 3^+, 4^-\}, \\ N_*(3) &= \{2^+, 4^-\}, \\ N_*(4) &= \{2^-, 3^-\}. \end{aligned} \quad (1)$$

One can note that there are following pairs in each neighborhood:  $\{1^-, 3^+\}$ ,  $\{3^+, 4^-\}$ , and  $\{1^-, 4^-\}$  in  $N_*(2)$ ,  $\{2^+, 4^-\}$  in  $N_*(3)$ , and  $\{2^-, 3^-\}$  in  $N_*(4)$ . These pairs give rise to the edges in  $(\Sigma)_2$ , where the edge 24, 34, and 13 get positive signs as atleast one of the vertices in each set is positively marked (thus does not form a  $\mathbf{P}$  pair).

The pairs  $\{1^-, 4^-\}$  and  $\{2^-, 3^-\}$  are the only  $\mathbf{P}$  pairs, then by Remark 2.1 these pairs give rise to the negative edges 14 and 23.

*1.2. Model and Assumptions.* Consider the vertices in a network system as channels or stations and edges as the signals or transmissions. Now, the problem of interference between stations  $a, b, c$  arise when the transmissions from  $a$  and  $c$  reaching at  $b$  are in the same space, at the same frequency and time. We assume that the two parameters space and frequency are the same for all transmissions but the time is the variation parameter. So, a positive edge between two channels represents that they always have a fixed time of transmission (which is different from the adjoining transmission) and a negative edge means a variable transmission (which is different from the fixed time transmission). The problem arises when two of the negative

edges are incident to a station as here there is a high risk of interference in the network. This network pattern can be observed using 2-path networks. The 2-path network  $(\Sigma)_2$  of the network  $\Sigma$  gives the possible interference pattern so a negative edge  $ab$  in the 2-path signed network means that there exist a vertex  $c$  such that  $ac$  and  $cb$  are the transmissions which could possibly interfere with each other at  $c$ . Thus, by analyzing 2-path networks, we can study and correct this problem of interference.

We thus study the theoretic Properties 2-path networks keeping our problem of interference in the background. In Section 1, we characterize the 2-path signed network, some of the results of which were presented in [35], give algorithms to construct 2-path signed network, detect if a network is isomorphic to a 2-path signed network, and collect  $\mathbf{P}$  pairs. Section 2 is on the balancedness of 2-path signed networks where we characterize signed network whose 2-path signed networks are balanced. The property of balanced network is used to identify and categorize all the channels in 2-path in exactly two groups such that the negative edges are across the groups and positive are in the same class. We also provide an algorithm for the same. Section 3 is dedicated to another property of signed network known as clusterability, where we check whether we can categorize our vertices in more than two clusters so that the negative edges are across the group. We follow the property of sign-regularity in Section 4.

## 2. 2-Path Signed Network

*2.1. Characterization of 2-Path Signed Network.* In this section, we give a characterization of 2-path signed network. We check if a given signed network is a 2-path signed network of some signed network. We then find its underlying signed network. Following characterization of 2-path networks was given by Acharya and Vartak.

**Lemma 1** (see [22]). *A connected network  $G$  with vertices  $v_i, i = 1, \dots, n$ , is a 2-path network of some network  $H$  if and only if  $G$  contains a collection of complete subgraphs  $G_1, G_2, \dots, G_n$  such that, for each  $i, j = 1, \dots, n$ , the following hold: (i)  $v_i \notin G_j$ , (ii)  $v_i \in G_j \iff v_j \in G_i$ , and (iii) if  $v_i v_j \in G$  then there exists  $G_k$  containing  $v_i v_j$ .*

Let  $v_i v_j$  be an edge in a signed network  $\Sigma$ . Let  $\mu_{ij}$  denote the marking on the vertex  $v_j$  in  $N_*(v_i)$  such that  $\mu_{ij} = \sigma(v_i v_j)$ ; then, we have the following theorem.

**Theorem 1.** *A connected signed network  $\Sigma$  with vertices  $v_1, v_2, \dots, v_n$  is a 2-path signed network of some signed network  $\Sigma'$  if and only if  $\Sigma$  contains a collection of complete signed subgraphs  $\Sigma_1, \Sigma_2, \dots, \Sigma_n$  with vertices  $v_i$  in  $\Sigma_j, j \neq i$  having the marking  $\mu_{ji}$  such that, for each  $i, j = 1, \dots, n$ , the following hold:*

- (i)  $v_i \notin \Sigma_i$
- (ii)  $v_i^{\mu_{ji}} \in \Sigma_j \iff v_j^{\mu_{ji}} \in \Sigma_i$ , where  $i \neq j$  and  $\mu_{ji} = \mu_{ij}$
- (iii) If  $v_i v_j \in E(\Sigma)$ , then there exists  $\Sigma_k$  containing  $v_i^{\mu_{ki}}, v_j^{\mu_{kj}}$  and if  $\sigma(v_i v_j) = -\iff \{v_i, v_j\}$  is a  $\mathbf{P}$  pair in  $\Sigma_k$

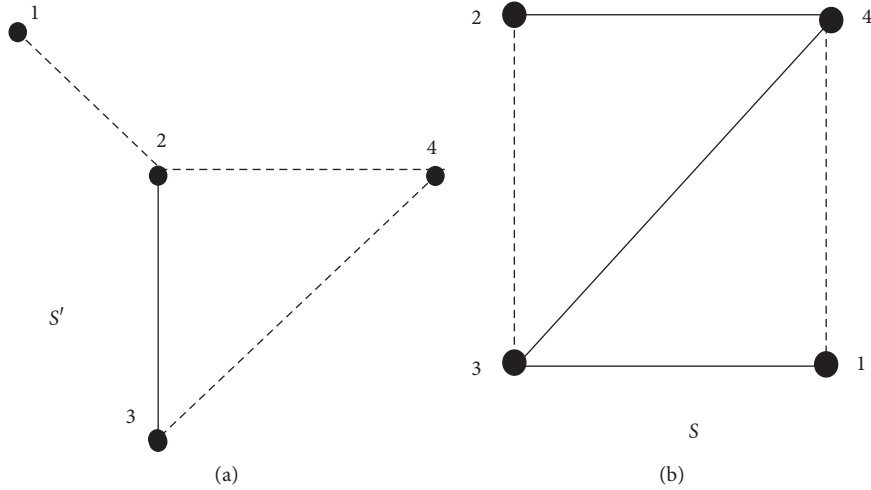


FIGURE 2: (a) Given signed network. (b) Its 2-path signed network.

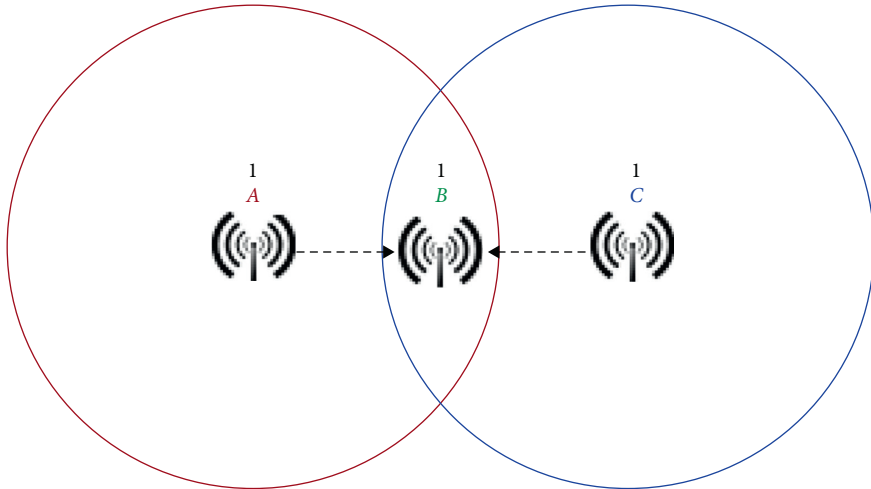


FIGURE 3: Interference at B due to A and C.

*Proof. Necessity.* Suppose  $\Sigma$  is a 2-path signed network of signed network  $\Sigma'$  with vertices  $v_1, v_2, \dots, v_n$ . To show that there exist a collection of  $n$  complete signed subgraphs  $\Sigma_1, \Sigma_2, \dots, \Sigma_n$  such that (i), (ii), and (iii) hold, consider the marked neighborhood  $N_*(v_i)$  so that  $\Sigma_i = \delta(N_*(v_i))$  for each vertex  $v_1, v_2, \dots, v_n$  in  $\Sigma'$ . Each  $N_*(v_i)$ ,  $i = 1$  to  $n$ , generates a complete signed subgraph in 2-path signed network  $\Sigma$  of  $\Sigma'$ , since  $N_*(v_i)$  consists of elements which have a common neighbor  $v_i$  in  $\Sigma'$ . Next,  $v_i \notin N_*(v_i)$  since the signed network  $\Sigma'$  is simple and  $v_i \notin N(v_i)$ . Let  $v_i^{\mu_{ij}} \in N_*(v_j)$  for some  $i, j$ ;  $i \neq j$ ,  $1 \leq i, j \leq n$ . Then,  $v_i v_j$  is an edge in  $\Sigma'$  such that  $\sigma(v_i v_j) = \mu_{ji}$ , and thus  $v_i^{\mu_{ij}}$  is in the marked neighborhood  $N_*(v_i)$ . Therefore, for  $i \neq j$ ,  $v_i^{\mu_{ki}} \in \Sigma_j \Leftrightarrow v_j^{\mu_{kj}} \in \Sigma_i$ . Next, since the mark  $\mu_{ij}$  depends on the sign of edge  $v_i v_j$  in  $\Sigma'$ , the vertex  $v_j \in N_*(v_i)$  gets the marking from the corresponding sign  $\sigma(v_i v_j) = \mu_{ij}$ ; hence,  $\mu_{ji} = \mu_{ij}$  and (ii) follows.

If  $v_i v_j$  is an edge in 2-path signed network  $\Sigma$  and thus in  $\Sigma^u$ , then from Lemma 1,  $v_i, v_j \in N(v_k)$  for some  $k$ . Let  $v_i v_j$  be an edge in  $\Sigma$  with a sign “-,” then  $v_i v_k$  and  $v_k v_j$  are negative

edges in  $\Sigma'$ , by definition of 2-path signed network. Thus, by Remark 1,  $\{v_i^-, v_j^-\} \subseteq N_*(v_k)$  and  $\{v_i, v_j\}$  is **P** pair. Hence, (iii) holds.

*Sufficiency.* Let  $\Sigma$  contain a collection of complete signed subgraphs  $\Sigma_1, \Sigma_2, \dots, \Sigma_n$  with vertices  $v_i$  in  $\Sigma_j$ ,  $j \neq i$  with the marking  $\mu_{ji}$ , satisfying the three properties (i), (ii), and (iii). Join each vertex  $v_i$  to the vertices of  $\Sigma_i$ ; let the obtained signed network be  $\Sigma'$ . Let  $v_j$  be in  $\Sigma_i$ , having the marking  $\mu_{ij}$ . Then, the sign of edge  $v_i v_j$  in  $\Sigma'$  is given by  $\mu_{ij}$ . From our construction of  $\Sigma'$  and property (ii), as  $v_j^{\mu_{ij}} \in \Sigma_i$ , we have  $v_i^{\mu_{ji}} \in \Sigma_j$  and  $\mu_{ij} = \mu_{ji}$ . Note that if  $v_j v_k \in E(\Sigma)$ , then by (iii), there exist  $\Sigma_i$  such that  $v_j^{\mu_{ij}} \in \Sigma_i$  and  $v_k^{\mu_{ik}} \in \Sigma_i$ . Also if sign of the edge  $v_j v_k$  is negative in  $\Sigma'$ , then  $\mu_{ij} = \mu_{ik} = -$  and  $\{v_j v_k\}$  is a **P** pair in  $\Sigma$ , whereas if the edge is positive in  $\Sigma'$ , then the marks  $\mu_{ij}$  and  $\mu_{ik}$  are according to the edge sign  $v_i v_j$  and  $v_i v_k$  in  $\Sigma$ , respectively.

Next, we will show that  $(\Sigma')_2 \cong \Sigma$ . Clearly, the vertex set of  $\Sigma$  and  $(\Sigma')_2$  is the same. We are left to show that their adjacencies along with their signs are preserved. For some

$k \neq i$  and  $k \neq j$ , let  $v_i v_j$  be an arbitrary edge in  $\Sigma$ , such that  $v_i, v_j$  is in  $\Sigma_k$ . Clearly,  $v_i v_k$  and  $v_k v_j$  are edges in  $\Sigma'$ . Thus, there is a path of length two between  $v_i$  and  $v_j$  in  $\Sigma'$ . Hence,  $v_i v_j$  is an edge in  $(\Sigma')_2$ .

Furthermore, let  $v_l v_m$  be an arbitrary edge in  $(\Sigma')_2$ . Clearly, there must be atleast one path of length two  $v_l v_k v_m$  in  $\Sigma'$ . As  $\Sigma'$  is constructed joining a vertex  $v_k$  to all those vertices present in  $\Sigma_k$ , therefore by (iii),  $v_l$  and  $v_m$  is in  $\Sigma_k$  (as both these vertices are adjacent to  $v_k$  in  $\Sigma'$ ). Hence,  $v_l v_m$  is an edge in one of the cliques and thus it is an edge in  $\Sigma$ . Hence, the adjacencies are preserved.

Finally, we show that the sign of  $v_i v_j$  is the same in  $(\Sigma')_2$  and  $\Sigma$ . Let  $v_i v_j$  be a negative edge in  $\Sigma$ ; then, by condition (iii),  $\{v_i, v_j\}$  is a **P** pair in  $\Sigma_k$ , for some  $k, k \neq i, j$ . By Property 2 and signing of the edges in  $\Sigma'$ , the edges formed by vertices incident to both  $v_i$  and  $v_j$  in  $\Sigma'$  will all be negative; thus, the edge  $v_i v_j$  in 2-path signed network  $(\Sigma')_2$  is negative. Similarly, if  $\sigma(v_i v_j) = -$  in  $(\Sigma')_2$ , then all the  $v_i v_j$ -paths of length two will be all-negative in  $\Sigma'$ . Let  $v_i v_k v_j$  be a 2-path in  $\Sigma'$ . As the edges incident to  $v_k$  in  $\Sigma'$  are formed by joining a vertex  $v_k$  with vertices in  $\Sigma_k$  and  $\sigma(v_k v_i) = \mu_{ki}$  and  $\sigma(v_k v_j) = \mu_{kj}$ , where  $v_i^{\mu_{ki}}, v_j^{\mu_{kj}} \in \Sigma_k$  and  $\mu_{ki} = \mu_{kj} = -$ , then, since all edges in 2-path between vertices  $v_i$  and  $v_j$  in  $\Sigma'$  are negative, then by (iii)  $\sigma(v_i v_j) = -$  in  $\Sigma$ . Similarly, the same holds for positive edge. If  $v_i v_j$  is a positive edge in  $(\Sigma')_2$ , then there exist atleast one vertex  $v_k$  in  $\Sigma'$ , where  $v_i v_k v_j$  is a path of length two. Then, clearly, at least one of the two vertices  $v_i$  and  $v_j$  forms a positive edge with  $v_k$  in  $\Sigma'$ , also in the complete sub signed network  $\Sigma_k$  containing both  $v_i$  and  $v_j$ , and thus  $v_i v_j \in E(\Sigma)$ , and by (iii) it cannot be a negative edge. Hence,  $\sigma(v_i v_j) = +$  in  $\Sigma$ . Conversely, if  $\sigma(v_i v_j) = +$  in  $\Sigma$ , then as done above for negative edge for some  $\Sigma_k$ , the mark of one of these two vertices would be positive, and thus there would be a heterogeneous path of length two between  $v_i$  and  $v_j$  in  $\Sigma'$  and thus a positive edge in  $(\Sigma')_2$ . Hence,  $(\Sigma')_2 \cong \Sigma$ .

Before going further, we prove the following lemmas.  $\square$

**Lemma 2.** *If a signed network  $\Sigma$  has an induced cycle of length  $2k, k \geq 3$ , then  $(\Sigma)_2$  contains two vertex disjoint cycles of length  $k$  each.*

*Proof.* Let  $v_1 t_1 v_2, \dots, v_k t_k v_1$  be an induced cycle of length  $2k, k \geq 3$ , in a signed network  $\Sigma$ . Then, there is a path of length two between each pair  $v_i$  and  $v_{i+1}$  for  $i = 1, \dots, k-1$  and  $v_k, v_1$ . Thus,  $v_1 v_2, \dots, v_k v_1$  is a cycle of length  $k$  in  $(\Sigma)_2$ . Similarly,  $t_1 t_2, \dots, t_k t_1$  is a cycle of length  $k$  in  $(\Sigma)_2$ . Clearly, there is no vertex in common in the two cycles. Hence, the result follows.  $\square$

*Remark 2.* If the signed network  $\Sigma$  consists of a cycle of length 4, then  $(\Sigma)_2$  contains two disjoint edges corresponding to the vertices of the cycle.

**Lemma 3.** *For a signed network  $\Sigma$ ,  $(\Sigma)_2$  contains an induced cycle of length  $k$  ( $k$  is odd and  $k > 3$ ) if either  $\Sigma$  has a cycle of length  $k$  or  $\Sigma$  has a cycle of length  $2k$ .*

*Proof.* First, assume that  $\Sigma$  has an induced cycle  $C: v_1 v_2, \dots, v_k v_1$  of length  $k$  ( $k$  is odd and  $k > 3$ ). Clearly,  $C': v_1 v_3, \dots, v_k v_2, \dots, v_{k-1} v_1$  is a cycle of length  $k$  in  $(\Sigma)_2$ .

Next, if  $\Sigma$  has an induced cycle  $v_1 t_1, \dots, v_k t_k$  of length  $2k$ , where  $v_i, t_i \in V(\Sigma)$  for  $i = 1, 2, \dots, k$ , clearly,  $v_1 v_2, \dots, v_k v_1$  and  $t_1 t_2, \dots, t_k t_1$  are two cycles of length  $k$  in  $(\Sigma)_2$ . Hence, result follows.  $\square$

**2.2. Algorithm to Detect **P** Pairs.** Here, in Algorithm 1, we present algorithm for detection of **P** pairs. For an input network  $\Sigma$ , these **P** pairs gives the negative edges in  $(\Sigma)_2$  and thus will be used as an input for all the other future algorithms.

**2.2.1. Complexity.** In Steps 5, 6, and 7, we use three different loops running upto  $n$  times (number of vertices). Thus, the complexity of this fragment is equal to  $O(n^3)$ .

Next, we use three nested loops in Steps 12, 13, and 15. Again the complexity is  $O(n^3)$ .

The loops at Step 19 and Step 20 are used for the array  $P$  and  $Q$  with a size of  $m = l - 2$  and  $g = t - 2$ , respectively. Since the maximum number of edges in a network is  $n(n-1)/2$  and  $P$  and  $Q$  collects vertices incident to edges, thus  $m \leq n(n-1)$  and  $g \leq n(n-1)$ . Therefore, the complexity of this fragment is  $O(n(n-1)) \leq O(n^2)$ .

Combining the above, the total complexity equals  $O(n^3) + O(n^3) + O(n^2)$ . Hence, the complexity involved in Algorithm 1 is  $O(n^3)$ .

### 2.2.2. Implementation of Algorithm

*Example 1.* In this example, we are given with a signed network, as shown in Figure 4(a), and we find all the **P** pairs for the given signed network  $\Sigma$ . Let us consider adjacency matrix for the signed network which is given as follows:

$$A = \begin{bmatrix} 0 & -1 & -1 & 0 \\ -1 & 0 & -1 & 1 \\ -1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}. \quad (2)$$

In order to detect the **P** pairs of  $\Sigma$ , we use Algorithm 1 where the inputs are the matrix  $A$  and size  $n = 4$ . The size of array  $P$  and  $Q$  is initially fixed as  $n^2$ . After initializing the array  $P$  and  $Q$  as zero,  $t = 1$  and  $l = 1$ , we enter the first, second, and third loop which run for 1 to 4. The three loops at Steps 5, 6, and 7 are used to obtain two distinct elements  $A(i, j)$  and  $A(j, k)$ . Next, in Step 9 we check pair of vertices  $i, j$  and  $j, k$  if  $A(i, j) = 1$  and  $A(j, k) = 1$  or  $A(i, j) = -1$  and  $A(j, k) = 1$  or  $A(i, j) = 1$  and  $A(j, k) = -1$ , then  $i$  and  $k$  enter as an consecutive entry in the array  $Q$ . The array  $Q$  collects the vertices forming positive 2-paths in  $\Sigma$ . For the above matrix, as we enter the loop, then for  $i = 1, j = 2, k = 4$ , we find that  $A(1, 2) = -1$  and  $A(2, 4) = 1$ . Thus, as  $t = 1, Q[1] = 1$  and  $Q[2] = 4$  and  $t$  is incremented to  $t = 3$ . Similarly, after completing all the loops we get  $Q$  as follows.

**Input:** Adjacency matrix  $A$  and dimension  $n$ .  
**Output:** Array  $P$  which has all collection of  $\mathbf{P}$  pair and array  $Q$  which has pair of vertices which are not  $\mathbf{P}$  pairs.  
**Process:**

- (1) Enter the order  $n$  and adjacency matrix  $A$  for the signed network  $\Sigma$
- (2) **for**  $i = 1$  to  $n^2$  **do**
- (3)  $P[i] = 0; Q[i] = 0,$
- (4)  $t = 1; l = 1;$
- (5) **for**  $i = 1$  to  $n$  **do**
- (6) **for**  $j = 1$  to  $n$  **do**
- (7) **for**  $k = 1$  to  $n$  **do**
- (8) **if**  $(i \neq j \neq k)$  **then**
- (9) **if**  $((A(i, j) = 1) \&\& (A(k, j) = -1)) \vee ((A(i, j) = 1) \&\& (A(k, j) = 1)) \vee ((A(i, j) = -1) \&\& (A(k, j) = 1))$  **then**
- (10)  $Q[t] = i; Q[t + 1] = k;$
- (11)  $t = t + 2;$
- (12) **for**  $i = 1$  to  $n$  **do**
- (13) **for**  $j = 1$  to  $n$  **do**
- (14) **if**  $(A(i, j) = -1)$  **then**
- (15) **for**  $k = 1$  to  $n$  **do**
- (16) **if**  $(A(k, j) = -1) \&\& (i \neq j)$  **then**
- (17)  $P[l] = i; P[l + 1] = k;$
- (18)  $l = l + 2;$
- (19) **for**  $i = 1$  to  $t - 2, i = i + 2$  **do**
- (20) **for**  $j = 1$  to  $l - 2, j = j + 2$  **do**
- (21) **if**  $((P[i] = Q[j]) \&\& (P[i + 1] = Q[j + 1]))$  **then**
- (22)  $P[i] = 0;$
- (23)  $P[i + 1] = 0;$

ALGORITHM 1: An algorithm to detect and collect  $\mathbf{P}$  pairs.

$Q$	1	4	2	4	2	3	3	4	3	2	4	3	4	2
-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---

One can note that the number of elements in array  $Q$  is given by the  $t - 2$  (since every time we enter Step 9,  $t$  is incremented by 2). In the given example,  $t = 16$ , and thus size of array  $Q$  is 14. Next, we enter loop at Step 13, to find all the edges  $ij$  and  $jk$  such that  $A(i, j) = -1$  and  $A(j, k) = -1$ . Once such pair is obtained then  $i$  and  $k$  are collected in Array  $P$  as done for array  $Q$ . For the given example, array  $P$  comes out to be

$P$	1	3	1	2	2	3	2	3
-----	---	---	---	---	---	---	---	---

The size of  $P$  is given by  $l - 2$  and in our example  $l = 10$ . Finally, we compare the elements of  $P$  and  $Q$  pairwise to find if there are pairs of vertices which is common in both and then we remove these pairs from  $P$ . This is done by using two loops one moves from 1 to  $t - 2$  (the size of array  $Q$ ) and the other moves from 1 to  $l - 2$  (size of array  $P$ ). For  $i = 5$  and  $j = 5$ , we see that  $P[5] = Q[5]$  and  $P[6] = Q[6]$ . Therefore,  $P[5] = 0$  and  $P[6] = 0$ . Proceeding in the same way, we obtain  $P$  for the given example as:

$P$	1	3	1	2	0	0	0	0
-----	---	---	---	---	---	---	---	---

Thus, the  $\mathbf{P}$  pair are 1, 3 and 1, 2. From Figure 4 and Remark 1, we can verify that the  $\mathbf{P}$  pair obtained in the algorithm is the same as that for the given signed network.

*Remark 3.* Each pair  $Q[i], Q[i + 1]$  in array  $Q$ , for  $i = 1$  to  $l - 1$ , is a pair of vertices in  $\Sigma$ , which have a path of length

two between them in  $\Sigma$ . Since they do not form a  $\mathbf{P}$  pair, they form a positive edge in  $(\Sigma)_2$  (by Remark 1).

*2.3. Algorithm to Construct 2-Path Signed Network.* Next in Algorithm 2, we obtain the 2-path signed network  $(\Sigma)_2$  for a given signed network  $\Sigma$ . We use the adjacency matrix  $A$  of  $\Sigma$ . We use Algorithm 1, to obtain the vertices forming negative edges (collected in array  $P$ ) and positive edges (collected in array  $Q$ ) of 2-path signed network. The adjacency matrix of 2-path signed network obtained is saved in matrix  $B$ .

*2.3.1. Complexity.* In Step 2, we use two loops to initialize the matrix  $B$  zero, and thus the complexity of these steps is  $O(n^2)$ .

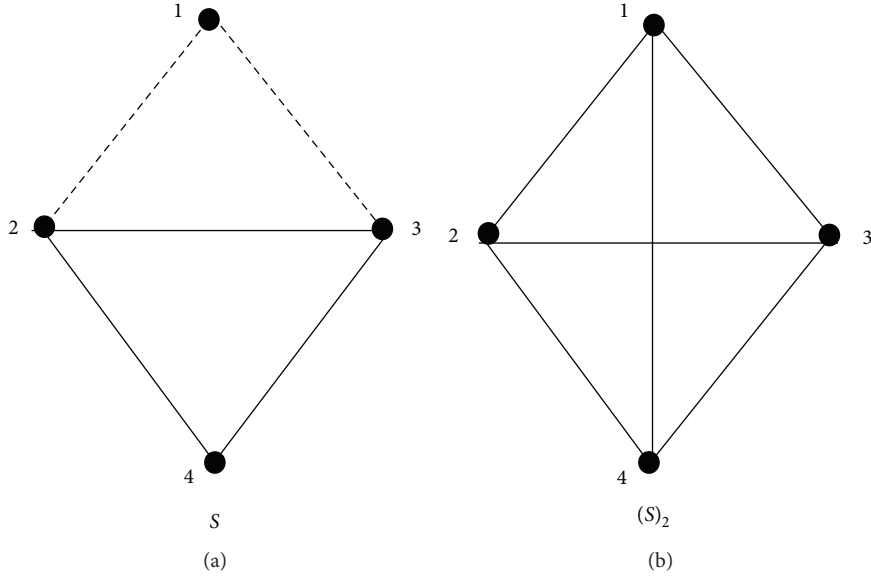
In Step 5, we use a loop which runs up to  $m - 1$ ,  $m \leq n(n - 1)$  (from previous algorithm). Thus, the complexity of this step is  $O(n(n - 1)) = O(n^2)$ . Similarly, in Step 10, we use a loop which runs up to  $g - 1$ ,  $g \leq n(n - 1)$ . Again complexity of this step is  $O(n(n - 1)) = O(n^2)$ .

The total complexity is  $O(n^2) + O(n^2) + O(n^2) = O(n^2)$ .

*2.3.2. Implementation of the Algorithm*

*Example 2.* We are given with a signed network  $\Sigma$ , as shown in Figure 4(a), with the adjacency matrix  $A$  as in Example 1. We have to find the adjacency matrix  $B$  of the 2-path signed network  $(\Sigma)_2$ . In Steps 2 and 3, we use two loops, each running from 1 to 4 as  $n = 4$  and assign zero to all the entries



FIGURE 4: (a) Given signed network  $\Sigma$ . (b) Its 2-path signed network.

**Input:** Adjacency matrix  $A$ , dimension  $n$  and array  $P$  from Algorithm 1.  
**Output:** Adjacency matrix  $B$  of 2-path signed network.  
**Process:**

- (1) Enter the order  $n$  and adjacency matrix  $A$  of for a given signed network  $\Sigma$ .
- (2) **for**  $i = 1$  to  $n$  **do**
- (3)     **for**  $j = 1$ :  $n$  **do**
- (4)          $B(i, j) = 0$
- for**  $t = 1$ :  $m - 1$ ,  $t = t + 2$  **do**
- (5)     **if**  $(P(t) \neq 0)$  **then**
- (6)         **if**  $((P[t] = j) \&\&(P[t + 1] = k))$  **then**
- (7)              $B(k, j) = -1$
- (8)              $B(j, k) = -1$
- (9)         **for**  $t = 1$ :  $g - 1$ ,  $t = t + 2$  **do**
- (10)         **if**  $((Q[t] = j) \&\&(Q[t + 1] = k))$  **then**
- (11)              $B(k, j) = 1$
- (12)              $B(j, k) = 1$

ALGORITHM 2: Algorithm to obtain a 2-path signed network for a given signed network.

of matrix  $B$ . Next, with the help of Algorithm 1, we obtain  $P$  as  $Q$  as

$P$	1	3	1	2	0	0	0	0	0					
$Q$	1	4	2	4	2	3	3	4	3	2	4	3	4	2

In Step 5, we enter the loop running upto  $m = 7$  for each pair  $P[t], P[t + 1]$  which saves the vertices forming negative edge in  $(\Sigma)_2$ . As an edge  $jk$  is given by entry  $j$ th row and  $k$ th column and by  $k$ th row and  $j$ th in the adjacency matrix, for each pair  $P[t] = j$  and  $P[t + 1] = k$ ,  $B(j, k) = B(k, j) = -1$ . Thus, for  $t = 1$ ,  $B(1, 3) = B(3, 1) = -1$ , similarly, for  $t = 3$   $B(1)[2] = B(2)[1] = -1$ . Thus, the matrix  $B$  after Step 9 is

$$A = \begin{bmatrix} 0 & -1 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (3)$$

Next, in Step 10, we enter the loop to find the vertices forming positive edges in  $(\Sigma)_2$ , which is given by pair of vertices in array  $Q$ . For  $t = 1$ ,  $Q[1] = 1$ ,  $Q[2] = 4$  and so  $B(1, 4) = B(4, 1) = 1$ . After running the loop for  $g = 13$ , matrix  $B$  is given by

$$A = \begin{bmatrix} 0 & -1 & -1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \quad (4)$$

which is the adjacency matrix for the 2-path signed network  $(\Sigma)_2$ , as shown in Figure 4(b).

### 3. Balance in 2-Path Signed Network

**3.1. Characterization of Balanced 2-Path Signed Network.** In this section, we provide with a characterization of balanced 2-path signed network. The characterization helps us to identify the groups in the interference network (2-path signed network) and find the balanced. We refer the following lemma given by Zaslavsky.

**Lemma 4** (see [54]). *A signed network in which every chordless cycle is positive and is balanced.*

**Theorem 2.** *For a signed network  $\Sigma$  of order  $n$ , the following statements are equivalent:*

- (i)  $(\Sigma)_2$  is balanced.
- (ii)
  - (a) Each homogeneous cycle in  $\Sigma$  is either positive or is a **P** cycle of length  $4k$ , for some positive integer  $k$ .
  - (b) Each heterogeneous odd cycle in  $\Sigma$  either does not contain **P** section or contains even number of **P** section of even length.
  - (c) For each heterogeneous even cycle (length greater than 4), the following conditions hold:
    - (c:1) If it has **P** sections of odd length  $l$ , then either  $l \equiv 1 \pmod{4}$  or if  $l \equiv 1 \pmod{4}$ ; then, there are even number of such **P** sections each separated by positive section of even length.
    - (c:2) If it has **P** sections of even length, then there are even number of such **P** sections each separated by positive section of even length.
  - (d) For each vertex  $u$  in  $\Sigma$  with  $d(u) \geq 3$ ,  $N_*(u)$  does not contain any **P** pairs.

*Proof.* (i)  $\implies$  (ii) Let  $(\Sigma)_2$  be balanced. If a cycle in  $\Sigma$  is a positive homogeneous cycle, then the cycle (or cycles in case the cycle in  $\Sigma$  is of even length) formed in its 2-path signed graph  $(\Sigma)_2$  will be a positive cycle. Let  $C$  be an all-negative cycle (**P** cycle), with length  $\neq 4k$  in  $\Sigma$ . Here, the length of  $C$  is either odd or  $2t$ , where  $t$  is odd. By Lemma 3, the corresponding cycle or cycles formed in  $(\Sigma)_2$  will be odd length all-negative, since cycle  $C$  is a **P** cycle in  $\Sigma$ . Thus, by Lemma 4,  $(\Sigma)_2$  is not balanced, which is a contradiction. Hence,  $\Sigma$  does not contain a **P** cycle with length  $\neq 4k$  as  $(\Sigma)_2$  is balanced. Thus, (a) of (ii) follows.

Next, let there exist a heterogeneous cycle  $C_1: v_1 v_2 \dots v_p v_1$ ,  $p$  being odd in  $\Sigma$ , with a **P** section

$v_1 v_2, \dots, v_k$ ,  $k$  being odd. Clearly, the length of **P** section is even. The cycle  $C_1$  in  $\Sigma$  will generate an odd cycle  $C'_1: v_1 v_3, \dots, v_{k-2} v_k, \dots, v_p v_2, \dots, v_{p-1} v_1$  in  $(\Sigma)_2$ . Now,  $v_1 v_3, v_3 v_5, \dots, v_{k-2} v_k$  are odd number of negative edges in  $C'_1$  of  $(\Sigma)_2$  and also  $v_2 v_4, \dots, v_{k-3} v_{k-1}$  are even number of negative edges in the cycle  $C'_1$  of  $(\Sigma)_2$  (see (b) in Figure 5, here  $k = 3$ ). Thus, the cycle  $C'_1$  in  $(\Sigma)_2$ , will be a negative cycle and hence  $(\Sigma)_2$  is not balanced, which is a contradiction. Similarly, for odd number of such **P** sections, the signed network  $(\Sigma)_2$  is unbalanced, whereas if there are even number of such **P** sections in the cycle of  $\Sigma$ ; then, there would be even number of negative edges in the corresponding cycle of  $(\Sigma)_2$ .

Next, let us consider a heterogeneous even cycle  $C_1: v_1 v_2, \dots, v_p v_1$  in  $\Sigma$ ,  $p$  even and  $v_1 v_2, \dots, v_l$  be a **P** section of odd length  $l$ . From Lemma 2, we know that a cycle of even length  $p$  in  $\Sigma$  gives rise to two cycles in  $(\Sigma)_2$  of length  $p/2$  each. Now, the **P** section  $v_1 v_2, \dots, v_l$  would correspond to  $(l-1)/2$  negative edges in both the cycles, clearly as  $(\Sigma)_2$  is balanced, that is, each of these cycles have even number of negative edges which is possible if  $(l-1)/2$  is an even number or there is another negative edge which is due to some other **P** section in the same cycle  $\Sigma$ . If  $(l-1)/2$  is even or  $l \equiv 1 \pmod{4}$ , then we are done, if not then we prove that there are even number of such **P** section. Clearly,  $l \equiv 1 \pmod{4}$ . Let  $v_i, \dots, v_k$ , where  $i \neq k$  and  $i, k \in \{l+1, l+2, \dots, p\}$ , be another **P** section of length  $s$ , such that  $s \not\equiv 1 \pmod{4}$ . Then, again as proved above there would be  $(s-1)/2$  negative edges in the cycles of  $(\Sigma)_2$ . Now, as each cycle has odd number of negative edges in  $(\Sigma)_2$  due to **P** sections  $v_1, \dots, v_p$  and  $v_i, \dots, v_k$  in  $\Sigma$ , therefore there are now even number of edges in each of the cycles generated by cycle  $C_1$  in  $(\Sigma)_2$ , which makes these cycles balanced.

Lastly, let  $u$  be a vertex in  $\Sigma$  with degree greater than or equal to 3. If neighborhood  $N_*(u)$  contains exactly one **P** pair, say  $\{v_1, v_2\}$ , then there exist atleast one vertex  $v_3$  in  $N_*(u)$  which does not form **P** pair with vertices  $v_1, v_2$ . Thus, by definition of 2-path,  $v_1 v_2 v_3$  will be a cycle in  $(\Sigma)_2$  with exactly one negative edge  $v_1 v_2$ . Thus, 2-path signed network is not balanced. Similarly, if neighborhood  $N_*(u)$  contains more than one **P** pair and there exist atleast one vertex in  $N_*(u)$  which does not form **P** pair, then again there is a cycle of length three with exactly one negative edge. Next, if all the vertices in  $N_*(u)$  are **P** pairs, then any three vertices  $v_1, v_2, v_3$  in  $N_*(u)$  give rise to a negative cycle of length three in  $(\Sigma)_2$  as  $\{v_1, v_2\}$ ,  $\{v_2, v_3\}$  and  $\{v_1, v_3\}$  are **P** pairs in  $N_*(u)$ . Thus, again making  $(\Sigma)_2$  unbalanced, which is a contradiction to our hypothesis. Therefore, no neighborhood of a vertex of degree greater than three in  $\Sigma$  contains a **P** pair.

(ii)  $\implies$  (i) Let, if possible, (a), (b), (c), and (d) hold. Let us consider the neighborhood for each vertex  $v_i$  in  $\Sigma$ . If the neighborhood consist of a single vertex, then it does not contribute to the edges in  $(\Sigma)_2$ . Next, if the neighborhood contains two vertices say  $v_j, v_k$ , then either  $v_j v_k$  is an edge in  $(\Sigma)_2$  which is part of a cycle or a tree. If the edge  $v_j v_k$  is part of the tree, then it is balanced by default. Next, if it is part of a cycle in  $(\Sigma)_2$ , then by Lemma 2 and Lemma 3, we know that this cycle in  $(\Sigma)_2$  is due to a cycle in  $\Sigma$ . If the corresponding cycle in  $\Sigma$  is homogeneous, then by condition (a) either it is

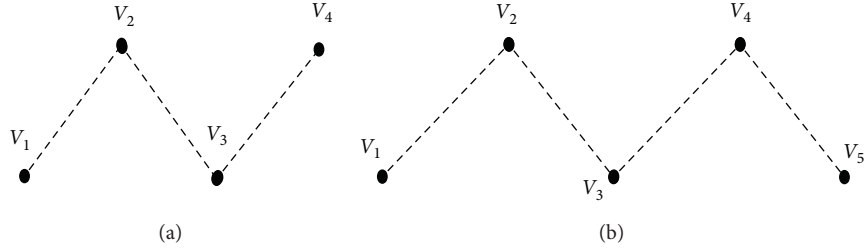


FIGURE 5: (a) A P section of length three. (b) A P section of length four.

positive or a P cycle of length  $4k$ ,  $k$  being odd. A positive cycle in  $\Sigma$  gives rise to a positive cycle in  $(\Sigma)_2$ , thus making the cycle balanced, else a cycle of even length greater than 4 in  $\Sigma$  gives rise to two cycles in  $(\Sigma)_2$  of equal length. Thus, the cycles formed in  $(\Sigma)_2$  due to the P cycle in  $\Sigma$  are of even length and thus balanced. If the cycle is heterogeneous in  $\Sigma$ , then the following cases arise:

- (1) If cycle in  $\Sigma$  is of odd length, then by (b) the cycle either does not contain P section or contains even number of P section of even length. If the cycle does not contain a P section then cycle in  $(\Sigma)_2$  is also positive hence balanced. Next, if the cycle in  $\Sigma$  contains odd number of P section of even length then the cycle in  $(\Sigma)_2$  contain even number of negative edges (as an even P section in  $\Sigma$  gives rise to odd negative edges in cycle of  $(\Sigma)_2$  as the negative even section in Figure 5(b) gives rise to three negative edges 13, 35, and 24 in 2-path signed network). Hence, again giving rise to a balanced cycle.
- (2) If cycle in  $\Sigma$  is of even length greater than 4, then by (c) either it has P sections of odd length  $l$  such that  $l \equiv 1 \pmod{4}$  and if  $l \not\equiv 1 \pmod{4}$  then there are even number of such P sections each separated by positive section of even length or as P sections of even length then there are even number of such P sections each separated by positive section of even length. In both the cases, even number of negative edges appear in both the cycles of  $(\Sigma)_2$ , thus making these cycles balanced.

Next, if the neighborhood of  $v_i$  for some  $i = 1$  to  $n$  contain more than two vertices, they give rise to a clique in  $(\Sigma)_2$ . Now, the negative edges of these cliques are due to the P pairs in each neighborhood but from (d) no neighborhood contains a P pair; thus, all the cliques are positive. Therefore, by virtue of the conditions, all cycles (and cliques) in  $(\Sigma)_2$  will be positive, and thus by Lemma 4,  $(\Sigma)_2$  will be balanced.  $\square$

## 4. Clusterability

**4.1. Clusterability in 2-Path Signed Networks.** In this section, we discuss the clusterability of a 2-path signed network.

**Lemma 5** (see [55]). *A signed network  $\Sigma$  is clusterable if and only if  $\Sigma$  contains no cycle with exactly one negative edge.*

**Theorem 3.** *For a given signed network  $\Sigma$  of order  $n$ , the following conditions are equivalent:*

- (i)  $(\Sigma)_2$  is clusterable.
- (ii) For all sequence of vertices  $v_1, v_2, \dots, v_r; 1 \leq r \leq n$  in  $\Sigma$  such that  $v_1, v_2 \in N(t_1); v_2, v_3 \in N(t_2); \dots; v_1, v_r \in N(t_r)$  for some  $t_1, t_2, \dots, t_r \in V(\Sigma)$ . If there exist a pair of vertices in sequence  $v_i, v_{i+1} \in N(t_i)$  having property P, then the sequence has atleast one pair of vertices  $v_l, v_{l+1} \in N(t_l), l \neq i$  satisfying property P, for some  $l, 1 \leq l \leq r$ .
- (iii)
  - (a) If  $\Sigma$  contains a heterogeneous cycle, then no even cycle in  $\Sigma$  contains exactly one P section of length  $< 5$  and no odd cycle contains exactly one P section of length 2.
  - (b) Each neighborhood of vertex  $v$  in  $\Sigma$  with  $d(v) \geq 3$  either does not contain a P pair or all vertices are P pairs.

*Proof.* (i)  $\implies$  (ii) Let for a given signed network  $\Sigma$ ,  $(\Sigma)_2$  be clusterable. Then, no cycle in  $(\Sigma)_2$  has exactly one negative edge. Let  $v_1, v_2, \dots, v_r; 1 \leq r \leq n$  be a sequence of vertices in  $\Sigma$  such that  $v_1, v_2 \in N(t_1); v_2, v_3 \in N(t_2); \dots; v_1, v_r \in N(t_r)$  for some  $t_1, t_2, \dots, t_r \in V(\Sigma)$ . Let  $v_i, v_{i+1} \in N(t_i)$  be a pair in sequence  $v_1, v_2, \dots, v_r$ ; for some  $r, 1 \leq r \leq n$  having property P. Now,  $C: v_1 v_2 \dots v_r v_1$  is a cycle in  $(\Sigma)_2$ , due to the sequence  $N(t_1), \dots, N(t_r)$  in  $\Sigma$ . Clearly,  $v_i, v_{i+1}$  will form a single negative edge in cycle  $C$  in  $(\Sigma)_2$ , which is not possible. Thus, there is atleast one more pair of vertices  $v_l, v_{l+1} \in N(t_l), l \neq i$  satisfying property P.

(ii)  $\implies$  (i) Let  $v_1, v_2, \dots, v_r; 1 \leq r \leq n$  be an arbitrary sequence of vertices in  $\Sigma$  such that  $v_1, v_2 \in N(t_1); v_2, v_3 \in N(t_2); \dots; v_1, v_r \in N(t_r)$  for some  $t_1, t_2, \dots, t_r \in V(\Sigma)$ . If there exist a pair of vertices  $v_i, v_{i+1} \in N(t_i)$  for some  $i$  having property P, then the sequence has atleast one other pair of vertices  $v_l, v_{l+1} \in N(t_l)$  for some  $l \in \{1, \dots, r\}$  satisfying property P. This sequence of vertices generates cycles in  $(\Sigma)_2$  such that no cycle has exactly one negative edge. Thus, by Lemma 5,  $(\Sigma)_2$  is clusterable.

(i)  $\implies$  (iii) Let  $(\Sigma)_2$  be clusterable. To prove that (iii) (a) and (iii) (b) hold, let, if possible, (a) does not hold. This implies, there exist an even heterogeneous cycle  $C$  in  $\Sigma$  of length  $2k$ , with P section of length  $< 5$ .  $C$  will give rise to two

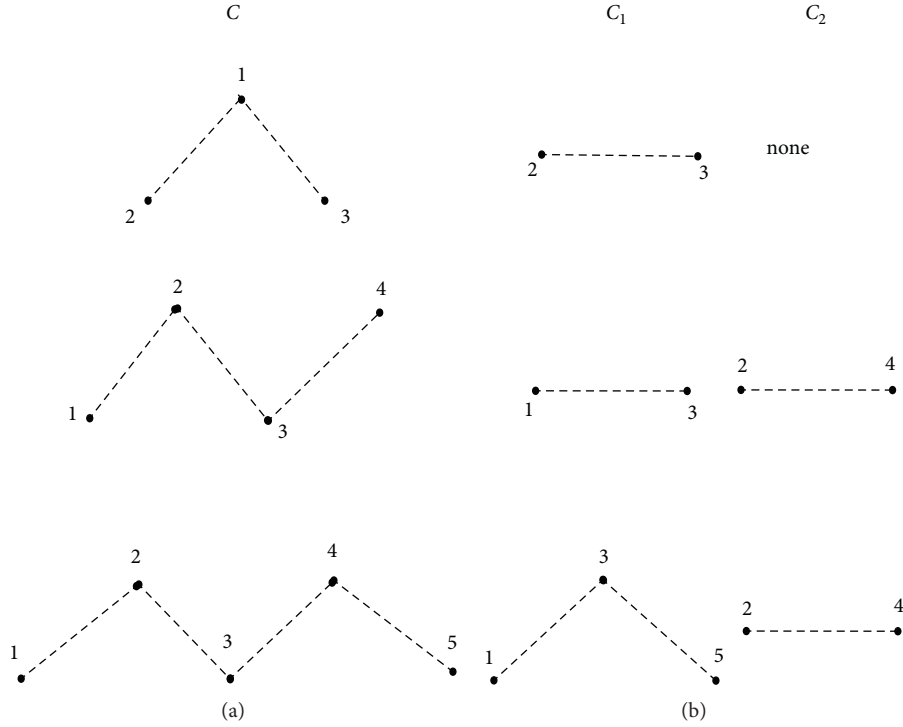


FIGURE 6: (a) A P section of length less than 5 in cycle  $C$  of  $\Sigma$ . (b) The corresponding P sections in cycle  $C_1$  and  $C_2$  in  $(\Sigma)_2$ .

cycles  $C_1$  and  $C_2$  of length  $k$  each in  $(\Sigma)_2$ , and clearly, at least one of the cycles will contain exactly one negative edge (see Figure 6), which is a contradiction to the hypothesis by Lemma 5. Next, if there exist an odd heterogeneous cycle with P section of length 2, then it will correspond to a single negative edge in the cycle of  $(\Sigma)_2$ , which is again not possible. Hence, (iii)(a) and (iii)(b) hold.

Next, we assume that (b) in (iii) does not hold. Let  $N_*(u)$  be a neighborhood of a vertex  $u$  in  $\Sigma$  with  $d(u) \geq 3$  and containing at least one P pair. Let  $v_1$  be a vertex in  $N_*(u)$ , which does not form a P pair with any other vertex in  $N_*(u)$ . Clearly, as  $d(u) \geq 3$ ,  $\exists v_2, v_3$  such that  $\{v_2, v_3\}$  is a P pair. The 2-path signed network  $(\Sigma)_2$ , will now contain a cycle  $v_1 v_2 v_3 v_1$  with exactly one negative edge  $v_2 v_3$ , since each neighborhood  $N(u)$  gives rise to clique  $\delta(N(u))$ . By Lemma 5, it is a contradiction to the hypothesis, whereas if all vertices in  $N_*(u)$  form P pairs, then all the edges in the corresponding clique will be negative, thus  $(\Sigma)_2$  remains clusterable.

(iii)  $\implies$  (i) Assume that condition (a) and (b) in (iii) hold. We have to show that  $(\Sigma)_2$  is clusterable. By Theorem 1, we know that  $(\Sigma)_2$  is obtained by taking the union of cliques generated by the neighborhood of vertices of  $\Sigma$ . Thus, each cycle in  $(\Sigma)_2$  is either due to cliques generated by  $N(t)$  for  $t \in V(\Sigma)$  such that  $|N(t)| \geq 3$  or due to induced cycles in  $\Sigma$ . Now, by condition (a), no heterogeneous even cycle in  $\Sigma$  contains exactly one P section of length  $< 5$ . Thus, each cycle formed in  $(\Sigma)_2$  has either no negative edge or has more than one negative edge. Also, if no odd cycle in  $\Sigma$  contains exactly one P section of length 2, then its corresponding cycle in  $(\Sigma)_2$  does not contain exactly one negative edge. From (b), it

is clear that no clique in  $(\Sigma)_2$  contains exactly one negative edge as the clique formed here is homogeneous; hence, by Lemma 5  $(\Sigma)_2$  is clusterable.  $\square$

## 5. Sign-Regularity

5.1. *Property of Sign-Regularity in 2-Path Signed Networks.* In this section, we establish a characterization of a sign-regular 2-path signed network. Note that

$$\rho_{v_i} = \left| \bigcup_{t \in V(\Sigma)} \{N(t) : v_i \in V(\Sigma)\} \right|, \quad (5)$$

$$|v_i| = \left| \left\{ \{v_i, v_j\} \in \mathbf{P}_* : v_j \in V(\Sigma) \right\} \right|.$$

**Theorem 4.** *For a signed network  $\Sigma$  of order  $n$ ,  $(\Sigma)_2$  is sign-regular if and only if*

- (i) *For all vertices  $v_i, v_j \in V(\Sigma)$  such that  $v_i \in N(v_j)$ ,  $\rho_{v_i}$  is identical for every  $v_i \in V(\Sigma)$*
- (ii) *If  $\mathbf{P}_*$  is collection of all P pairs, then  $|v_i|$  is identical for all  $1 \leq i \leq n$*

*Proof. Necessity.* Let for a given signed network  $\Sigma$ ,  $(\Sigma)_2$  be sign-regular. Then, number of positive and negative edges incident to each vertex in  $(\Sigma)_2$  is identical. Since  $(\Sigma)_2$  is obtained by taking the union of cliques generated by the neighborhood of vertices of  $\Sigma$ , thus the total number of vertices in each neighborhood containing  $v_i$  is the same for each vertex  $v_i \in V(\Sigma)$ . Also,  $\rho_{v_i}$  gives the total number of

**Input:** Adjacency matrix  $A$  of  $\Sigma$ , number of vertices  $n$  and array  $P$  from Algorithm 1.

**Output:** Whether the 2-path of a given network is sign-regular or not.

**Process:**

```

(1) Enter the adjacency matrix  $A$  with elements  $A(i, j)$  for a given signed network  $\Sigma$  along with the order  $n$  of signed network  $\Sigma$ .
(2)  $h = 1$ ;
(3) for  $i = 1: n$  do
(4)    $count[i] = 0$ ;
(5)    $count1[i] = 0$ ;
(6)    $countrow[i] = 0$ ;
(7) for  $i = 1: n$  do
(8)   for  $j = 1: n$  do
(9)     if  $(A(i, j) \neq 0)$  then
(10)       $countrow[i] = countrow[i] + 1$ ;
(11) for  $i = 0$  to  $n$  do
(12)   for  $j = 1: n$  do
(13)     if  $(A(i, j) \neq 0)$  then
(14)       $count[j] = count[j] + countrow[i] - 1$ ;
(15) for  $i = 1: n$  do
(16)   for  $j = 1: n$  do
(17)     if  $(count[i] \neq count[j])$  then
(18)      Print The 2-path network is not regular.
(19)       $h = 0$ ; break;
(20) for  $i = 1: n$  do
(21)   for  $j = 1: l$  do
(22)     if  $(i = P[j])$  do
(23)       $count1[i] = count1[i] + 1$ .
(24) for  $i = 1: n$  do
(25)   for  $j = 1: n$  do
(26)     if  $count1[i] \neq count1[j]$  then
(27)      Print "The two path network is not regular."
(28)       $h = 0$ ;
(29)      break;
(30) if  $(h \neq 0)$  then
(31)   Print "The network is sign-regular."

```

ALGORITHM 3: Algorithm to check sign-regularity of 2-path signed network.

edges incident to a vertex  $v_i, i \in \{1, \dots, n\}$ . Hence,  $\rho_{v_i}$  is identical  $\forall v_i \in V(\Sigma)_2$ . Therefore, (i) holds. Since  $\mathbf{P}_*$  consists of all pair of vertices satisfying property  $\mathbf{P}$ , hence number of negative edges incident to each vertex in  $(\Sigma)_2$  must be equal. Thus, the vertex  $v_i$  appearing in the number of  $\mathbf{P}$  pairs in  $\mathbf{P}_*$ ,  $|v_i|$  is the same for every  $i \leq i \leq n$ .

*Sufficiency.* Let (i) and (ii) hold. Now, (i) suggests that, for all the vertices  $v_i$  in  $V(\Sigma)$  such that  $v_i \in N(v_j)$ , the union of all these neighborhoods which gives the total vertices adjacent to  $v_i$  in  $(\Sigma)_2$  have the same cardinality. Thus, each vertex in  $(\Sigma)_2$  is adjacent to the same number of vertices. Next, we know that the elements of  $\mathbf{P}_*$  generate all the negative edges of  $(\Sigma)_2$  and  $|v_i|$  gives the number of negative edges incident to  $v_i$  in  $(\Sigma)_2$ . By (ii), the cardinality  $|v_i|$  is the same for each vertex  $v_i, i = 1, \dots, n$ . Thus, the number of negative edges incident to  $v_i$  is the same for each  $i = 1$  to  $n$ . Since the number of total edges and negative edges incident to vertex  $v_i$  is the same for each  $i$ , the number of positive edges will also be same. Therefore,  $(\Sigma)_2$  is sign-regular.  $\square$

*5.2. Algorithm to Detect if 2-Path Signed Network of a Signed Network Is Sign-Regular.* Algorithm 3 detects if, for a given signed network  $\Sigma$ , its 2-path  $(\Sigma)_2$  is sign-regular, by using results of Theorem 4. Consider the adjacency matrix  $A$  and its order  $n$  as input. The vector  $countrow$  gives the number of nonzero elements in each row (degree of the vertices). The array  $count$  gives the number of edges for each vertex  $v$  present in some neighborhood of a vertex  $u$ . Also, vector  $count1$  counts the number of negative edges in  $(\Sigma)_2$  for each vertex  $u$ .

*5.2.1. Complexity.* From Steps 3 to 6, initialization of the vector arrays uses a single loop which runs upto  $n$ . Thus, the complexity of these steps =  $O(n)$ .

In Steps 7 and 8, two loops are used upto Step 10. Also, at Steps 11, 15, and 24, again two loops running upto  $n$  are used independent of each other. Thus, they have combined complexity =  $O(n^2)$ .

In Step 21, the  $\mathbf{P}$  pairs are collected, and we know that the complexity of this step is of order  $n^3$ . Finally, the two loops in

Steps 20 and 21 go upto  $n$  and  $l$ , respectively. Thus, the complexity of this step =  $O(nl)$ . Total complexity =  $O(n) + O(n^2) + O(n^3) + O(nl) = O(n^3)$ .

### 5.2.2. Implementation of Algorithm

*Example 3.* In this example, we implement Algorithm 3, for a signed network, as shown Figure 4(a). We want to check whether, for a given signed network, its corresponding 2-path signed network is sign-regular. This is done by the adjacency matrix (as in Example 1) of the given signed network  $\Sigma$ . Now, in the given signed network,  $n=4$ , so the vector arrays *count* and *count1* are initialized as zero along with variable  $h$  which is initialized as 1. After we enter the loop at Step 7 and Step 8, we fetch each nonzero value of the matrix. For  $i=1$  and  $j=2$   $count[2] = 1$  (as 2 and 3 are in the neighborhood of 1 and  $count[2] = 0 + 2 - 1$ ) and  $count[3] = 1$ . Next, for  $i=2$   $count[1] = 0 + 3 - 1 = 2$ ,  $count[3] = 3$  and  $count[4] = 2$ . Similarly, moving further, when  $i=3$ , we get  $count[1] = 2 + 3 - 1 = 4$ ,  $count[2] = 1 + 3 - 1 = 3$ , and  $count[4] = 2 + 3 - 1 = 4$ . For  $i=4$ ,  $count[2] = 4$  and  $count[3] = 4$ . In Step 14, we check that if there exist  $i, j$ , such that  $count1[i] \neq count1[j]$ , but since this is not true we move to Step 20 and collect all  $\mathbf{P}$  pairs. Next, in Step 21, we check whether the number of negative edges incident to each vertex is identical for all the vertices. We count the number of appearance of each vertex in array  $P$  of Algorithm 1; this is done by array *count1* as  $\mathbf{P}$  is the following:

$P$	1	3	1	2	0	0	0	0
-----	---	---	---	---	---	---	---	---

Thus,  $count1[1] = 2$ ,  $count[2] = 1$ ,  $count[3] = 1$ , and  $count[4] = 0$ . Next, in Step 26, we check these entries of array *count1* and initially find that  $count1[1] \neq count[2]$ . Hence, the number of negative edges incident to vertex 1 and 2 is not the same; thus, the 2-path signed network is not sign-regular. The same is clear from Figure 4(b), hence the proof.

## 6. Conclusion

In this paper, we studied various characterizations of 2-path signed networks and other allied properties such as balancedness, clusterability, and sign-regularity. We designed a model using the 2-path signed networks on radio frequency interference. We assumed that the vertices represented channel/stations and transmission between them was represented by edges. The negative sign in signed network  $\Sigma$  was given to the edge  $uv$  if the transmission from  $u$  and  $v$  takes place at the same time and frequency, otherwise  $uv$  was given a positive sign. The paper explored both theoretic and applicable aspect of 2-path signed networks. In our work, we not only focused on the characterization and other results due to the signing but also the algorithms which can be readily used in real world problems.

## Data Availability

No data were used to support the findings of the study.

## Conflicts of Interest

All the authors declare that they have no conflicts of interest regarding the publication of this paper.

## Acknowledgments

Authors express gratitude to Professor Thomas Zaslavsky for his careful reading, valuable comments, and fruitful suggestions that improved the paper throughout. The first author's work is supported by the Research Grant from DST (MTR/2018/000607) under Mathematical Research Impact Centric Support (MATRICS) for a period of 3 years (2019–2022).

## References

- [1] T. A. McKee and F. R. McMorris, *Topics in Intersection Graph Theory*, SIAM, Philadelphia, PA, USA, 1999.
- [2] M. Pal, "Intersection graphs: an introduction," 2014, <http://arxiv.org/abs/1404.5468>.
- [3] D. Hermelin, J. Mestre, and D. Rawitz, "Optimization problems in dotted interval graphs," *Discrete Applied Mathematics*, vol. 174, pp. 66–72, 2014.
- [4] P. Diaconis, S. Holmes, and S. Janson, "Interval graph limits," *Annals of Combinatorics*, vol. 17, no. 1, pp. 27–52, 2013.
- [5] J. Spinrad, "On comparability and permutation graphs," *SIAM Journal on Computing*, vol. 14, no. 3, pp. 658–670, 1985.
- [6] S. Paul, M. Pal, and A. Pal, "L (2, 1)-labeling of permutation and bipartite permutation graphs," *Mathematics in Computer Science*, vol. 9, no. 1, pp. 113–123, 2015.
- [7] M. Bonamy, M. Johnson, I. Lignos, V. Patel, and D. Paulusma, "Reconfiguration graphs for vertex colourings of chordal and chordal bipartite graphs," *Journal of Combinatorial Optimization*, vol. 27, no. 1, pp. 132–143, 2014.
- [8] F. Gavril, "The intersection graphs of subtrees in trees are exactly the chordal graphs," *Journal of Combinatorial Theory, Series B*, vol. 16, no. 1, pp. 47–56, 1974.
- [9] J. Köbler, S. Kuhnert, and O. Verbitsky, "Solving the canonical representation and star system problems for proper circular-arc graphs in logspace," *Journal of Discrete Algorithms*, vol. 38, pp. 38–49, 2016.
- [10] M. C. Lin, F. J. Soulignac, and J. L. Szwarcfiter, "Normal helly circular-arc graphs and its subclasses," *Discrete Applied Mathematics*, vol. 161, no. 7, pp. 1037–1059, 2013.
- [11] V. Klee, "What are the intersection graphs of arcs in a circle?" *The American Mathematical Monthly*, vol. 76, no. 7, pp. 810–813, 1969.
- [12] J. Fox and J. Pach, "Applications of a new separator theorem for string graphs," *Combinatorics, Probability and Computing*, vol. 23, no. 1, pp. 66–74, 2014.
- [13] D. Sinha and M. Acharya, "Characterization of signed graphs whose iterated signed line graphs are balanced or s-consistent," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 32, pp. 1–10, 2015.
- [14] T. Zaslavsky, "Consistency in the naturally vertex-signed line graph of a signed graph," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 39, pp. 307–314, 2016.
- [15] C. L. Monma and V. K. Wei, "Intersection graphs of paths in a tree," *Journal of Combinatorial Theory, Series B*, vol. 41, no. 2, pp. 141–181, 1986.
- [16] A. Asinowski, E. Cohen, M. C. Golumbic, V. Limouzy, M. Lipshteyn, and M. Stern, "Vertex intersection graphs of

- paths on a grid,” *Journal of Graph Algorithms and Applications*, vol. 16, no. 2, pp. 129–150, 2012.
- [17] J. Lee and B. Park, “Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 81–90, 2012.
- [18] F. Librino, M. Levorato, and M. Zorzi, “An algorithmic solution for computing circle intersection areas and its applications to wireless communications,” *Wireless Communications and Mobile Computing*, vol. 14, no. 18, pp. 1672–1690, 2014.
- [19] T. Erlebach and K. Jansen, *Scheduling of Virtual Connections in Fast Networks*, Univ., Mathematik/Informatik, 1996.
- [20] T. Przytycka, G. Davis, N. Song, and D. Durand, “Graph theoretical insights into evolution of multidomain proteins,” *Journal of Computational Biology*, vol. 13, no. 2, pp. 351–363, 2006.
- [21] J. Zhao, O. Yağan, and V. Gligor, “Monotone increasing properties and their phase transitions in uniform random intersection graphs,” 2015, <http://arxiv.org/abs/1502.00405>.
- [22] B. D. Acharya and M. N. Vartak, *Open Neighborhood Graphs*, Indian Institute of Technology Department of Mathematics Research Report, Bombay, India, 1973.
- [23] F. Heider, “Attitudes and cognitive organization,” *The Journal of Psychology*, vol. 21, no. 1, pp. 107–112, 1946.
- [24] H. Fritz, *The Psychology of Interpersonal Relations*, Psychology Press, London, UK, 2013.
- [25] H. Frank, “On the notion of balance of a signed graph,” *The Michigan Mathematical Journal*, vol. 2, no. 2, pp. 143–146, 1953.
- [26] H. Frank, M.-H. Lim, and D. C. Wunsch, “Signed graphs for portfolio analysis in risk management,” *IMA Journal of Management Mathematics*, vol. 13, no. 3, pp. 201–210, 2002.
- [27] H.-P. Kriegel, P. Kröger, and A. Zimek, “Clustering high-dimensional data,” *ACM Transactions on Knowledge Discovery from Data*, vol. 3, no. 1, p. 1, 2009.
- [28] I. Giotis and V. Guruswami, “Correlation clustering with a fixed number of clusters,” in *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithm*, Miami, FL, USA, January 2006.
- [29] D. Sinha and S. Anshu, “An algorithm to detect balancing of iterated line sigraph,” *SpringerPlus*, vol. 4, no. 1, p. 1, 2015.
- [30] B. D. Acharya, “Signed intersection graphs,” *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 13, no. 6, pp. 553–569, 2010.
- [31] M. Acharya and D. Sinha, “Common-edge sigraphs,” *AKCE International Journal of Graphs and Combinatorics*, vol. 3, no. 2, pp. 115–130, 2006.
- [32] M. Acharya and D. Sinha, “Characterization of signed line digraphs,” *Discrete Applied Mathematics*, vol. 161, no. 9, pp. 1170–1172, 2013.
- [33] K. A. Germina, S. Hameed K, and T. Zaslavsky, “On products and line graphs of signed graphs, their eigenvalues and energy,” *Linear Algebra and its Applications*, vol. 435, no. 10, pp. 2432–2450, 2011.
- [34] M. K. Gill and G. A. Patwardhan, “Switching invariant two-path signed graphs,” *Discrete Mathematics*, vol. 61, no. 2-3, pp. 189–196, 1986.
- [35] D. Sinha and D. Sharma, “Algorithmic characterization of signed graphs whose two path signed graphs and square graphs are isomorphic,” in *Proceedings of the 2014 International Conference on Soft Computing Techniques for Engineering and Technology (ICSCTET)*, Nainital, India, August 2014.
- [36] D. Sinha and D. Sharma, “On square and 2-path signed graph,” *Journal of Interconnection Networks*, vol. 16, no. 1, Article ID 1550011, 2016.
- [37] D. Sinha and D. Sharma, “Characterization of 2-path product signed graphs with its properties,” *Computational Intelligence and Neuroscience*, vol. 2017, 8 pages, 2017.
- [38] P. Siva Kota Reddy and M. S. Subramanya, “Note on path signed graphs,” *Notes on Number Theory and Discrete Mathematics*, vol. 15, no. 4, pp. 1–6, 2009.
- [39] G. Sheng, Y. Su, and W. Wang, “A new fractal approach for describing induced-fracture porosity/permeability/compressibility in stimulated unconventional reservoirs,” *Journal of Petroleum Science and Engineering*, vol. 179, pp. 855–866, 2019.
- [40] H. Zhao, L. Xu, Z. Guo et al., “A new and fast waterflooding optimization workflow based on insim-derived injection efficiency with a field application,” *Journal of Petroleum Science and Engineering*, vol. 179, pp. 1186–1200, 2019.
- [41] A. Dey, L. H. Son, A. Pal, and H. V. Long, “Fuzzy minimum spanning tree with interval type 2 fuzzy arc length: formulation and a new genetic algorithm,” *Soft Computing*, vol. 24, no. 6, pp. 3963–3974, 2020.
- [42] A. Dey, S. Mondal, and T. Pal, “Robust and minimum spanning tree in fuzzy environment,” *International Journal of Computing Science and Mathematics*, vol. 10, no. 5, pp. 513–524, 2019.
- [43] A. Dey, L. Son, G. P. Kumar, S. Quek, and S. Gai Quek, “New concepts on vertex and edge coloring of simple vague graphs,” *Symmetry*, vol. 10, no. 9, p. 373, 2018.
- [44] E. Malesinska, *Graph Theoretical Models for Frequency Assignment Problems*, Shaker, Maastricht, Germany, 1997.
- [45] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, USA, 1996.
- [46] R. Coase, W. Meckling, and J. Minasian, *Problems of Radio Frequency Allocation*, RAND Corporation, Santa Monica, CA, USA, 1995.
- [47] J. Pan, H. Wang, X. Gao et al., “Radio-frequency interference estimation using equivalent dipole-moment models and decomposition method based on reciprocity,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 1, pp. 75–84, 2015.
- [48] S. Misra, R. D. De Roo, and C. S. Ruf, “An improved radio frequency interference model: reevaluation of the kurtosis detection algorithm performance under central-limit conditions,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4565–4574, 2012.
- [49] H. Wang, V. Khilkevich, Y.-J. Zhang, and J. Fan, “Estimating radio-frequency interference to an antenna due to near-field coupling using decomposition method based on reciprocity,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 55, no. 6, pp. 1125–1131, 2013.
- [50] H. Frank, *Graph Theory*, Addison Wesley, Reading, MA, USA, 1969.
- [51] D. B. West, *Introduction to Graph Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 2001.
- [52] T. Zaslavsky, “A mathematical bibliography of signed and gain graphs and allied areas,” *The Electronic Journal of Combinatorics*, vol. 1000, 2012.
- [53] D. Sinha and D. Sharma, “On 2-path signed graphs,” in *Proceedings of the International Workshop on Computational Intelligence (IWCI)*, Dhaka, Bangladesh, December 2016.
- [54] T. Zaslavsky, “Signed analogs of bipartite graphs,” *Discrete Mathematics*, vol. 179, no. 1–3, pp. 205–216, 1998.
- [55] J. A. Davis, “Clustering and structural balance in graphs. Social networks. A developing paradigm,” pp. 27–34, 1977.

## Research Article

# Legal Judgment Prediction Based on Multiclass Information Fusion

Kongfan Zhu, Rundong Guo , Weifeng Hu, Zeqiang Li, and Yujun Li 

*School of Information Science and Engineering, Shandong University, Qingdao 266200, China*

Correspondence should be addressed to Yujun Li; [liyujun@sdu.edu.cn](mailto:liyujun@sdu.edu.cn)

Received 25 June 2020; Revised 20 August 2020; Accepted 4 October 2020; Published 26 October 2020

Academic Editor: Shirui Pan

Copyright © 2020 Kongfan 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.

Legal judgment prediction (LJP), as an effective and critical application in legal assistant systems, aims to determine the judgment results according to the information based on the fact determination. In real-world scenarios, to deal with the criminal cases, judges not only take advantage of the fact description, but also consider the external information, such as the basic information of defendant and the court view. However, most existing works take the fact description as the sole input for LJP and ignore the external information. We propose a Transformer-Hierarchical-Attention-Multi-Extra (THME) Network to make full use of the information based on the fact determination. We conduct experiments on a real-world large-scale dataset of criminal cases in the civil law system. Experimental results show that our method outperforms state-of-the-art LJP methods on all judgment prediction tasks.

## 1. Introduction

Legal judgment prediction (LJP) aims to predict the judgment results according to the information based on fact determination, which consists of the fact description, the basic information of defendant, and the court view. LJP techniques can provide inexpensive and useful legal judgment results to people who are unfamiliar with legal terminologies, and they are also helpful for the legal consulting. Moreover, they can serve as a handy reference for professionals (e.g., lawyers and judges), which can improve their work efficiency.

LJP is regarded as a classic text classification problem and has been researched for many years [1]. For example, Liu et al. proposed to extract shallow textual features (e.g., Chinese characters, words, and phrases) for charge prediction [2]. Katz et al. predicted the US Supreme Court's decisions based on efficient features from case profiles [3]. Luo et al. combined the fact description with the corresponding law articles to predict the charges [4]. Although great progress has been made in the LJP, there still exist some problems, such as multiple subtasks, topological dependencies between subtasks, and cases of similar descriptions

with different penalties. Zhong et al. pointed out that law articles prediction was one of the fundamental subtasks in some countries (e.g., China, France, and Germany) with the civil law system, and these subtasks had a strict order in the real world [5]. Further, Yang et al. proposed a neural model for the interaction between subtask results [6].

Despite these efforts in designing efficient features and employing advanced Natural Language Processing (NLP) techniques, LJP still confronts two major challenges.

**1.1. The Lack of External Information.** Some existing works propose various mechanisms to extract information from the fact description, such as the Word Collection Attention mechanism. Some other works propose various frameworks to build the dependencies between subtasks, such as DAG Dependencies of Subtasks and MPBF. However, for the judgment document in Figure 1, there are many other information items that can be utilized except the fact description. Such information is called the external information including the basic information of defendant and the court view. Therefore, how to utilize the external information effectively is a major challenge.



<p>被告人洪流，男，1986年9月18日出生于湖北省黄梅县，……同年6月28日被取保候审。</p>	<p>The basic information of defendant</p>	<p>The defendant, Hong Liu, male, was born in huangmei county, hubei province on september 18, 1986. ……and was released on bail pending trial on june 28 of the same year.</p>
<p>武汉市汉阳区人民检察院指控，2013年6月9日12时许，被告人洪流在武汉市汉阳区琴台大道塞纳河畔小区2号楼楼下，盗走王某华停放在此处价值人民币2275元的大众牌电动车1辆。……</p>	<p>The fact description</p>	<p>The people's procuratorate of hanyang district of wuhan city accused the defendant of stealing at 12 o'clock on june 9, 2013, the under the no. 2 building of the seine river district, qintai avenue, hanyang district, wuhan city. one volkswagen electric car. ……</p>
<p>本院认为，被告人洪流以非法占有为目的，秘密窃取他人价值人民币2275元的财物，数额较大，其行为已构成盗窃罪。……辩护人辩称被告人洪流归案后认罪态度较好，可以从轻处罚的观点，与本案事实、证据和法律规定相符，本院予以采纳。</p>	<p>The court view</p>	<p>The court believes that the defendant's torrent secretly steals other people's property worth RMB 2275 for the purpose of illegal possession, which is a large amount, and his act has constituted the crime of theft. ……The defender argued that the defendant had a better attitude of pleading guilty after returning to the case, and could accept the fact that the case was light, consistent with the facts, evidence and legal provisions of the case.</p>

FIGURE 1: A judgment document example in China (original Chinese text and its English translation).

1.2. *Encoding Long Document Is Difficult.* The fact description in judgment document is often long document containing the long-term dependency problem. Many existing models, such as Recurrent Neural Network (RNN) [7] and Convolutional Neural Network (CNN) [8], which perform well in the text processing are unable to deal with the long-term dependency problem. There are some keywords in the judgment document that are very important for LJP. It is very difficult to find them in the judgment document.

In order to resolve the above challenges, in this paper, we propose the Transformer-HAN-Multi-Extra (THME) Network. It contains a structured data encoder to extract the semantics of the external information as well as a Transformer-Hierarchical Attention Network (TH) encoder to encode the fact description. Specifically, as shown in Figure 1, from the basic information of the defendant, we can get the defendant's gender, age, and education level and the content related to the criminal records of the defendant by using regular expressions. Similarly, we can get some objective attributes of a case, such as amount, plot, and consequences, from the court view. Based on the statistical analysis of large samples, we can find the relationship between the data and the terms of penalty as is shown in Table 1, where the symbol "+" represents "related." For example, given the same conditions, male's terms of penalty is longer than female's for certain cases. We use the symbol "↑" to denote positive correlation. For example, the more serious the case's plot is, the longer the defendant's terms of penalty will be. We use the symbol "↓" to denote negative correlation. For example, the better the defendant's guilty attitude is, the shorter the defendant's terms of penalty will be. It is worth noting that the case's conclusion in judgment document is significant for terms of penalty but it cannot be used as an input to predict the terms of penalty. If it is used as an input to predict the terms of penalty, it seems like that the cat shuts its eyes when stealing. Therefore, we first use the external information to predict the case's conclusion and then use it together with the external information to predict the terms of

penalty. Meanwhile, according to the data attributes, we divide the data into continuous and discrete types. Then, we extract the required information via the continuous data encoder and the discrete data encoder. In order to reduce the information loss in the process of converting sentences into fixed-length vectors, an attention mechanism is adopted. But, it cannot solve the polysemy problem. Then, we choose a proper Transformer [9]. Transformer has attention structure; it has advantages over the RNN in solving long-term dependency problem and performs better than attention on polysemy. The Hierarchical Attention Network (HAN) can catch the keywords in a long document easily [10]. Thus, we can combine the Transformer with the HAN to solve the long-term dependency problem. Experimental results show that the performance of Transformer-HAN is better than Gate Recurrent Unit (GRU)-HAN.

The main contributions of this paper are summarized as follows:

- (i) We propose a novel text processing structure, namely, Transformer-HAN, to improve the text encoding ability. This model can solve the long-term dependency problems better than the GRU-HAN. Transformer-HAN encoder uses the attention mechanism in addition to the necessary fully connected layer of the parameter matrix, and it works much faster than the encoder structure based on GRU and Long Short-Term Memory (LSTM).
- (ii) We propose a structured data encoder. To introduce the external information as an auxiliary, we extract fact-related data from the defendant's basic information and the court view as supplementary information of the model. According to different attributes of data, we design both continuous and discrete data encoders. Experiments show that information based on fact determination can effectively improve the judgment prediction, especially for the prediction of the terms of penalty.

TABLE 1: Origins and categories of structured data.

Origin	Category
The basic information of defendant	Defendant’s gender (+), defendant’s education level (↑), defendant’s previous convictions (+), defendant’s number of previous convictions (↑), defendant’s terms of penalty of previous convictions (↑), defendant’s penalty of previous convictions (↑)
The court view	Case’s amount of money involved (↑), case’s conclusion (↑), case’s amount (↑), case’s plot (↑), case’s consequence (↑), defender’s guilty attitude (↓)

- (iii) Experimental results show that the THME Network can effectively improve the prediction accuracy of few-shot data. The macro-average indicators of the three tasks of law article prediction, charge prediction, and terms of penalty prediction are relatively improved compared with other models, which indicates that the prediction accuracy of few-shot data has been greatly improved.

The rest of this paper is organized as follows. Section 2 briefly reviews the related work. In Section 3, we propose the overall THME framework and detailed methods. The experimental results and analyses are presented in Section 4. Finally, Section 5 contains the concluding remarks.

## 2. Related Work

*2.1. Legal Judgment Prediction.* With the development of Chinese legal digitalization process, as one of the most critical task steps in LegalAI, LJP has become more and more important. Thanks to the development of machine learning and text mining techniques, more researchers formalize this task under text classification frameworks. Most of these studies attempt to extract textual features [11–13] or introduce some external knowledge [4, 14]. However, these methods can only utilize shallow features and manually designed factors; usually the effect of these methods becomes worse when applied to other scenarios. Therefore, researchers take advantage of other technologies to improve the interpretability and generalization of the model. For example, Jiang et al. utilized the deep reinforcement learning to derive short snippets of documents from the fact descriptions to predict charges [15], and Chen et al. proposed a Legal Graph Network (LGN) to achieve high-precision classification of crimes [16]. Due to the rareness of some types of cases in real life, the few-shot problem is inevitable. While some researchers hardly solve this problem using machine learning, others find that neural networks have good results. For example, Chen et al. proposed a neural network model by embedding law articles and fact descriptions into the same embedding space in the same way [17]. Yang et al. proposed a repeated interactional mechanism to simulate the process of judge’s decision [18].

*2.2. Multitask Learning.* Multitask models have many beneficial effects for deep learning tasks. Sulea et al. proposed multiple tasks, which include law articles predictions, charge predictions, and terms of penalty predictions, to test the application of machine learning in the judicial field [19].

Zhong et al. proposed a topological structure network, which can simulate the judge’s judgment process to improve the performance of various tasks. Yang et al. designed a Multi-Perspective Bi-Feedback Network (MPBFN) to enhance the connection between tasks and allow tasks’ results to flow in both directions. Wang et al. set the relationship between law articles as a tree structure via a Hierarchical Matching Network (HMN) and matched relevant law articles via a two-layer matching network [20], which can improve the work efficiency.

The emergence of multitask learning has promoted the development of LJP; however, due to the lack of external information, it has also resulted in unsatisfactory prediction of terms of penalty. In this work, we propose a framework to utilize the external information effectively. Different from most existing works, we extract the information from both the fact description and the external information and merge them together into a topological classifier to predict the three subtasks of LJP.

## 3. Method

In this section, we will describe the THME Network. We first give the essential definitions of the LJP task and the composition of THME Network in Sections 3.1 and 3.2, respectively. We describe a text encoder for fact descriptions in Section 3.3. We introduce the structured data encoder in Section 3.4. Finally, the classifier is proposed in Section 3.5.

*3.1. Problem Formulation.* In most tasks of the Chinese text processing, the char-granularity processing is superior to the word-granularity processing [21], so for each judgment document, we set each Chinese character as a token. The fact description is a token sequence  $T = (t_1, t_2, t_3, \dots, t_N)$ , where  $N$  is the number of tokens. This can reduce the complexity of model and make it fit easier. Besides the input  $T$ , the basic information of the defendant and the court view are also deemed as external inputs of the structured data encoder. Given these inputs, we will predict the judgment results of applicable law articles, charges, and terms of penalty, which is a multitask classification problem.

*3.2. Overview.* Our THME consists of three parts, i.e., the text encoder, the structured data encoder, and the classifier. The text encoder is composed of text embedding layer, text convolution layer, main encoder layer, and information extraction layer. Due to different attributes of the structured

data, we divide structured data into discrete data and continuous data, for which we propose discrete data encoder and continuous data encoder, respectively. The classifier is implemented with a topological structure, which utilizes the topological dependencies between subtasks in LJP. The general framework of the THME is shown in Figure 2.

We employ a text encoder to extract the information from the fact description; the fact description is embedded into CNN, so that advanced features are gradually extracted from the shallow textual features.  $c_{ij}$  represents the  $j$ -th Chinese character in the  $i$ -th sentence. The main encoder layer is actually Transformer-HAN, which includes two layers: the first layer aggregates token-level features into sentence-level features, and the second layer aggregates sentence-level features into text-level features. Finally, we generate four hidden-layer states  $T_1, T_2, T_3$  corresponding to three subtasks of LJP and  $T_4$  corresponding to the case's conclusion which is critical in predicting the terms of penalty through the information extraction layer. Next, we employ the regular expression to extract the discrete data and the continuous data from the external information. Then, we standardize the continuous data, embed the discrete data, and input them into the discrete data encoder and continuous data encoder, respectively. The outputs of these two encoders are combined to generate the structured data vector  $T'_m, T'_m$  and the hidden-layer state  $T_4$  are concatenated into a full connection network to predict the case's conclusion  $T_{dc}$ . The case's conclusion vector  $T_{dc}$  and the structured data vector  $T'_m$  make up the output of the structured data encoder  $T_m$ . Finally,  $T_m$  and the hidden-layer state of all subtasks in LJP  $T_1, T_2, T_3$  are concatenated into the classifier with topological structure to predict the law articles, charges, and terms of penalty.

**3.3. Text Encoder for Fact Description.** We employ a text encoder to generate the vector of fact description as the input of the classifier. We will briefly introduce this encoder which is composed of lookup layer, convolution layer, Transformer-HAN layer, and information extraction layer.

**3.3.1. Lookup and Convolution.** Taking a token sequence  $T$  as input, the encoder computes a simple text representation through two layers, i.e., lookup layer and convolution layer.

(1) *Lookup.* We first convert each token  $t_i$  in  $T$  into a natural number  $d_i \in N$  by preprocessed dictionary mapping. The token sequence  $T$  is converted into an integer sequence  $D = (d_1, d_2, d_3, \dots, d_N)$ . Next, we propose an initialized word embedding sequence  $E = (e_0, e_1, e_2, \dots, e_s)$ ,  $e_i \in R^k$ , where  $s$  is the size of dictionary.  $d_i$  is mapped to  $x_i$  via the word embedding sequence  $E$ . Thus, we can obtain the text embedding sequence  $X = (x_1, x_2, x_3, \dots, x_N)$ ,  $x_i \in R^k$ , where  $k$  is the length of word embedding.

(2) *Convolution.* For  $X$ , we make a convolution operation with the convolution matrix  $W \in R^{m \times (l \times k)}$  given by

$$c_i = W \cdot x_{i:i+l-1} + b_c, \quad (1)$$

where  $x_{i:i+l-1}$  is the concatenation of word embeddings in the  $i$ -th window,  $b_c \in R^m$  is the bias vector,  $m$  is the number of filters, and  $l$  is the size of a sliding window. We apply the convolution over each window  $i$  and finally obtain  $C = (c_1, c_2, c_3, \dots, c_N)$ . The Chinese character vector after convolution has  $n$ -gram features; that is to say, the Chinese character vector after convolution has context features and is no longer isolated.

**3.3.2. Transformer-HAN Encoder and Information Extraction.** (1) *Transformer-HAN encoder.* Transformer is currently the most mainstream information extractor, mainly due to its unique attention mechanism, which achieves the true bidirectional encoding. However, the number of parameters of the multilayer Transformer encoder is very huge. In order to fully take advantage of Transformer and meanwhile constrain the number of parameters, we design the Transformer-HAN as our main encoder.

Transformer-HAN encoder is divided into two layers: the first layer uses Transformer for Chinese character-granularity coding, then uses the attention mechanism to extract the most important information in each word embedding, and combines them into sentence vectors. The second layer uses Transformer for sentence-granularity coding, then uses the attention mechanism to extract the most important information in sentence vectors, and combines them into a chapter-granularity vector. Therefore, the fact description is divided into  $m$  sentences  $C = (c_1, c_2, \dots, c_m)$ , and the  $i$ -th sentence consists of  $n$  Chinese characters  $c_i = (c_{i1}, c_{i2}, c_{i3}, \dots, c_{in})$ , where  $m \times n = N$ .

Since the Transformer encoder is less sensitive to the position of Chinese characters, we need to add the position embedding to the word embedding before input. For Chinese character in the  $j$ -th sentence  $c_j$ , we calculate its position vector  $P_j$  as

$$P(\text{pos}, 2i) = \sin\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right), \quad (2)$$

$$P(\text{pos}, 2i + 1) = \cos\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right),$$

where  $\text{pos}$  is the position of this Chinese character in the sentence,  $i$  is the index of the  $i$ -th value in its word embedding, and  $d_{\text{model}}$  is the dimension of its word embedding. The position vectors of all Chinese characters in the sentence  $c_j$  form the sequence  $P_j$ . Then, we merge the position sentence  $P_j$  with  $c_j$  to obtain the sentence sequence with the information of position  $C_{pj}$  given by

$$C_{pj} = P_j \oplus c_j, \quad (3)$$

where  $\oplus$  is an element-wise addition operation.

The Transformer encoder is composed of Multihead Attention (MHA), Add & Norm Layer, and Feed Forward (FF). Multihead Attention is composed of Self-Attention, for which the inputs  $Q$ ,  $K$ , and  $V$  are the same. Multihead

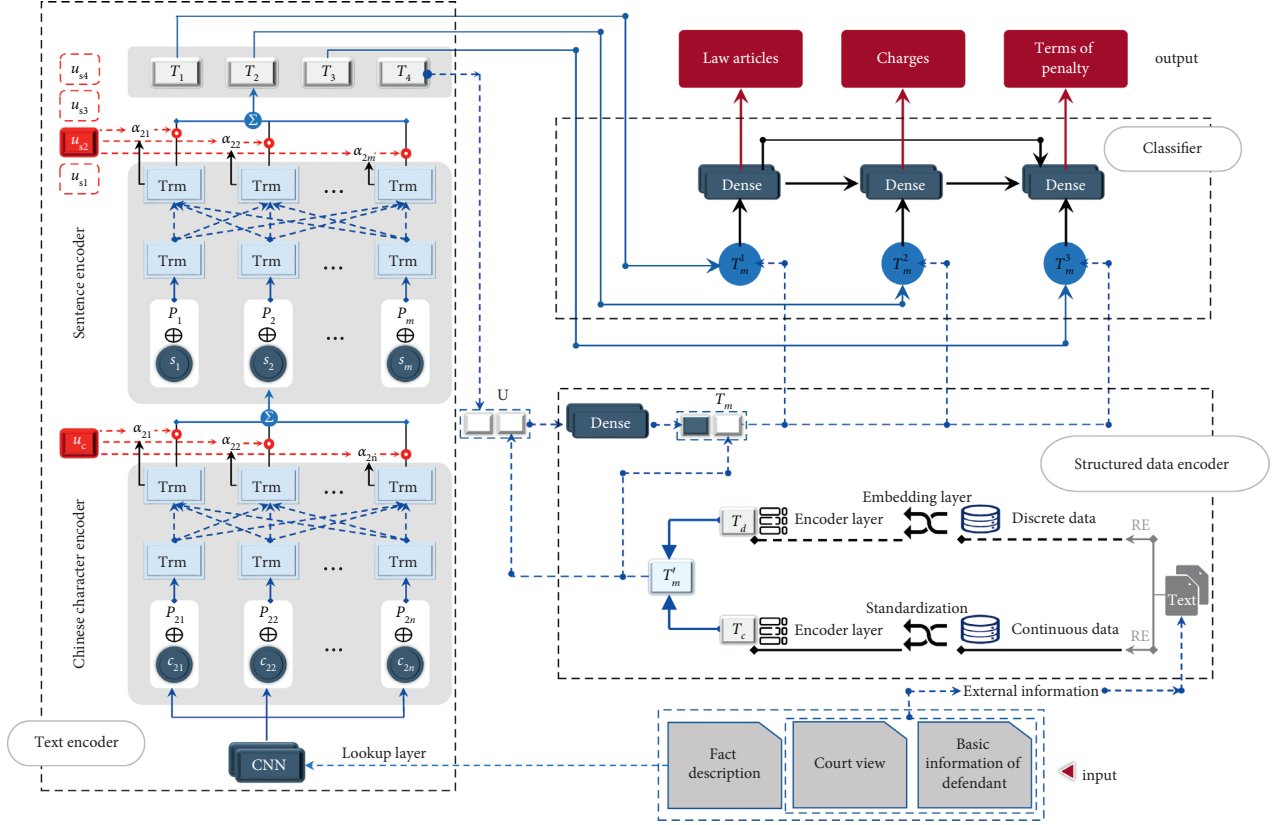


FIGURE 2: General THME framework. Trm denotes the Transformer, and RE denotes the regular expression.  $u_{ic}$  represents attention vector in the  $i$ -th sentence;  $u_{s1}, u_{s2}, u_{s3}$  represent attention vectors which extract textual features of three subtasks of LJP ( $T_1, T_2, T_3$ ) from sentence-level sequence; and  $u_{s4}$  represent attention vectors which extract textual features of case's conclusions ( $T_4$ ) similarly.

Attention converts  $Q, K$ , and  $V$  into  $Q', K'$ , and  $V'$  through linear transformation by using a parameter matrix. Next, we apply the Self-Attention mechanism to extract the semantic information. This process is repeated  $h$  times. The results are concatenated together, and then the linear transformation is performed. The calculation process is given as follows:

$$\begin{aligned}
 Q &= C_{pj}, K = C_{pj}, V = C_{pj}, \\
 Q'_i &= QW_i^Q, K'_i = KW_i^K, V'_i = VW_i^V, \\
 \text{Head}_i &= \text{Attention}(Q'_i, K'_i, V'_i) = \text{softmax}\left(\frac{Q'_i K_i'^T}{\sqrt{d_k}}\right) V'_i, \\
 \text{MHA}(Q, K, V) &= \text{concat}(\text{Head}_1, \text{Head}_2, \dots, \text{Head}_h) W^0,
 \end{aligned} \tag{4}$$

where  $\text{concat}()$  is the vector concatenation operation,  $d_k$  is the size of head, and  $W^0, W_i^Q, W_i^K, W_i^V \in R^{k \times (k/h)}$  are the parameter matrices.

Add & Norm Layer contains the Add layer and the Norm layer. First, we merge the input of Multihead Attention  $C_{pj}$  with the output of MHA and obtain the fact semantic vector  $M_j$  as

$$M_j = C_{pj} \oplus \text{MHA}. \tag{5}$$

There are two reasons for this: First, it can make up for the lack of information. Second, it is equivalent to introducing a highway in the network. When the network is backpropagating, a part of it can be directly propagated into the original information without going through the complex network, preventing gradient explosion or gradient disappearance. Then, we employ the Layer Normalization [22] to normalize  $M_j$  and obtain  $M'_j = (m'_{j1}, m'_{j2}, m'_{j3}, \dots, m'_{jn})$ ,  $m'_{ji} \in R^k$ . Therefore, we obtain the sentence sequence  $M_j = (m_{j1}, m_{j2}, m_{j3}, \dots, m_{jn})$  as

$$M_j = \text{Relu}((M'_j W^1 + b_1) W^2 + b_2), \tag{6}$$

where  $W^1, W^2 \in R^{k \times k}$  are the parameter matrices and  $b_1, b_2$  are the basic vectors. Then, we use the attention vector to extract the main information. In order to get the sentence vector  $s_j$ , we initialize an attention vector  $u_w \in R^n$  and obtain  $s_j$  as

$$\begin{aligned}
 a_i &= u_{ic} \cdot m_{ji}, \\
 a'_i &= \text{softmax}(a_i) = \frac{\exp(a_i)}{\sum_{j=1}^n \exp(a_j)}, \\
 s_j &= \sum_{i=1}^n a'_i m_{ji}.
 \end{aligned} \tag{7}$$

Similarly, we get the sentence sequence  $S = (s_1, s_2, \dots, s_m)$ . The sentence encoder is basically the same as the

Chinese character encoder. The difference is that the token vector is replaced with a sentence vector which is produced by the Chinese character encoder.

Since we still use the Transformer to encode the sentence sequence, we first calculate the sentence's position vector  $P_s$  and merge it with the sentence sequence  $S$  by

$$C_{P_s} = P_s \oplus S. \quad (8)$$

As the input of the Transformer,  $C_{P_s}$  passes the Transformer's MHA, Add & Norm Layer, and Feed Forward to obtain a new sentence sequence  $S' = (s'_1, s'_2, \dots, s'_m)$ , which has higher-level characteristics and more comprehensive and useful information.

(2) *Information extraction.* Finally, for our three subtasks of LJP and case's conclusion, we need four different attention vectors to extract four different kinds of information from the same information sequence. We first initialize four attention vectors  $u_{s1}, u_{s2}, u_{s3}, u_{s4} \in R^m$  and obtain the vector  $T_j \in R^n$  as

$$\begin{aligned} s_i^n &= \tanh(s_i' W^{T_j} + b_{T_j}), \\ t_i &= u_{s_j} \cdot s_i^n, \\ t_i' &= \text{softmax}(t_i) = \frac{\exp(t_i)}{\sum_{j=1}^l \exp(t_j)}, \end{aligned} \quad (9)$$

$$T_j = \sum_{i=1}^l t_i' s_i^n,$$

where  $W^{T_j}$  is the fully connected matrix and  $b_{T_j}$  is the bias vector.

**3.4. Structured Data Encoder.** The deep learning model is like a judge. We train the model and keep feeding data to the model, just like constantly showing different cases to the judge and training the professional quality of the judge. However, most of the previous work only gave the model to "see" the fact description. In practice, the judge would not sentence the defendant only based on the fact description at the time of judging. In the process of judgment prediction, we sometimes need some explicit data to convict and sentence the defendant. For example, information such as the defendant's guilty attitude, whether to commit recidivism, and the amount of money involved directly affect the final judgment. Based on the above facts, we use the regular expression to extract discrete data and continuous data from the external information, as shown in Tables 2 and 3. In order to well integrate data into THMA, we design both the discrete data encoder and the continuous data encoder, as shown in Figure 3.

**3.4.1. Continuous Data Encoder.** We normalize each category of continuous data as

$$c_i' = \frac{c_i - \mu_c}{\sigma_c}, \quad (10)$$

where  $\mu_c$  is the mean of continuous data and  $\sigma_c$  is the variance. We can obtain the continuous data sequence  $C' = (c_1', c_2', c_3', \dots, c_g')$ , where  $g$  is the number of types of continuous data. Then, we employ a full connection network to fuse different types of continuous data and obtain the continuous data vector  $T_c$  as

$$T_c = \text{Relu}(C' W^c + b_c), \quad (11)$$

where  $W^c$  is the fully connected matrix,  $b_c$  is the bias vector, and  $T_c \in R^p$ .

**3.4.2. Discrete Data Encoder.** Since there are few discrete data categories, we use the word embedding method to create a discrete data vector space for each category of discrete data. We convert each category of discrete data into its word embedding  $d_i' \in R^w$ . Similarly, we obtain the discrete data vector  $T_{di}$  as

$$T_{di} = \text{Relu}(d_i' W^{di} + b_{di}), \quad (12)$$

where  $W^{di}$  is the fully connected matrix,  $b_{di}$  is the bias vector, and  $T_{di} \in R^p$ . The discrete data sequence is then represented as

$$T_d = (T_{d1}, T_{d2}, \dots, T_{dq}), \quad (13)$$

where  $q$  is the number of categories of discrete data.

**3.4.3. Case's Conclusion Prediction.** The specific content of the case's conclusion is presented in Table 4.

In order to predict the case's conclusion, we firstly obtain the combination of discrete data sequence and continuous data vector as  $T_m'$ , given by

$$T_m' = \text{Concat}(T_c, T_d). \quad (14)$$

Case's conclusion is very helpful for LJP, especially for the prediction of terms of penalty. For prediction of case's conclusion, the input  $U$  is the concatenation of the case's conclusion corresponding vector  $T_4$  and  $T_m'$ . Similarly, we obtain the vector of case's conclusion  $T_{dc}$  as

$$T_{dc} = \text{Relu}(U W^{dc} + b_{dc}), \quad (15)$$

where  $W^{dc}$  is the fully connected matrix,  $b_{dc}$  is the bias vector, and  $T_{dc} \in R^{dc}$ . Finally, we obtain the output of the structured data encoder as

$$T_m = \text{Concat}(T_m', T_{dc}). \quad (16)$$

**3.5. Classifier.** When a judge decides a case, he/she often first searches for the legal basis related to this case such as the fact description. Then, according to the relevant laws, the conviction is made. Finally, intergrating all the evidence and facts, the judge passes the sentence. Therefore, there are topological dependencies among multitask results [5]. We evaluate the performance on three LJP subtasks, including law articles (denoted as  $t_1$ ), charges (denoted as  $t_2$ ), and

TABLE 2: Continuous data.

Category	Case's amount of money involved	Number of previous convictions	Terms of penalty of previous convictions	Penalty of previous convictions
Attribute	Case's information	Information of previous convictions	Information of previous convictions	Information of previous convictions

TABLE 3: Discrete data.

Category	Case's amount	Case's plot	Case's consequences	Defendant's guilty attitude	Defendant's previous convictions
Attribute	Huge amount	The plot is lighter	Causes serious consequences	Surrenders oneself	Several
	Larger amount	The plot is bad	Causes significant losses	Admits actively	
	Extremely huge amount	The plot is serious	Cause particularly serious consequences	Guilty attitude is good	
	Huge quantity	The plot is particularly bad	Causes particularly significant losses	Guilty attitude is very good	
	Larger quantity	The plot is particularly serious			
	Extremely huge amount	The plot is slight			

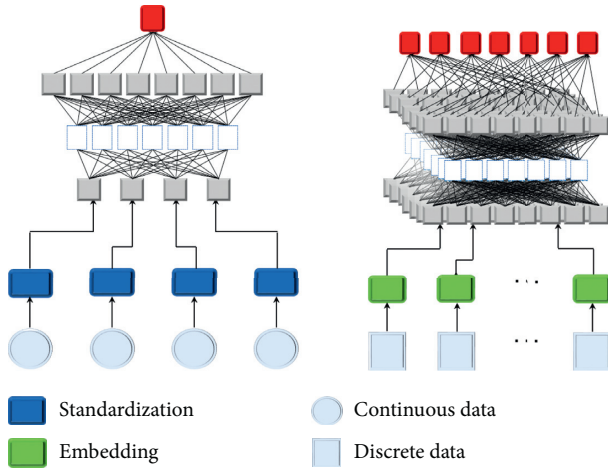


FIGURE 3: Continuous data encoder and discrete data encoder.

TABLE 4: Case's conclusion.

Category	Case's conclusion
Attribute	Light punishment
	Lenient punishment
	Mitigated punishment
	Severe punishment
	Heavier punishment

terms of penalty (denoted as  $t_3$ ). Note that we implement the classifier with dependency in Figure 2; i.e.,

$$\begin{aligned}
 H_1 &= \phi, \\
 H_2 &= t_1, \\
 H_3 &= t_1, t_2,
 \end{aligned} \tag{17}$$

where  $H_i$  represents the input of  $t_i$  and  $\phi$  is the empty set. This means that the charge prediction depends on law articles, and the terms of penalty prediction depend on both law articles and charges. Such explicit dependencies conform to the judicial logic of human judges, which will be verified in later sections. In order to combine the fact description and the structured data, we concatenate the structured data vector  $T_m$  and the  $i$ -th subtask's corresponding vector  $T_i$  to obtain the vector  $T_m^i$  as

$$T_m^i = \text{Concat}(T_m, T_i), \quad i = 1, 2, 3. \tag{18}$$

Considering the topological dependencies between subtasks, we predict the law article first, then the charge, and finally the terms of penalty. We obtain the law article's vector  $T_l$  as

$$\begin{aligned}
 T_l^1 &= \text{Relu}(T_m^1 W_m^1 + b_m^1), \\
 T_l &= \text{softmax}(\text{Relu}(T_l^1 W_l^1 + b_l^1)).
 \end{aligned} \tag{19}$$

The processes of charge prediction and terms of penalty prediction are similar with the law article prediction. Different from the law article prediction, the input of the charge prediction is the concatenation of  $T_m^2$  and  $T_l^1$ , while the input of terms of penalty prediction is the concatenation of  $T_m^3$ ,  $T_l^1$ , and  $T_{ch}^1$ . Finally, we obtain  $T_l \in R^x$ ,  $T_{ch} \in R^y$ , and  $T_p \in R^z$ , where  $x, y, z$  are the number of categories of label for subtasks 1, 2, 3, respectively. In order to learn parameters of THME model, we use the Adam algorithm [23]. We adopt the cross-entropy loss in the training process as follows:

$$L_i^i = -[y \log \hat{y} + (1 - y) \log (1 - \hat{y})], \tag{20}$$

where  $\hat{y}$  is the prediction result,  $y$  is the real result,  $l$  is the law articles prediction, and  $i$  is the  $i$ -th sample. Equation (20) represents the loss function of one sample in the prediction of the law articles. When there are multiple samples, we add

all the losses together to form the total loss of the law articles. We have three subtasks, so the sum of losses of the three subtasks constitutes the final loss of the model. We train our model in an end-to-end fashion and utilize the dropout [24] to prevent overfitting.

## 4. Experiments

In this section, we verify the effectiveness of our proposed model. We first introduce the datasets and the data processing. Then, we provide the necessary parameters of our model. Finally, we did some experiments to verify the advantage of our model and the importance of external information.

*4.1. Dataset Construction.* Since there are no publicly available LJP datasets in previous works, we collect and construct an LJP dataset CJO. CJO consists of criminal cases published by the Chinese government from China Judgment Online<sup>1</sup>. The data used in this experiment is all from the judgment documents published by the Supreme People’s Court of China. Before the formal data processing, we first clean the data. Our experiment aims at criminal offense, so other types of judgment documents except criminal offense are screened out. Then, we filter out the multi-criminal judgment documents. The structure of the multi-criminal judgment documents is complicated, and we will research it in our future work. The terms of penalty for a single-criminal judgment document are up to 25 years, so we screen out the judgment documents with the terms of penalty more than 25 years (except death penalty and life imprisonment). Finally, we screened 5480000 judgment documents and obtained 750000 available data pieces. We used the selected 750000 pieces of data for experiments.

Our model’s inputs include the token sequence  $T$ , the discrete data, and the continuous data. However, we find that our processing approach is not suitable for the terms of penalty of previous convictions. It cannot solve the problem of uneven distribution. Therefore, we discretize the terms of penalty. The specific method is shown in Table 5.

For the majority data in the CJO dataset, their terms of penalty are no longer than 12 months. Meanwhile, the amount of data decreases as the terms of penalty increase. Especially for those with terms of penalty longer than 3 years, the amount of data has dropped significantly. In order to solve the problem of uneven distribution, we use small intervals where data is dense and large intervals where data is sparse, so as to ensure the stability of the amount of data in each interval.

*4.2. Baselines.* To evaluate the performance of our proposed THME framework, we employ the following text classification models and judgment prediction methods as baselines:

- (i) Fact-Law Attention Model [4]: It was proposed by Luo et al. in 2017. The main idea is embedding the law article into the model and then using the fact

TABLE 5: Terms of penalty conversion table.

Terms of penalty	Conversion result
No penalty	0
0–6 months	1
6–9 months	2
9–12 months	3
1–2 years	4
2–3 years	5
3–5 years	6
5–7 years	7
7–10 years	8
10–25 years	9
Death or life imprisonment	10

descriptions to extract the relevant law article to help the model get good results.

- (ii) TOPJUDGE [5]: It was proposed by Zhong et al. in 2018. The main idea is using the topological dependencies between subtasks to improve the task effect.
- (iii) MPBFN-WCA [6]: It was proposed by Yang et al. in 2019. The main idea is that repeated iterations between subtasks can reduce the error accumulations, thereby improving the effectiveness of the tasks.

*4.3. Experimental Settings.* We set the word embedding size  $k$  as 256. For the discrete data encoder, the dimension of the discrete data embedding  $w$  is 32. The dimension of the output vector of the discrete data encoder  $q$  is 64, the dimension of the output vector of the continuous data encoder  $p$  is 64, and the dimension of the case’s conclusion’s vector  $dc$  is 256.

We use the TensorFlow framework to build neural networks. In the training part, we set the learning rate of Adam optimizer as 0.0001 and the dropout probability as 0.5. The padding length of the text  $N$  is 320 tokens, the length of each sentence  $n$  is 16 tokens, and each text is divided into 20 sentences. We set the batch size as 256 for all models. We train each model for 256 epochs, and if overfitting occurs, we will terminate the training early.

We employ accuracy (Acc.), macro – precision(MP), macro – recall(MR), and macro –  $F_1$  ( $F_1$ ) as evaluation metrics. Here, the macro-precision/recall/ $F_1$  is calculated by averaging the precision/recall/ $F_1$  of each category.

*4.4. Results and Analysis.* All the models are repeated 3 times, and we evaluate the performance on three LJP subtasks, including law articles, charges, and terms of penalty and report the average values as the final results for clear illustration. Experimental results on the test set of CJO are shown in Table 6. It is shown that THME achieves the best performance on all metrics. Thus, the effectiveness and robustness of our proposed framework are verified. Compared with TOPJUDGE and MPBFN-WCA, THME takes advantage of the information of the fact determination and

thus achieves promising improvements. It indicates that the external information enables the model to learn rules that are not in the original fact description. Compared with Fact-Law Model, our model takes advantage of the correlation among relevant subtasks and achieves significant improvements. Thus, it is important to properly model topological dependencies between different subtasks.

**4.5. Ablation Study.** To further illustrate the significance of modules in our framework. Compared to THME, we designed the following models:

- (i) Transformer-HAN-Single-Extra (THSE): We decompose the multitask model into a single-task model to verify the superiority of the multitask model.
- (ii) Transformer-HAN-Single (THS): In order to reflect the role of continuous data and discrete data based on the fact description in a single-task, we design THS to compare the effect with THSE.
- (iii) Transformer-HAN-Multi (THM): In order to reflect the role of continuous data and discrete data based on the fact description in multitasking, we design THM to compare the effect with THME.
- (iv) GRU-HAN-Multiextra (GHME): In order to prove the role of Transformer in the model, we design the GHME model and the THME to compare their effects.

As shown in Table 7, compared with THS, THM can improve the performance by 1.52%, 4.8%, and 4.53% for law article prediction, charge prediction, and terms of penalty prediction in our dataset, respectively. Thus, multitask model is beneficial to improve the performance of each task. THSE performs better than THS, especially in terms of penalty prediction. THSE has enhanced the performance by 2.51%. Thus, the structured data based on the fact description plays an important role, even if the single-task model is also significantly better than the multitask model without the addition of structured data. Hence, the structured data plays a more important role compared with the multitask structure.

Through comparing GHME and THS, we can see that THS performs better, which indicates that the performance of Transformer is better than the traditional GRU model in handling long documents and the effect of Transformer-HAN on LJP is greater than that of the multitask topological structure and the external information. This also proves that the proposed Transformer-HAN is a state-of-the-art model to deal with long-term dependency problems.

**4.6. Information Source Study.** To further show the significance of the external information and explore the impacts of the information source, we evaluate the performance of THME under various information sources. We remove all the external information (fact), court view (-court view), defendant’s information (-defendant’s information), and

case’s conclusion (-case’s conclusion), respectively. Results are summarized in Table 8.

It is shown that the performance of THME gets worse for all tasks after removing either origin of information. More specifically, when we remove all the external information, tremendous decrease is observed for the terms of penalty prediction. This demonstrates that the external information is beneficial for terms of penalty prediction. When we remove the defendant’s information, the performance is better than when removing the court view. This also demonstrates that the court view is more significant than the defendant’s information and it plays a decisive role in LJP. The case’s conclusion comes from the court view. When we remove the case’s conclusion, the effect of THME is worse than the situation of removing the defendant’s information, which is similar to the situation of removing the court view. This demonstrates that the case’s conclusion plays a very important role in LJP.

**4.7. Error Analysis and Solution.** Prediction errors induced by our proposed model can be traced down into the following causes.

**4.7.1. Data Imbalance.** Data imbalance is a natural phenomenon, because the number of cases with long terms of penalty is significantly less than those with short terms of penalty. Although we have adopted effective techniques to discretize the terms of penalty to reduce the impact of data imbalance, for the subtasks of law articles and charges, our model achieves more than 90% on accuracy, while only about 75% for macro-F1. This issue is much more severe on the subtask of the terms of penalty, for which our model yields a poor performance of only 40% macro-F1. The bad performance is mainly due to the imbalance of category labels; e.g., there are only a few training instances where the term is “life imprisonment or death penalty.” Most judgment prediction approaches perform poorly (especially for recall) on these labels as listed in Figure 4.

**4.7.2. Terms of Penalty Problem.** It can be seen from the results that although our model surpasses other models in terms of penalty prediction, the effects of terms of penalty prediction is still very poor. The accuracy rate is only 56.89%, and the macro-average index is even less than 50%. Such an index is far from meeting the actual needs. The actual cases are often multiple criminal cases, which are much more complicated than the cases we are analyzing, but complex cases often contain more information, which also provide us with ideas for solving the problem of terms of penalty prediction. In multiple criminal cases, we can split the case into multiple subcases and then comprehensively consider the categories of subcases, the number of subcases, and the severity of subcases to provide more information for terms of penalty prediction. The specific implementation method remains to be explored.



TABLE 6: Judgment prediction results on CJO.

Tasks	Metrics	Law articles				Charges				Terms of penalty			
		Acc.	F1	MP	MR	Acc.	F1	MP	MR	Acc.	F1	MP	MR
Baselines	Fact-Law	88.95	70.56	72.31	71.22	92.35	78.67	80.47	78.83	52.15	37.62	38.51	38.43
	TOPJUDGE	89.55	63.38	63.78	65.72	93.90	66.51	66.60	67.26	53.71	38.71	37.43	41.32
	MPBFN-WCA	88.35	71.54	74.09	70.90	89.78	57.54	66.78	55.81	46.00	20.40	23.74	22.83
<b>Ours</b>	<b>THME</b>	<b>90.93</b>	<b>77.26</b>	<b>80.24</b>	<b>76.43</b>	<b>96.36</b>	<b>81.77</b>	<b>83.05</b>	<b>81.04</b>	<b>56.88</b>	<b>45.71</b>	<b>47.66</b>	<b>46.07</b>

TABLE 7: Ablation study on CJO.

Tasks	Metrics	Law articles				Charges				Terms of penalty			
		Acc.	F1	MP	MR	Acc.	F1	MP	MR	Acc.	F1	MP	MR
	THS	87.69	72.88	75.31	71.26	92.23	74.71	75.01	71.33	52.10	31.89	34.30	32.01
	THM	89.40	74.13	76.10	73.60	95.14	78.97	80.07	78.28	52.78	37.78	38.73	39.14
	THSE	90.57	69.75	70.55	70.09	96.12	75.02	75.59	74.83	53.41	38.48	41.46	40.09
	GHME	86.02	50.03	55.81	50.26	94.11	53.92	59.14	53.85	49.88	27.59	32.17	28.52
<b>Ours</b>	<b>THME</b>	<b>90.93</b>	<b>77.26</b>	<b>80.24</b>	<b>76.43</b>	<b>96.36</b>	<b>81.77</b>	<b>83.05</b>	<b>81.04</b>	<b>56.88</b>	<b>45.71</b>	<b>47.66</b>	<b>46.07</b>

TABLE 8: Comparable results of the effect of different information sources in the model.

Tasks	Metrics	Law articles				Charges				Terms of penalty			
		Acc.	F1	MP	MR	Acc.	F1	MP	MR	Acc.	F1	MP	MR
Baselines	Fact	89.51	73.72	76.75	74.43	95.11	78.53	79.44	78.41	50.94	36.95	39.35	39.66
	-Court view	90.00	74.99	78.03	74.60	95.39	79.48	80.89	78.92	55.57	42.40	45.98	42.55
	-Defendant’s information	90.90	76.02	<b>81.41</b>	75.52	96.35	81.35	82.74	80.71	55.61	43.86	<b>48.05</b>	43.84
	-Case’s conclusion	90.40	75.13	77.68	75.06	96.13	80.15	81.85	79.36	55.08	43.22	46.75	43.19
<b>Ours</b>	<b>THME</b>	<b>90.93</b>	<b>77.28</b>	80.26	<b>76.44</b>	<b>96.36</b>	<b>81.77</b>	<b>83.06</b>	<b>81.04</b>	<b>56.89</b>	<b>45.72</b>	47.68	<b>46.08</b>

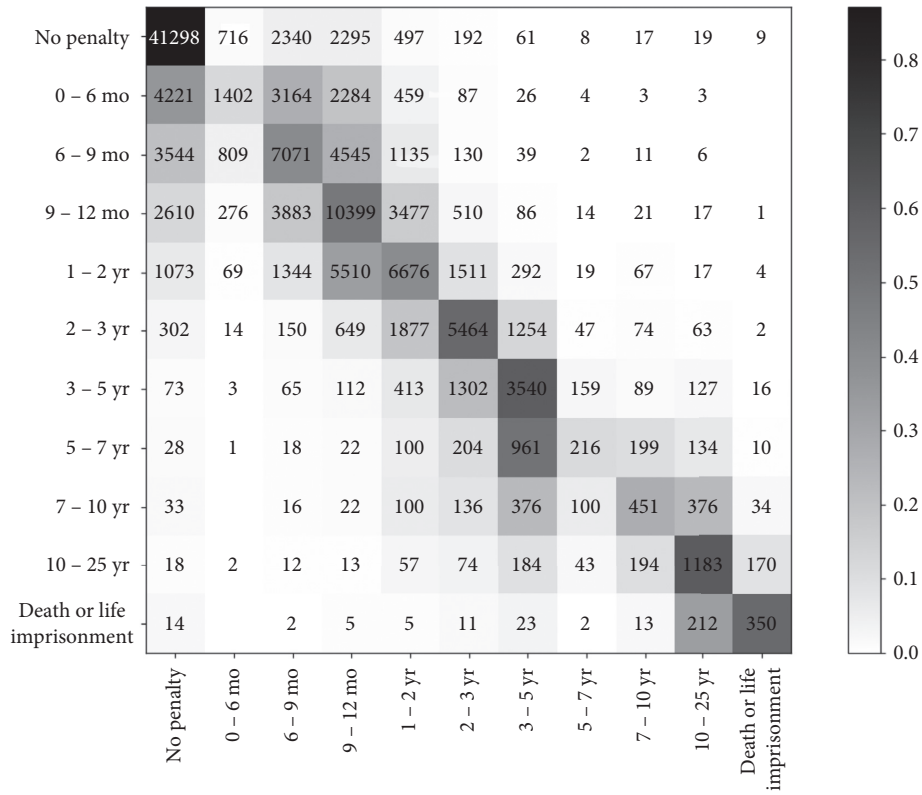


FIGURE 4: The confusion matrix in the subtask of predicting the terms of penalty. The rows denote the prediction truth while the columns denote the ground truth.

## 5. Conclusion

In this paper, we have studied the multi-extra and multi-task of LJP with topological dependencies between subtasks and address the problem of insufficient information and insufficient coding in LJP. Based on the topological structure between multiple tasks, we extract the information from the fact description via the Transformer-HAN encoder, extract the external information from the judgment document by the structured data encoder, and then integrate them into the classifier to reduce the misjudgment of penalty prediction. Experimental results show that our model achieves significant improvements over baselines for all judgment prediction tasks.

In the future, we will seek to explore the following directions: (1) It is interesting to explore the multitask legal prediction with multiple labels and multiple defendants. In recent years, the rise of knowledge graphs and graph neural networks (GNN) has made this possible [25–28]. (2) We will explore how to incorporate various factors into LJP, such as defendant’s subjective viciousness, defendant’s criminal means, and defendant’s identity, which are not considered in this work. (3) When a judge decides a case, similar cases are crucial to the judgment result for this case. Therefore, we can also recommend similar judgment documents to judges [29–31]. (4) With more and more research on the transfer learning, GPT, Bert, and other natural language models are also produced and continuously improve the ability to extract information from the text. The use of transfer learning in the process of dealing with the fact descriptions may improve the effectiveness of models [32–34].

## 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.

## Authors’ Contributions

Kongfan Zhu and Rundong Guo contributed equally to the paper.

## Acknowledgments

This work was supported in part by the Key Research and Development Program of China under Grant no. 2018YFC0831000 and no. 2017YFC0803400.

## References

- [1] J. A. Segal, “Predicting supreme court cases probabilistically: the search and seizure cases, 1962–1981,” *American Political Science Review*, vol. 78, no. 4, pp. 891–900, 1984.
- [2] C. Liu, C. Chang, and J. Ho, “Case instance generation and refinement for case-based criminal summary judgments in Chinese,” *Journal of Information Science and Engineering*, vol. 20, no. 4, pp. 783–800, 2004.
- [3] D. M. Katz, M. J. Bommarito, and J. Blackman, “A general approach for predicting the behavior of the supreme court of the United States,” *PLoS One*, vol. 12, no. 4, 2017.
- [4] B. Luo, Y. Feng, J. Xu, X. Zhang, and D. Zhao, “Learning to predict charges for criminal cases with legal basis,” in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2727–2736, Copenhagen, Denmark, September 2017.
- [5] H. Zhong, Z. Guo, C. Tu, C. Xiao, Z. Liu, and M. Sun, “Legal judgment prediction via topological learning,” in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3540–3549, Brussels, Belgium, November 2018.
- [6] W. Yang, W. Jia, X. Zhou, and Y. Luo, “Legal judgment prediction via multi-perspective bi-feedback network,” in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence IJCAI-19*, pp. 4085–4091, Melbourne, Australia, August 2019.
- [7] W. Zaremba, I. Sutskever, and O. Vinyals, *Recurrent Neural Network Regularization*, <https://arxiv.org/abs/1409.2329>.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [9] A. Vaswani, N. Shazeer, N. Parmar et al., “Attention is all you need,” in *Advances in Neural Information Processing Systems*, pp. 5998–6008, 2017.
- [10] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, “Hierarchical attention networks for document classification,” in *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1480–1489, San Diego, CA, USA, June 2016.
- [11] Y.-H. Liu, Y.-L. Chen, and W.-L. Ho, “Predicting associated statutes for legal problems,” *Information Processing & Management*, vol. 51, no. 1, pp. 194–211, 2015.
- [12] Q. Bao, H. Zan, P. Gong, J. Chen, and Y. Xiao, “Charge prediction with legal attention,” in *Proceedings of the CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 447–458, Dunhuang, China, October 2019.
- [13] Y. Shen, J. Sun, X. Li, L. Zhang, Y. Li, and X. Shen, “Legal article-aware end-to-end memory network for charge prediction,” in *Proceedings of the 2nd International Conference on Computer Science and Application Engineering*, pp. 1–5, Hohhot, China, October 2018.
- [14] H. Chen, D. Cai, W. Dai, Z. Dai, and Y. Ding, “Charge-based prison term prediction with deep gating network,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, no. 1, pp. 6361–6366, Hong Kong, China, November 2019.
- [15] X. Jiang, H. Ye, Z. Luo, W. Chao, and W. Ma, “Interpretable rationale augmented charge prediction system,” in *Proceedings of the Coling 2018*, pp. 146–151, Santa Fe, NM, USA, August 2018.
- [16] S. Chen, P. Wang, W. Fang, X. Deng, and F. Zhang, “Learning to predict charges for judgment with legal graph,” in *Artificial Neural Networks and Machine Learning–ICANN 2019*, pp. 240–252, Springer, Berlin, Germany, 2019.
- [17] Y. S. Chen, S. W. Chiang, and T. Y. Juang, “A few-shot transfer learning approach using text-label embedding with legal attributes for law article prediction,” *EasyChair, Technical Representative*, vol. 1344, 2019.

- [18] Z. Yang, P. Wang, L. Zhang, L. Shou, and W. Xu, "A recurrent attention network for judgment prediction," in *International Conference on Artificial Neural Networks*, pp. 253–266, Springer, Berlin, Germany, 2019.
- [19] O.-M. Sulea, M. Zampieri, S. Malmasi, M. Vela, P. L. Dinu, and V. J. Genabith, *Exploring the Use of Text Classification in the Legal Domain*, ASAIL@ICAIL, London, UK, 2017.
- [20] P. Wang, Y. Fan, S. Niu, Z. Yang, and J. Guo, "Hierarchical matching network for crime classification," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 325–334, Paris, France, July 2019.
- [21] X. Li, Y. Meng, X. Sun, Q. Han, A. Yuan, and J. Li, "Is word segmentation necessary for deep learning of Chinese representations?," in *Proceedings of the Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 3242–3252, Florence, Italy, July 2019.
- [22] J. L. Ba, R. Kiros, and E. G. Hinton, *Layer Normalization*, <https://arxiv.org/abs/1607.06450>.
- [23] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," <https://arxiv.org/abs/1412.6980>.
- [24] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [25] S. Pan, R. Hu, S.-f. Fung, G. Long, J. Jiang, and C. Zhang, "Learning graph embedding with adversarial training methods," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2475–2487, 2019.
- [26] J. Shaoxiong, P. Shirui, C. Erik, M. Pekka, and P.S. YU, "A survey on knowledge graphs: representation, acquisition and applications," <https://arxiv.org/abs/2002.00388>.
- [27] S. Pan, J. Wu, X. Zhu, C. Zhang, and S. Y. Philip, "Joint structure feature exploration and regularization for multi-task graph classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 715–728, 2015.
- [28] T. Guo, S. Pan, X. Zhu, and C. Zhang, "Cfond: consensus factorization for co-clustering networked data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 4, pp. 706–719, 2018.
- [29] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4186–4199, 2019.
- [30] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 1–13, 2018.
- [31] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [32] X. Ben, P. Zhang, Z. Lai, R. Yan, X. Zhai, and W. Meng, "A general tensor representation framework for cross-view gait recognition," *Pattern Recognition*, vol. 90, pp. 87–98, 2019.
- [33] X. Ben, C. Gong, P. Zhang, R. Yan, Q. Wu, and W. Meng, "Coupled bilinear discriminant projection for cross-view gait recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 734–747, 2020.
- [34] X. Ben, C. Gong, P. Zhang, X. Jia, Q. Wu, and W. Meng, "Coupled patch alignment for matching cross-view gaits," *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 3142–3157, 2019.

## Research Article

# Influence of Social Networks on Citizens' Willingness to Participate in Social Governance: Evidence from China

Rui Nan and Fan Ouyang 

*School of Law and Humanities, China University of Mining and Technology, Beijing 100083, China*

Correspondence should be addressed to Fan Ouyang; 201123@cumtb.edu.cn

Received 24 June 2020; Revised 16 August 2020; Accepted 3 September 2020; Published 14 October 2020

Academic Editor: Fei Xiong

Copyright © 2020 Rui Nan and Fan Ouyang. 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.

Social networks are social structure constituted by a set of social actors with embedded relationships, which has a significant impact on both perceptions and behaviors among individuals and groups. The influence of the social networks on citizens' willingness to participate in social governance is manifested in two aspects: one is that social networks directly affects the citizens' willingness to participate; the other is the social capital made up by social networks, social trust, and social norms affects the citizens' willingness to participate. Drawing on a transprovincial survey regarding citizen participation in the social networks, this paper uses the Ordered Logistic model to explore how does social networks affect the citizens' willingness to participate. The results show that (1) social networks have a significant impact on citizens' willingness to participate in social governance, specifically, the stronger the social networks are, the higher level of the citizens' willingness to participate will be; (2) social networks, together with social trust and social norms constitute social capital, and the social capital has a significant impact on the citizens' willingness to participate in social governance, which is manifested; the higher the social capital stock is, the higher level of the citizens' willingness to participate will be; and (3) from the lens of demographic characteristics, those who are male, high educated, CCP members, or from the eastern region of China are more willing to participate.

## 1. Introduction

Social networks refer to the complex interpersonal networks which arise from the interaction between individual members of a society. They are social relationship system that are focused on one or some members of a society and can support their action and also a collection of various social relationships [1]. The social networks can be generally divided into informal networks and formal networks. The informal networks are a kind of interpersonal network formed through the interaction of relatives, friends, or acquaintances, showing a loose relationship not subjected to formal system constraints. The formal networks are a group pattern networks established by interest groups, nongovernmental organizations, and other institutions, which have obvious system rule constraints [2, 3]. Practice shows that no matter what kind of social networks, they will directly affect the behavior and willingness of individuals and groups.

With the advancement of industrialization and globalization, China's social groups become more diverse and subsequently create a more diversified and personalized public demands. Hence, the social problems, which are more complicated than ever, have brought new risks and uncertainties. The citizen participation has become a pivotal approach to solve social problems and reduce social risks, as well as a policy tool to absorb various interests and value demands. From the perspective of practice, at this stage, the awareness and ability of citizen participation in China have been gradually improved, and the depth and breadth of participation have been increasingly enhanced as well. However, there are still some problems, such as weak awareness of citizen participation, weak ability, lagging laws, and regulations, which not only directly restrict the effectiveness of citizen participation in social governance but also frustrate the initiative and enthusiasm of citizen participation to a certain extent. Therefore, it is of great significance to

advocate orderly and effective citizen participation in the implementation of social governance refinement and improve the level of social governance in China.

Orderly and efficient citizen participation depends on the certain social networks. The citizen participation in social governance refers to the participation in the field of social governance, which is an important form of social interaction. It refers to all activities in which the citizens widely participate in social governance and try to affect public policy and public life through feedback and interaction on the public policy making and the decision of public affairs or public governance. The citizen participation in social governance is not only an individual behavior but also a collective behavior centered on the public. Willingness refers to the motivation or tendency of an individual or a group to take a certain action, and the action is determined by willingness [4].

The citizen participation in social governance is a typical collective behavior. Social networks have an impact on the willingness of individuals and groups from two aspects. On the one hand, through intensive and effective social interaction, the social networks reduce the risk of information asymmetry, increase the cooperation among the citizens, promote the transformation of individual behavior into collective action, and enhance the citizens' willingness to participate [5]. On the other hand, the social capital refers to the structure, content, and perception of one's social relationships in the networks [6, 7]. Therefore, the social networks, together with social trust and social norms constitute social capital by strengthening cooperation, mutual benefit and trust among the citizens, increase the stock of social capital, enhance the citizens' willingness to participate, and then promote the conversion of willingness into behavior [1, 5]. Therefore, this study responds to the following questions:

- (1) Whether social networks affect the citizens' willingness to participate in social governance? If so, in what manner and to what degree?
- (2) Whether social capital affects the citizens' willingness to participate in social governance? If so, in what manner and to what degree?
- (3) Serving as control variables, how does demographic characteristics affect the citizens' willingness to participate through social networks and social capital?

In response to the questions above, the dependent variable introduced in this research is "the citizens' willingness to participate in social governance," the independent variable is the social networks and its social capital, and the control variable is the demographic characteristics. The Ordered Logistic model is used to test the influence of social networks on the citizens' willingness to participate in social governance; and the influence of social capital composed of social networks, social norms, and social trust on the citizens' willingness; how the control variable affects the citizens' willingness through social networks and social capital.

It is known that both social networks and social governance in China have Chinese characteristics. By using certain model and taking China as the sample, this paper explores the influence and causes of China's social network on citizen participation in social governance, which may provide useful reference and enlightenment for other countries in the world. This is also the academic contribution of this paper.

## 2. Related Work

There is a rich tradition in both social science and natural science research studies that have focused on the individual or collective participation, in the perspective of social networks and social capital. Comparatively speaking, it is common to study citizen participation through social capital.

Usually in the sphere of social science, it is publicly accepted that the theoretical formulations on social capital of Bourdieu, Coleman, and Putnam have greatly contributed to the basic concept, introducing and conceptualizing it into the research of social science. In Bourdieu's opinion, social capital is the resource that provides each of the members with the backing of collective-owned capital, a credential which entitled them to credit [1], which provided a classical definition of "social capital." Coleman [8] followed Bourdieu's steps, but he emphasized the social structure which circumscribes social action, and the utility-maximizing pursuit of his/her self-interest that determines actors' goals. Putnam [9] conducted research on Italy's institutional innovation, whereas he discovered that the social capital is something which could be self-accusing and cumulative so that cooperation, trust, reciprocity, collective benefit, and civic engagement would be formed and strengthened; in addition, the networks of civic engagement fostered strong reciprocity norms. Lin [10] also has made a great contribution to the basic concepts of social networks and social capital, by trying to build a network theory of social capital, and he argued that social capital is another form of neo-capital theories, which indicate access to and use of resources embedded in social networks, namely, the use of social capital by individuals. Unlike others, he discovered two conceptual elements of social capital and accordingly found out approaches in measuring social capital as assets, which expanded the individual perspective that the predecessors insisted. Obviously the basic concept they formed is used in this paper; however, the connotation of social capital and social networks in this paper is still inevitably influenced by the institutional arrangement of collectivism in People's Republic of China and the unique culture and tradition of "guanxi," which sees the individual as part of a community and a set of family, hierarchical and friendly relationships.

Based on these foundation and framework, multiple research studies have been conducted from various angles of view, such as sociology, management, political science, and economics, most of which take a certain area, or country, or enterprise as a case to study.

Some of them are still focused on the theoretical construction. Fukuyama [11] noticed that, in the Chinese parts

of East Asia and much of Latin America, social capital resides largely in families and a rather narrow circle of personal friends; therefore, social capital in those countries have a narrow radius of trust. Ostrom [12] collected both experimental and empirical evidence to find out the explanations to collective actions, which is development and growth of social norms. Glaeser et al. [13] tried to probe into an economic approach to social capital, and they found out the interpersonal externalities could be generated by social capital, which bring about the difference between social capital and other forms of capital; in addition, the social capital accumulation patterns are consistent with the standard economic investment model. Roberts and Devine [14] suggested that a major factor in the individuals' volunteering, who are involved in numerous forms of voluntary activity, is the pleasure and enjoyment they would gain from the activity itself; therefore, a constant process of dialogue within formal structures of participation is needed. Fu [15] found out that social capital and trust are mutually reinforcing, namely, social capital generates trusting relationship while in turn trust produces social capital. Those help our research design to take notice on the difference among regions and the difference between rural and urban areas.

In recent year, it is more popular to take a certain area, or country, or enterprise as the case to probe into. Brewer [16] conducted a research based on the data from 1996 American National Election Study, comparing public servants and other citizens with regard to attitudes and behaviors related to social capital, with the conclusion that public servants are far more active in civic affairs than other citizens. Skoric et al. [17] followed Putnam's pattern, trying to examine whether the social networks on the Internet could pave the paths for revival of political participation both online and offline in Singapore. Torri [18] examines GMCL, an Indian community-based enterprise led by women which is formed by a network of self-help groups, arguing that group and social forms of entrepreneurship should not become the paradigm in developmental policies for women. Fung and Hung [19] explored developing sustaining social capital for community development based on their case study on the Tin Shui Wai North in Hongkong. Asteria et al. [20] discovered by a quantitative approach that puts efforts to increase women's participation in urban planning and environmental management require the support of social capital. Suh and Heidi [21] revealed a positive association between outgroup trust and protest participation, moderated by both functioning institutions and state repression.

The above studies provide different social science' perspectives to discuss social networks and citizen participation; however, they are mainly based on the Western social background, which may not adequate to explain the reality of social networks in China and how it influences the citizens' willingness to participate and what measures could be taken to foster and reinforce the social networks so that the social capital could grow.

Meanwhile, the concepts of social capital, participation, social networks, social norms, social recommendation, etc., also began to attract the attention of the sphere of natural science. Zúñiga's et al. [22] research disclosed that

information use of social networks sites exerted a significant and positive impact on individuals' activities aimed at engaging in civic and political action. Hu et al. [23] introduced evolutionary opinion dynamics from the field of statistical physics into the recommender system, measuring user influence according to their topological role in the social networks, in order to predict the unknown ratings. Xiong et al. [24] also conducted experiments on real-world datasets to compare the proposed method with the state-of-the-art models, and the results indicate that the social recommendation method makes notable improvements in rating prediction. Li et al. [25] found the importance of users' ratings and determine their influence on entire social networks analyzing the topology of social networks to investigate users' influence strength on their neighbors. The above research studies provide a technical insight for our research and an adequate approach to prove and support our research.

### 3. Data and Model

*3.1. Theoretical Analysis and Hypothesis.* To some extent, the interpersonal networks formed by social networks are both a valuable resource and a media. The stronger the social networks are, the more rapid and effective the information transmission and diffusion are. They can reduce the information bias in the citizen participation in social governance, effectively alleviate moral hazard, reduce the transaction cost of social interaction, promote the formation of orderly cooperative behavior, and thus enhance the citizens' willingness to participate in social governance. In addition, social networks, together with social trust and social norms, constitute social capital. It is not only a sort of resource, "social structure resource possessed by individuals," or "a resource embedded in a social structure that could be acquired or mobilized in the purposeful actions," whereas others consider social capital as a kind of ability, "the ability that individuals maintains precious resources in the networks or in a much wider social structure, through the membership" [5, 8, 10]. The social capital, through the coordination of social networks, restriction of social norms, and the support of social trust, promotes the public's collective action, which helps to improve the public's rule obedience, and promotes the public's sense of self-interest to public welfare which effectively solves the collective action dilemma, and improves the citizens' willingness to participate in social governance [26, 27]. Among them, "trust is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently of his capacity ever to be able to monitor it) and in a context in which it affects his own action" [28]; norms are "shared understandings about actions that are obligatory, permitted, or forbidden" [29]; networks are the carrier of social capital, postulating "that human behavior is embedded in a network of interpersonal relations" [30]. Based on these, the following research hypotheses are proposed:

$H_1$ : the stronger the social networks are, the stronger the citizens' willingness to participate in social governance is.

$H_2$ : the higher stock of social capital is, the stronger the citizens' willingness to participate in social governance is.

$H_{2a}$ : the higher level of social trust is, the stronger the citizens' willingness to participate in social governance is.

$H_{2b}$ : the higher level of social norms is, the stronger the citizens' willingness to participate in social governance is.

Based on previous studies, the citizens' willingness to participate in social governance can be affected by the demographic characteristics, including gender, age, education, income, and region [31–33]. Therefore, the following hypothesis is formulated:

$H_3$ : demographic characteristics (gender, age, education, political status, income, and region) strongly affect the willingness of citizen participation in social governance.

**3.2. Data Source and Validation.** The research utilized the Stratified Sampling method to collect the first-hand data. The stratification of this survey was conducted firstly on the basis of the division of three economic belts (Eastern, central, and Western) in China (five provinces from each), and then the sample was drawn according to the population proportion and the sex ratio (roughly equivalent) (according to the level of economic development and geographical location, China can be divided into three regions). On the basis of the three economic belts, the first layer comprised of five more developed provinces or municipalities in Eastern China: Beijing, Guangdong, Zhejiang, Shandong, and Fujian; the second layer comprised of five medium-developed provinces in Central China: Hubei, Anhui, Henan, Hunan, and Jilin; the third layer was comprised of five underdeveloped provinces or autonomous regions in Western China: Yunnan, Shaanxi, Sichuan, Inner Mongolia, and Gansu. The survey was conducted both online and offline, among which the online survey covered all the 15 provinces, municipalities or autonomous regions, while the offline survey covered Beijing, Guangdong, and Zhejiang in the Eastern region, Hubei and Anhui in the Central region, and Yunnan and Shaanxi in the Western region. There were 773 valid questionnaires out of the 1037 questionnaires (74.54%) obtained from online survey, while there were 232 valid questionnaires out of the amount of 250 questionnaires (92.8%) obtained from offline survey.

In order to ensure the reliability and validity of the survey, the respondents were limited to 18–65 years old, and Stratified Sampling method was adopted in the survey process. The test results were shown that the Cronbach Alpha Coefficient is 0.891, which means the internal consistency is good; while the KMO = 0.783, and the significance level of Bartlett's test of sphericity = 0.000, indicating the survey is of high validity.

### 3.3. Variable Designing and Measurement

**3.3.1. Dependent Variables.** The dependent variable is “the citizens' willingness to participate in social governance (willingness),” measured mainly by asking the respondents the question “your willingness of participating in social governance.” The respondents could choose among the four options, “not willing at all,” “willing to participate in all the available activities,” “willing to participate in the activities of easy access,” and “willing to participate in the activities of easy access, with self-interests involved,” which was an ordinal categorical variable, on a scale from 1 to 4.

**3.3.2. Independent Variables.** The independent variable is the social networks and its social capital. Social capital was divided into three parts, social networks, social trust, and social norms, according to Putnam's definition, and measured by the three parts as well. They were, respectively, marked as below: *networks, capital, trust, and norms*. The scale for measuring these variables is designed according to the characteristics of China's social capital and social networks based on the previous research [34–38]. The scale focuses on the measurement of community microsocial capital and social networks, focusing on the impact of collective participation behavior from the macroperspective.

As for the measurement of social networks, there were also two questions: Q1 was “how many neighbors do you greet frequently in your community (village),” and the options were “<30,” “30–39,” “40–49,” “50–60,” and “≥60,” on a scale from 1 to 5; Q2 was “how many neighbors are considered to be your friends in your community (village),” and the options were “<5,” “5–7,” “8–10,” “11–15,” and “>15,” on a scale from 1 to 5.

Likert Scale was used to measure “social trust,” mainly on the levels of trust of plural participants in social governance; then, four questions were set up, about the levels of trust on the four kinds of participants' participation in social governance, that is, the government (or CPC committee), social organizations, self-organizations, and other publics. There were five options: “completely trust,” “somewhat trust,” “moderately trust,” “somewhat distrust,” and “distrust completely,” on a scale from 1 to 5.

Concerning “social norms,” two questions were set up based on the previous research studies: Q1 was “how difficult do you think it is to acquire the relevant information on social governance?,” along with five options that were “completely difficult,” “somewhat difficult,” “moderately difficult,” “somewhat easy,” and “completely easy,” on a scale from 1 to 5; Q2 was “to what degree do you understand the overall policy and regulations (or institutions) on social governance?,” along with five options, which were “completely do not understand,” “somewhat do not understand,” “moderately understand,” “somewhat understand,” and “completely understand,” on a scale from 1 to 5 (each dimension is set with 4–5 items, and only items with passing factor load (at least greater than 0.4) is listed here).

**3.3.3. Control Variables.** The control variables are gender, age, education, political status, occupation, income, urban or rural status, and region, respectively, marked as *gender, age,*

education, CPC membership, occupation, income, urban or rural, and region. The variable' scaling and statistical description are as shown in Table 1.

**3.4. Model.** According to the practices, the citizens' willingness to participate in social governance could be divided into four categories: "no willing at all," "willing to participate in all the available activities," "willing to participate in the activities of easy access," and "willing to participate in the activities of easy access, with self-interests involved," which are typical ordinal categorical variables with clear stratifications; therefore, the Ordinal Logistic model was used as the primary empirical model in this research. The logic diagram is shown in Figure 1.

The specific process is as follows [39]:

$$y_i^* = \beta_i x_i + \mu_i. \quad (1)$$

In this model,  $i$  refers to the  $i$ th specific citizen;  $y_i^*$ , an unobservable variable, refers to the different levels of willingness of citizen participation in social governance;  $x_i$  is the vector of a group of explanatory variables which might influence the citizens' willingness to participate in social governance;  $\beta_i$  is the coefficient of correspondence to the explanatory variables; and  $\mu_i$  is the random error term.

There are four categorical levels on the citizens' willingness to participate in social governance, as the following:  $y_i = 1$  refers to "no willing at all,"  $y_i = 2$  refer to "willing to participate in all the available activities,"  $y_i = 3$  refers to "willing to participate in the activities of easy access," and  $y_i = 4$  refers to "willing to participate in the activities of easy access, with self-interests involved." There are three cutoff points (thresholds) as below:

$$y_i = \begin{cases} 1, & y_i \leq \mu_1, \\ 2, & \mu_1 < y_i \leq \mu_2, \\ 3, & \mu_2 < y_i \leq \mu_3, \\ 4, & y_i > \mu_3. \end{cases} \quad (2)$$

Assuming  $\mu_i$  is subject to the *Logistic* distribution, and the distribution function is  $F_{(x)}$ , that is,

$$\begin{cases} p_r(y_i = 1 | x_i) = 1 - F_{(\beta_i x_i)}, \\ p_r(y_i = 2 | x_i) = F_{(\mu_1 - \beta_i x_i)}, \\ p_r(y_i = 3 | x_i) = F_{(\mu_2 - \beta_i x_i)}, \\ p_r(y_i = 4 | x_i) = 1 - F_{(\mu_3 - \beta_i x_i)}, \end{cases} \quad (3)$$

$$F_{(\beta_i x_i)} = \frac{\exp(\beta x)}{1 + \exp(\beta x)}.$$

In order to better analyze the alteration of the level of the citizens' willingness to participate in social governance, the marginal effect was also adopted to analyze the degree of influence on the citizens' willingness to participate so that the change of social networks and social capital has

$$\begin{cases} \frac{\partial p_r(y_i = 1 | x_i)}{\partial x_i} = -f_{(-\beta_i x_i)} \beta_i, \\ \frac{\partial p_r(y_i = 2 | x_i)}{\partial x_i} = f_{(\mu_1 - \beta_i x_i)} - f_{(-\beta_i x_i)} \beta_i, \\ \frac{\partial p_r(y_i = 3 | x_i)}{\partial x_i} = f_{(\mu_2 - \beta_i x_i)} - f_{(-\beta_i x_i)} \beta_i, \\ \frac{\partial p_r(y_i = 4 | x_i)}{\partial x_i} = f_{(\mu_3 - \beta_i x_i)} \beta_i. \end{cases} \quad (4)$$

The measurement of social capital involves many indexes, so it is necessary to make factor analysis. The specific process is as follows.

Let social capital have  $m$  common factors, namely,  $f_1, f_2, \dots, f_m$ . General assumption  $X_1, X_2, \dots, X_p$  is the sum of these common factors and their special factors  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$  composition:

$$X_i = a_{i1} f_1 + a_{i2} f_2 + \dots + a_{im} f_m + \varepsilon_i, \quad (i = 1, 2, \dots, p, m \leq p). \quad (5)$$

It is expressed in the matrix form as follows:

$$X = AF + \varepsilon, \quad (6)$$

where

$$\begin{aligned} X &= \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_p \end{bmatrix}, \\ F &= \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_p \end{bmatrix}, \\ A &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pm} \end{bmatrix}, \\ \varepsilon &= \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_p \end{bmatrix}. \end{aligned} \quad (7)$$

$E(X) = 0$ ,  $\text{cov}(X, X) = R$ ,  $R$  is the correlation coefficient matrix,  $E(F) = 0$ , and  $\text{cov}(F, F) = I_m$  ( $M$ -order unit matrix), that is, there is no correlation between the common factors and the variance is 1.

$$E(\varepsilon) = 0, \text{cov}(\varepsilon, \varepsilon) = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sigma_p^2 \end{bmatrix} = \emptyset, \text{ that is, each}$$

special factor is not related, and the variance does not need to be equal.



TABLE 1: Ordinal logistic model variable scaling and statistics.

Category	Variables	Scaling	Ratio (%)
Dependent	Willingness of participate	1 = no willing at all	36.82
		2 = willing to participate in all the available activities	19.50
		3 = willing to participate in the activities of easy access	35.22
		4 = willing to participate in the activities of easy access, with self-interests involved	8.46
Independent	Social networks	PCS <sup>1</sup>	—
	Social capital	PCS	—
	Social trust	PCS	—
	Social norms	PCS	—
Control	Gender	0 = male	53.33
		1 = female	46.67
	Age	1 = 18–29	42.49
		2 = 30–39	29.35
		3 = 40–49	21.89
		4 ≥ 50	6.27
	Education	1 = junior secondary school or lower	20.50
		2 = senior secondary school or specialized secondary school	27.06
		3 = short-cycle courses in HEIs or above	52.44
	CPC membership	0 = No	32.44
		1 = Yes	67.56
	Occupation	1 = government, public institutions, and state-owned enterprises	27.16
		2 = private enterprises	39.20
		3 = farming	11.64
		4 = others	21.99
		1 ≤ 3000 yuan	36.72
		2 = 3000–4500 yuan	22.89
		3 = 4501–9000 yuan	26.47
		4 ≥ 9000 yuan	13.93
	Income (CNY)	0 = urban	70.55
1 = rural		29.45	
Urban or rural	1 = eastern	43.38	
	2 = central	36.02	
Region	2 = central	36.02	
	3 = western	20.60	

<sup>1</sup> PCS was the main component factor score

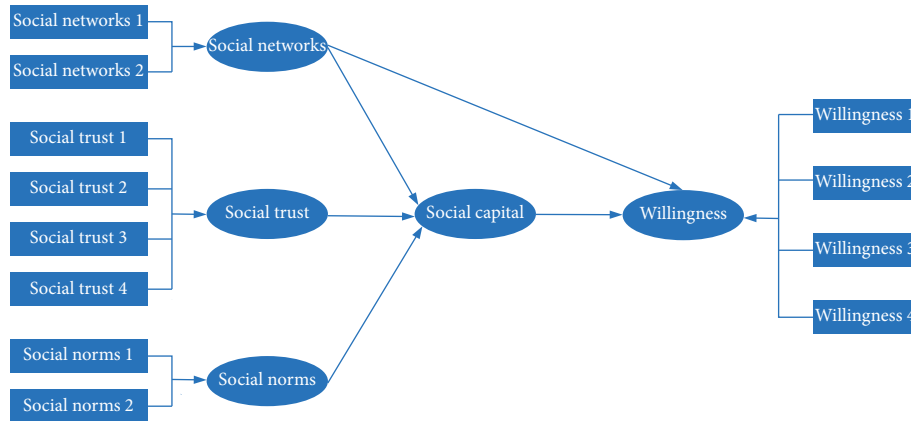


FIGURE 1: Model logic diagram.

$\text{cov}(F, \varepsilon) = 0$ , that is, the common factor and the special factor are not related.  $X$  is an observable random vector, while  $F$  and  $\varepsilon$  are unobservable random vectors.

In the factor analysis model, matrix  $A$  is called factor load matrix. Record  $A = (a_{ij})_{p \times m}$ :

$$\begin{aligned}
 \text{cov}(X_i, f_j) &= \text{cov}(a_{i1}f_1 + a_{i2}f_2 + \cdots + a_{im}f_m + \varepsilon_i, f_j) \\
 &= \text{cov}(a_{ij}f_j, f_j) = a_{ij}\text{cov}(f_j, f_j) \\
 &= a_{ij}, \quad (i = 1, 2, \dots, p, j = 1, 2, \dots, m),
 \end{aligned}$$

(8)

where  $a_{ij}$  is the correlation coefficient between variable  $X_i$  and common factor  $f_j$ , i.e., factor load. It is the load of the  $i$  variable  $X_i$  on the  $j$  factor  $f_j$ .

#### 4. Results and Analysis

Based on Sherry Arnstein's "Ladder of Citizen Participation" [40], along with the practice of social governance in China, the citizens' willingness to participate in social governance was divided into three types according to the levels of participation intensity (Table 2).

It is shown in Table 2 and Figure 2 that, according to the levels of participation intensity, the willingness to participate in social governance could be divided into three types which are strong participation, weak participation, and limited participation. Level 1 and level 2, respectively, corresponded to the willingness of "no willing at all" and "willing to participate in all the available activities," with the participation degrees of extremely weak and extremely strong, which were put into the two types of "weak participation" and "strong participation." In the middle, there were level 3 and level 4, respectively, corresponding to the willingness of "willing to participate in the activities of easy access," and "willing to participate in the activities of easy access, with self-interests involved," with the degrees of "relatively strong" and "relatively weak," which were put into the type of "limited participation." Concerning the ratio, the type of strong participation possessed the least proportion of 19.5%; the ratio of the public of limited participation was relatively higher, which was 43.68%; and the ratio of weak participation reached as high as 36.82%, which indicated the overall level of the citizens' willingness to participate in social governance is still on the low side in China.

*4.1. Descriptive Statistics.* It is obvious in Table 3 that, from the nationwide perspective, on the citizens' willingness to participate in social governance, 36.82% of the respondents are not willing to participate in and 63.18% of them are willing to participate in, among which 19.50% of them have strong willingness of participation, namely, intending to participate in all the available activities; 35.22% of them willing to participate in the activities of easy access, and 8.46% of them intend to willing to participate in the activities of easy access, with self-interests involved. In the Eastern region in China, the proportion of respondents who have "no willingness at all" fell to around 20% (23.39%), while that of respondents approximates to (or exceeds) 50% (45.03% and 50.72%) in Central and Western regions. It indicates that, the nationwide willingness of citizen participation in social governance is relatively strong, though more falling into the limited-participation type, which including the two willingness types of "willing to participate in the activities of easy access" and "willing to participate in the activities of easy access, with self-interests involved." From the comparison of three regions, the citizens' willingness to participate in social governance in the Eastern region is

relatively stronger, the central region takes the second place, and the Western region is relatively weaker; in addition, there are more limited-participation citizens in the Eastern region, with their willingness mainly displayed as limited participation.

It is shown in Figure 3 as follows:

- (1) Gender and willingness to participate: the ratio of female respondents who are not willing to participate in social governance is higher than that of male, and the ratio of male is higher than female on the three levels of relatively stronger willingness of participation, especially on the "categorical level = 2," and the participation willingness of males is apparently higher than that of females. It suggests that the male's willingness of participation in social governance is higher than that of female.
- (2) Age and willingness to participate: the ratio of the respondents above 40 years old (middle aged) who are not willing to participate in social governance is relatively higher, while the younger respondents make up a smaller proportion. On the three different levels of willingness, the ratios of respondents from different age groups are roughly equal, without apparent disparity, though they are still concentrated on the latter two levels, namely, the limited participation, which indicates that, comparatively, the participation willingness of middle-aged citizens is weaker than that of young citizens; however, among the citizens from different age groups, there is little difference on the three levels of willingness to participate in China and more on the limited-participation type (for the division of middle-aged and young people, this paper adopts the standard of all China Youth Federation and considers that 18–40 years old are young people).
- (3) Education and willingness to participate: The ratio of the respondents with college degree or above (well educated) with no willing to participate in social governance is well below the other two categories, and the ratio of the respondents with degrees of senior secondary school or lower is relatively higher. To be more specific, for the well-educated respondents, the ratio of citizens who are "willing to participate in the activities of easy access" accounts for a higher proportion, whereas the ratio of the citizens who are "willing to participate in the activities of easy access, with self-interests involved" is relatively insensitive to the education factor, which indicated that, relatively, the citizens with higher education have stronger willingness to participate in social governance in China.
- (4) Political status and willingness to participate: the ratio of the citizens who are CPC members is lower than the non-CPC members on the "no willing at all" level, and the CPC members' participation willingness relatively concentrates on the "categorical level = 3," which suggests the willingness of

TABLE 2: Types of willingness to participate in social governance.

Categorical level	Willingness	Types of willingness to participate
1	No willing at all	Weak participation
2	Willing to participate in all the available activities	Strong participation
3	Willing to participate in the activities of easy access	Limited participation
4	Willing to participate in the activities of easy access, with self-interests involved	

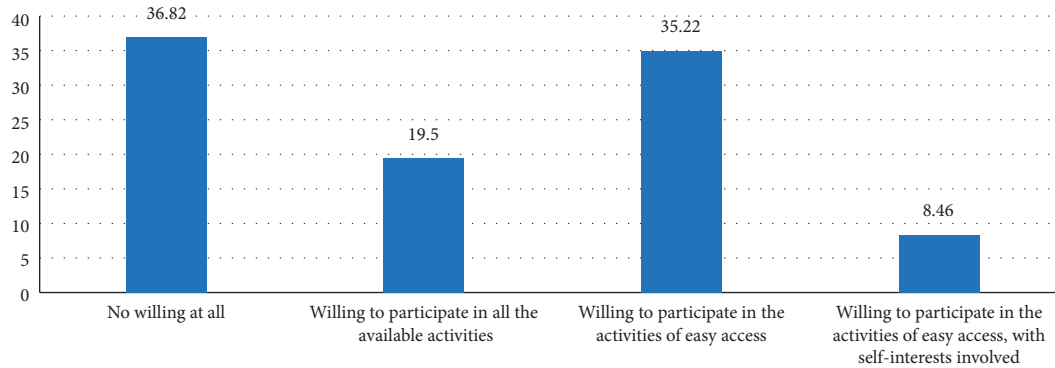


FIGURE 2: Statistical analysis of willingness in social governance.

TABLE 3: Analysis on the citizens' willingness to participate in social governance.

Region	Gross sample	Ratio of the citizens from different categorical levels (%)			
		1	2	3	4
National	1005	36.82	19.50	35.22	8.46
Eastern	436	23.39	9.86	43.81	22.94
Central	362	45.03	6.91	29.56	18.51
Western	207	50.72	8.21	27.05	14.01

participation in social governance of the citizens with CPC membership is stronger than the citizens without CPC membership.

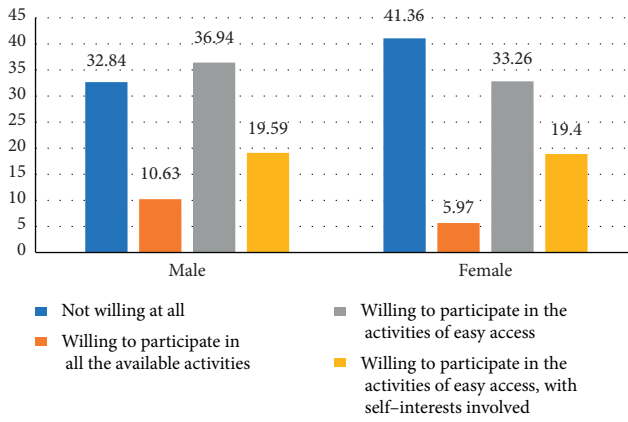
- (5) Occupation and willingness to participate: the respondents in farming account for a relatively higher proportion (49.57%) on the level "no willing at all," whereas the participation willingness of the respondents within the system (government, public institutions, and state-owned enterprises) is relatively stronger (38.83% and 21.61%). It reveals that the citizens within the system has comparatively stronger willingness of participation in social governance, that of the citizens working in private enterprises ranks the next, and that of the citizens in farming is relatively lower.
- (6) Income and willingness to participate: the proportion of "no willing at all" among the respondents of high-income (>9000) group and low-income (<3000) groups is relatively larger, whereas the proportion among the respondents of middle income is relatively small. All the groups concentrate on the "categorical level=3," while middle-income group accounts for the largest proportion on the "categorical level=2," which reveals that the citizens

from middle-income group have stronger willingness of participation in social governance; moreover, the degree of their willingness is relatively higher (in this paper, according to the per capita M2 (9360.5 yuan) and per capita GDP (4493.5 yuan) of China in 2016, the personal income level is divided into low-income ( $I < 3000$  yuan), low-income (3000 yuan  $< I < 4500$  yuan), middle-income (4500 yuan  $< I < 9000$  yuan), and high-income ( $I > 9000$  yuan), in which  $I$  is the average monthly income).

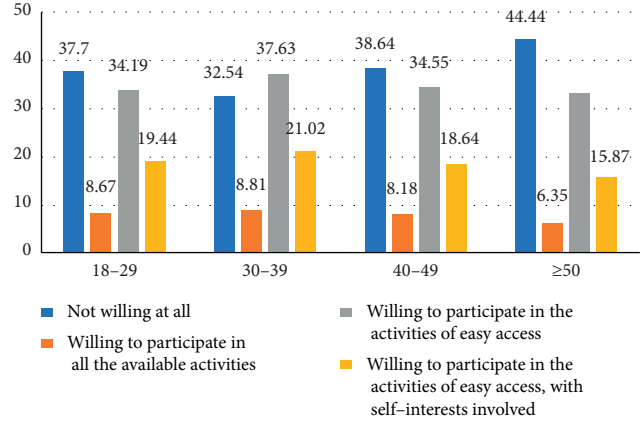
- (7) Urban or rural status and willingness to participate: the respondents living in rural areas account for a relatively higher ratio of the level of "no willing at all," while the urban respondents concentrating on the latter two levels, namely, limited-participation. On the "categorical level=2," the ratio of the rural respondents is higher than that of the urban respondents. As a whole, the willingness of the citizens in the urban areas is relatively stronger, yet the willingness of the citizens in the rural areas is stronger than that of the citizens in the urban areas when it comes to the type of "willing to participate in the activities of easy access."

**4.2. Analysis on Econometric Results.** In general, descriptive statistics is a kind of data analysis tools, direct and simple, and easy to operate. However, it is difficult for descriptive statistics to describe the correlations of multiple variables, which should be analyzed through a certain econometric model, taking the results of descriptive statistics as evidence.

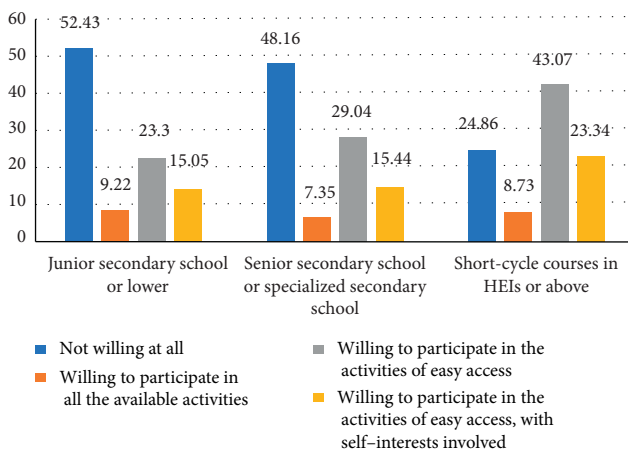
**4.2.1. Measurement of Social Capital.** There are many indicators involved in the measurement of social capital, so conducting factor analysis would be necessary. The result



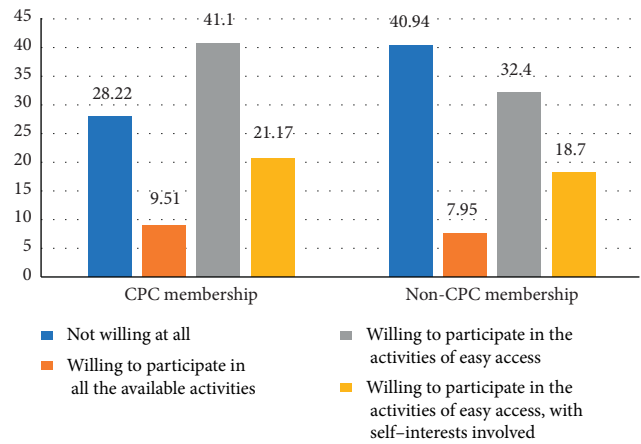
(a)



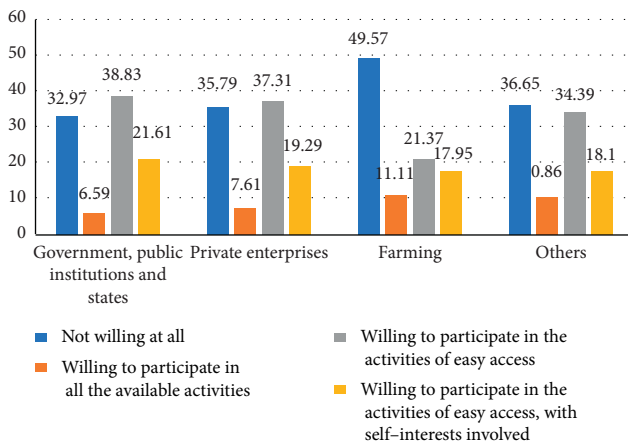
(b)



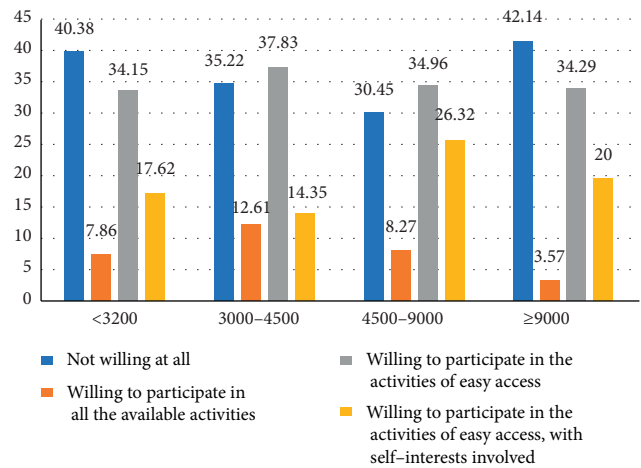
(c)



(d)



(e)



(f)

FIGURE 3: Continued.

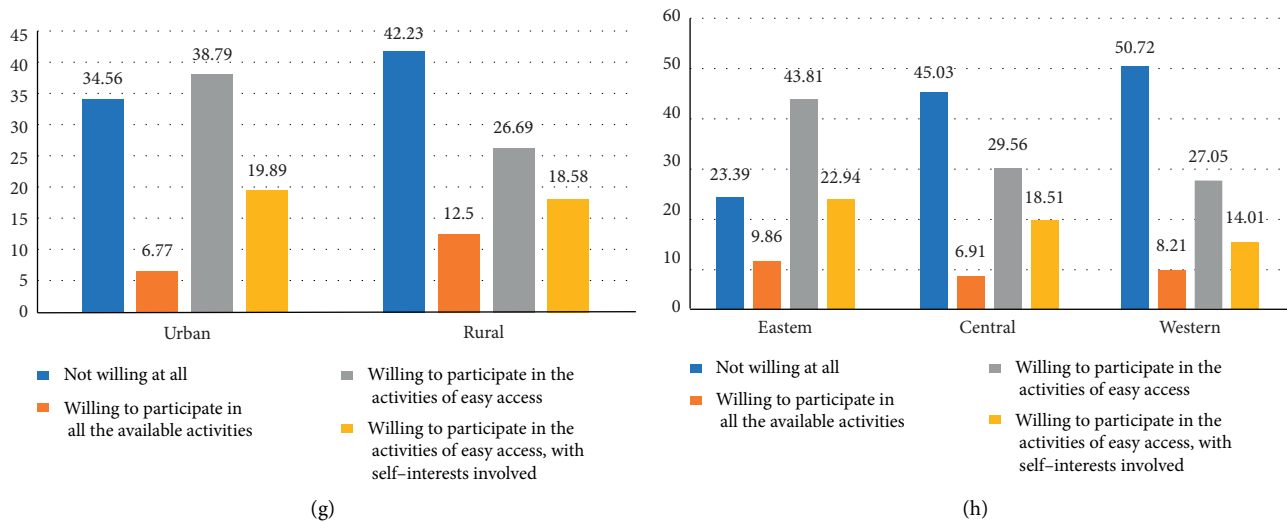


FIGURE 3: Analysis of the citizens' willingness to participate in social governance based on multifactors. (a) Statistics of citizen participation willingness of different genders. (b) Statistics of citizen participation willingness of different ages. (c) Statistics of citizen participation willingness of different education backgrounds. (d) Statistics of citizen participation willingness of different political situations. (e) Statistics of citizen participation willingness of different occupations. (f) Statistics of citizen participation willingness of different incomes. (g) Statistics of citizen participation willingness of urban and rural. (h) Statistics of citizen participation willingness of different regions.

of adaptability test shows that the KMO value of the sample test is 0.783 (greater than 0.6), indicating that it is adequate to build the factor analysis model, and the significance level of Bartlett's test of sphericity to the samples is 0.000 ( $<0.01$ ), indicating the correlation between the variables is significant, so as to be suitable for factor analysis. Orthogonal rotation was conducted by the varimax method, in accordance with the principle of eigen values  $> 1$ , and the first three Eigen factor values are all greater than 1.00, and the cumulative variance contribution rate amounts to 72.563%, greater than 70%; then, three common factors could be extracted, and the details are shown in Table 4:

It is suggested in Table 4 that the first common factor has great factor loadings on "trust on the government," "trust on the social organizations," "trust on NGOs," and "trust on the public," concentrating on the dimension of social trust, so it is called "social trust factor;" the second common factor has great factor loadings on "how many neighbors do you greet frequently in your community (village)" and "how many neighbors that are considered to be your friends in your community (village)," concentrating on the dimension of social networks, so it is called "social network factor;" and the third common factor has great factor loadings on "how difficult do you think it is to acquire the relevant information on social governance?" and "to what degree do you understand the overall policy and regulations (or institutions) on social governance?," concentrating on the dimension of social norms, so it is called "social norm factor." As a whole, the result is compatible with the theoretical hypothesis of this research, namely, the three-dimension interpretation of social capital, and it also demonstrates the adequate structural validity of the questionnaire.

**4.2.2. Analysis of the Influence of Social Networks and Its Social Capital on the Citizens' Willingness to Participate in Social Governance.** In order to better test the robustness of the model result, OLS was introduced to conduct comparative analysis while estimating [41], on the basis of the consistency on the direction of parameter estimation and significance between OLS and the Maximum Likelihood Estimate of Ordinal Logistic. The specifics are shown in Table 5.

The Nested Model was used to verify the research hypothesis: model 1 and model 3 are in the same group, comprised of control variables and dependent variables—social networks, social trust, and social norms—and their influence on the citizens' willingness to participate in social governance was analyzed by the method of Maximum Likelihood Estimate based on OLS and Ordinal Logistic. Likewise, model 2 and model 4 are in the same group, comprised of control variable and dependent variable—social capital—and the influence of the whole of social capital including social networks on the citizens' willingness to participate in social governance was analyzed by the method of Maximum Likelihood Estimate based on OLS and Ordinal Logistic, in order that the impact of the social capital's endogeneity problem could be excluded.

In general, with the significance level of 5%, the results of either Ordinal Logistic estimation, or OLS estimation, indicate that, the social networks and social capital have significant influence on the citizens' willingness to participate in social governance, so do the demographic characteristics "gender," "education," "political status," and "region," whereas the other characteristics "age," "occupation," "income," and "urban or rural status" do not.

Since the results of parameter estimation in the ordinal logistic model could not clarify the direction and degree of influence between variables at different categorical levels, but

TABLE 4: Result of factor analysis on social capital.

Measurement item	Principal factor		
	1	2	3
How many neighbors do you greet frequently in your community (village)	-0.047	0.847	-0.060
How many neighbors that are considered to be your friends in your community (village)	0.011	0.839	-0.078
How difficult do you think it is to acquire the relevant information on social governance?	-0.019	-0.015	0.831
To what degree do you understand the overall policy and regulations (or institutions) on social governance?	0.000	-0.123	0.804
Trust on the government	0.834	-0.023	0.016
Trust on the social organizations	0.894	-0.007	-0.035
Trust on the NGOs	0.878	0.004	-0.018
Trust on other publics	0.866	-0.033	0.002
Eigen values	3.018	1.438	1.349
% of variance	37.727	17.978	16.859
Cumulative % of variance	37.727	55.705	72.563

Method of extraction: PCA. Rotation: Kaiser-Varimax rotation.

TABLE 5: Regression analysis results of social networks, social capital, and demographic characteristics on the citizens' willingness to participate in social governance.

Variables	Ordinal logistic model				OLS model	
	(1)	(2)	(3)	(4)	(5)	(6)
Gender	-0.4463* (0.1197)	-0.4855* (0.1207)	-0.4324* (0.1218)	-0.2358* (0.0618)	-0.2473* (0.0612)	-0.2177* (0.0613)
Age	-0.0395 (0.0640)	-0.0620 (0.0646)	-0.0701 (0.0648)	-0.0212 (0.0331)	-0.0311 (0.0328)	-0.0348 (0.0326)
Education	0.4832* (0.0781)	0.4606* (0.0786)	0.4828* (0.0792)	0.2397* (0.0397)	0.2239* (0.0394)	0.2335* (0.0392)
Political status	0.4031* (0.1282)	0.3691* (0.1287)	0.3973* (0.1294)	0.2125* (0.0656)	0.1904* (0.0653)	0.2041* (0.0654)
Occupation	0.0370 (0.0601)	0.0538 (0.0604)	0.0500 (0.0608)	0.0282 (0.0309)	0.0372 (0.0306)	0.0351 (0.0305)
Income	-0.0740 (0.0572)	-0.0773 (0.0574)	-0.0879 (0.0577)	-0.0395 (0.0291)	-0.0416 (0.0288)	-0.0461 (0.0290)
Urban or rural	-0.0126 (0.1451)	-0.0566 (0.1463)	-0.1192 (0.1478)	-0.0036 (0.0763)	-0.0250 (0.0753)	-0.0608 (0.0748)
Region	-0.5001* (0.0800)	-0.5065* (0.0804)	-0.5143* (0.0808)	-0.2543* (0.0417)	-0.2545* (0.0413)	-0.2576* (0.0412)
Networks	—	—	0.3041* (0.0624)	—	—	0.1555* (0.0328)
Trust	—	—	0.1578** (0.0628)	—	—	0.0692** (0.0330)
Norms	—	—	0.1495** (0.0599)	—	—	0.0798* (0.0305)
Capital	—	0.5026* (0.1005)	—	—	0.2426* (0.0509)	—
_cons	—	—	—	2.1500* (0.1899)	2.2114* (0.1872)	2.2086* (0.1872)
$F/x^2$	108.64	134.19	145.62	15.32	17.46	16.38
$Prob$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$R^2/Pseudo R^2$	0.0428	0.0529	0.0574	0.0983	0.1197	0.1314

Note. \* and \*\*, respectively, represent the significance levels of 1% and 5%; data within () is robust standard errors.

only give limited information on significance and symbols, this research further estimates the marginal effect of each variable, in order to directly identify each variable's influence on the citizens' willingness to participate in social governance, as shown in Table 6.

Based on the combination of Tables 5 and 6, the empirical conclusion is drawn as follows:

- (1) *The influence of social networks on the citizens' willingness to participate in social governance.* In general, there is also a significant positive correlation between the social networks and the citizens' willingness to participate in social governance, namely, the more abundant the social networks are, the stronger the willingness of citizen participation in

TABLE 6: Marginal effect of the ordinal logistic model.

Variables	Categorical level = 1		Categorical level = 2		Categorical level = 3		Categorical level = 4	
	Margin	Std. err.	Margin	Std. err.	Margin	Std. err.	Margin	Std. err.
Gender	0.0997*	0.0281	0.0050***	0.0030	-0.0771*	0.0218	-0.0276*	0.0080
Age	0.0162	0.0149	0.0010	0.0010	-0.0126	0.0116	-0.0045	0.0042
Education	-0.1112*	0.0184	-0.0063***	0.0033	0.0865*	0.0148	0.0310*	0.0057
Political status	-0.0895*	0.0284	-0.0080***	0.0043	0.0702*	0.0227	0.0272*	0.0097
Occupation	-0.0115	0.0140	-0.0007	0.0009	0.0090	0.0109	0.0032	0.0039
Income	0.0202	0.0133	0.0012	0.0010	-0.01575	0.0104	-0.0056	0.0037
Urban or rural	0.0277	0.0345	0.0013	0.0014	-0.0214	0.0266	-0.0075	0.0091
Region	0.1185*	0.0188	0.0068**	0.0035	-0.0922*	0.0152	-0.0330*	0.0058
Networks	-0.0701*	0.0143	-0.0040***	0.0022	0.0545*	0.0115	0.0195*	0.0043
Trust	-0.0364**	0.0145	-0.0021	0.0013	0.0283**	0.0113	0.0101**	0.0041
Norms	-0.0344**	0.0138	-0.0020	0.0012	0.0268**	0.0108	0.0096**	0.0039
Capital	-0.1163*	0.0233	-0.0058***	0.0034	0.0896*	0.01844	0.0326*	0.0071

\*, \*\*, and \*\*\*, respectively, represent the significance levels of 1%, 5%, and 10%.

social governance is. The reasons are as follows: social networks are a kind of intangible resources, the citizens who possess abundant social networks could easily maintain and timely share the information on social governance through social networks; as a result, frequent social interactions may boost the information exchange among the citizens, so as to mitigate the risk of asymmetric information, reduce the cost of participation at a certain degree, and to realize individual development and improvement of life through enriched social networks; thus, the citizens “make a profit by utilizing some sort of resources owned by others” [5], and then the willingness of citizen participation will be promoted. Under such circumstances, the more enriched the social networks are, the more information that the citizens could obtain from social governance and the higher the efficiency will be; in addition, if the social networks could be utilized effectively that the public could benefit from them, then the willingness of the citizen participation in social governance will be stronger.

On all the different categorical levels, the influence of social networks on the citizens’ willingness to participate in social governance is relatively significant. On “categorical level = 1” and “categorical level = 2,” they are displayed as negative correlations, whereas on “categorical level = 3” and “categorical level = 4,” they are displayed as positive correlations. To the public of the first three categorical levels (1, 2, 3), social networks promote the willingness of the citizen participation step-by-step, which further proves that the willingness of citizen participation will be promoted as long as the degree of social networks increases; therefore, the willingness relies on the maturity of the social networks to a relatively great degree. However, on “categorical level = 4,” the influence slightly drops off, which might be because of the “interest” factor which is usually valued more than other factors to the interest-driven citizens when making decisions.

(2) *The influence of social capital on the citizens’ willingness to participate in social governance.* In general, there is a significant positive correlation between the social capital and the citizens’ willingness to participate in social governance, namely, the more the stock of social capital is, the stronger the willingness of citizen participation in social governance is. It is mostly because, the citizen participation in social governance in China is still in its primary stage, which means the degree of participation is relatively limited; the willingness of citizen participation in social governance will be apparently affected by the cost-benefit factor to a great extent, so the more the stock of social capital is, the less the transaction costs for the citizens to participate in social governance will be, the lower the costs of participation will be, the less the risk of asymmetric information will be, and the greater probability of collective action will be, which will accelerate the process of turning willingness of participation into action; therefore, the willingness of citizen participation in social governance is largely reliable on the stock of social capital.

On different categorical levels, the influence of social capital on the citizens’ willingness to participate in social governance shows a tendency of gradually strengthening: on “categorical level = 1” and “categorical level = 2,” the influence of social capital on the willingness of citizen participation in social governance is a negative correlation, while on “categorical level = 3” and “categorical level = 4,” there is a positive correlation. To the citizens on the first three categorical levels, the influence of social capital on the willingness of participation is gradually increasing, which indicates that, with the increase of the stock of social capital, the willingness of citizen participation in social governance increases. The influence of social trust and social norms on the citizens’ willingness is as follows:

(1) The influence of social trust on the citizens’ willingness to participate in social governance:

there is also a significant positive correlation between social trust and the citizens' willingness to participate in social governance, that is, the higher the degree of social trust is, the stronger the willingness of citizen participation is. That is because the pattern of multiple participants has been formed that may lead to collective actions in China, which are greatly reliable on the degree of trust, for "other things being equal, the greater the communication (both direct and indirect) among participants is, the greater their mutual trust and the easier they will find it to cooperate" [42], and social trust could promote the probability of voluntary cooperation among members through mutual influences. In addition, most of the citizen participations in social governance are voluntary, whose effect and quality depend on the mutual cooperation, which cannot be separated from the support of mutual trust [43–45]. On different categorical levels, social trust has different levels of influence on the citizens' willingness to participate in social governance: on "categorical level = 2," there is no significant influence of social trust on strong-engaging citizens, with the possible reason that, to strong-engaging citizens, the degree of participation are greatly influenced by their original motivations or jobs (tasks); therefore, they are not sensitive to other external factors; on "categorical level = 1," there is a negative correlation between social trust and the willingness of citizen participation in social governance, whereas on "categorical level = 3" and "categorical level = 4," the correlations turn into positive and the reliable level of the willingness on social trust is relatively high, which further proves that, with the increase of social trust level, the willingness of citizen participation in social governance increases accordingly, especially to the citizens of latter two categorical levels (limited participation), who have more options [46].

- (2) The influence of social norms on the citizens' willingness to participate in social governance: in general, there is also a significant positive correlation between the social norms and the willingness of citizen participation in social governance, namely, the higher the degree of social norms is, the stronger the willingness of citizen participation in social governance is. The current phenomenon of disordered participation and symbolic participation are largely because of the absence of social norms in China. For an orderly citizen participation in social governance, on one hand the social norms are the reciprocal norms, with the networks as basis and the trust as the core, which could provide the citizens with benign social norms and orders to engage in, accelerate the citizens to concur with a certain policy decision, and remove the barriers

of citizen participation, thereby promote the willingness of citizen participation in social governance; on the other hand, social norms reflect individuals' consent with the referent group on its decision-making behavior, then obeying the norms will earn respect and reputation within the group; otherwise, they would be possible to be excluded and isolated. Consequently, the social norms will constrain the individuals' behavior of participation in social governance and motivate the citizens to obey the public interest instead of certain personal interest; thus, their willingness of participation will be promoted effectively. It indicates that the higher the degree of social norms is, the more powerful the institution and order of citizen participation are, the greater the consent of participating into decision-making is, and the stronger the willingness of participation is.

The influence of social norms on the citizens' willingness to participate in social governance differentiates on different categorical levels: just like the social trust, on "categorical level = 2," social norms does not impact significantly on the participation willingness of "strong-participation" type, with the possible reason that, to the citizens of this type, their own participation willingness is relatively strong because of their original motivations or jobs (tasks); therefore, they rely less on the maturity of social norms. Likewise, on "categorical level = 1," there is a negative correlation between social norms and participation willingness, whereas on "categorical level = 3" and "categorical level = 4," the correlation turns into positive, and the participation willingness relies more on the social norms, which further verifies that, with the increasing social norms, the willingness of citizen participation in social governance enhances accordingly, especially to the citizens of the latter two categorical levels; in addition, the maturity of social norms apparently improves the willingness of citizen participation in social governance.

- (3) *The influence of demographic characteristics on the citizens' willingness to participate in social governance.* In general, there is a significant correlation between the gender, education, political status, region factors, and the citizens' willingness to participate in social governance, while there is no such significant correlation between the age, occupation, income, urban or rural status factors, and the willingness. Specifically, (1) gender factor has significant impact on the willingness of citizen participation in social governance. Compared with the male citizens, the female citizens have weaker willingness and the impact differentiates on different categorical levels, and it may be mainly because of the different gender roles given by the society that the male citizens are more appropriate



for the management in political and public affairs; therefore, the male citizens have stronger willingness (in China, men and women are endowed with different roles by the society. This is due to the influence of the superiority of men and the inferiority of women in Chinese traditional Confucian culture, that is, men play a dominant role in family and social life, while women are in a subordinate position. Therefore, politics and public affairs are suitable places for typical male characteristics). (2) There is a positive correlation between education and the willingness, and with the increase of the willingness, education has more significant impact on the citizens' willingness to participate in social governance, which indicates that the more educated the citizens are, the stronger their willingness of participating in social governance is. The reason lies in that the more educated citizens will act in a more rational way, their degree of self-identification is higher, their ability of participation is higher, and they have more opportunities to participate in [47]. (3) Political status factor has significant impact on the willingness. Compared to the citizens with non-CPC membership, the citizens with CPC membership have stronger willingness of participation, and the citizens on different categorical levels differentiate on the participation willingness. The main reason may lie in that CPC party members regularly or irregularly attend various kinds of studies or activities, which may dramatically improve the willingness of citizen participation in social governance. (4) Region factor has significant impact on the willingness of citizen participation. Compared to the citizens from the Eastern region, the citizens from Central and Western in China have lower willingness, and the citizens on different categorical levels differentiate on the participation willingness. The main reason lies in that the division of regions were much based on the economic development level, which may be shown as the "U" characteristic among Eastern, Central, and Western regions, and higher level of political participation is always accompanied with higher level of development, and political participation tends to be evaluated higher in a society of greater socioeconomic development. Therefore, the willingness of participation in the relatively developed Eastern region is stronger than that in the less developed Central and Western regions.

## 5. Discussion

Social networks have a significant impact on the citizens' willingness to participate in social governance, as well as the gender, education, political status, and region factors of demographic characteristics, while age, occupation, income, and urban or rural status factors do not. The primary conclusions are as follows. (1) Social networks are invisible resources, the citizens who own abundant social networks

are easy to obtain and to share the relevant information of social governance timely through the social networks, while frequent social interaction could boost the width, depth, and efficiency of information exchange among the citizens, to effectively mitigate the risk of asymmetric information and the cost of participation, and to improve the benefit, thereby enhance the willingness of citizen participation. In addition, the strength of the willingness of citizen participation relies on the maturity of social networks greatly. (2) Social capital composed of social networks, social trust, and social norms will also have a significant impact on the citizens' willingness to participate in social governance. The higher stock of social capital means lower transaction costs for the citizens to participate in social governance, lower participation costs, lower risk of asymmetric information, and greater probability of collective actions, which in all will accelerate the process of turning the willingness of citizen participation into action; therefore, to a large extent, the willingness of citizen participation in social governance relies on the stock of social capital, namely, with the increase of the stock of social capital, the willingness of citizen participation in social governance is strengthened. In addition, the other two components of social capital also influence the willingness of citizen participation through social networks. (3) The age, occupation, income, and urban or rural status factors are not significantly correlated with the citizen participation willingness, while gender, education, political status, and region factors are significantly correlated with that, namely, comparatively, the citizens who are male, well educated, CPC members, and from the Eastern region have stronger willingness of participating in social governance.

Compared with the existing literature, more studies focus on political participation or social participation, reflecting the impact on policy agenda and government behavior. China's social governance is a kind of governance oriented by people's livelihood services, with distinct Chinese characteristics. The citizen participation in social governance is different from the traditional political participation and social participation. It is a kind of public participation relying on the community and preferring to the public service supply decision-making, which is also reflecting Chinese wisdom and citizen participation's Chinese characteristics. At the same time, China's social network also has its own characteristics. Therefore, the influence of China's social networks on the citizens' willingness to participate in social governance has a strong sample significance is studied, which may provide useful reference and enlightenment for the citizen participation in other countries all over the world.

Due to the limited energy and ability, the samples selected in this research are limited. There are still some defects and deficiencies in the representativeness and typicality of the samples, which need to be further improved in the future.

## 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 study was supported by the National Social Science Fund of China (Grant no. 18CGL033). We also gratefully acknowledge many useful suggestions by Iveta Reinholde.

## References

- [1] B. P. Bourdieu, "The forms of capital," in *Handbook of Theory and Research for the Sociology of Capital*, J. G. Richardson, Ed., pp. 241–258, Greenwood Press, New York, NY, USA, 1986.
- [2] M. S. Granovetter, "The strength of weak ties," *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.
- [3] R. D. Putnam, *Bowling Alone: The Collapse and Revival of American Community*, Simon & Schuster, New York, NY, USA, 2000.
- [4] I. Ajzen and M. Fishbein, "Scaling and testing multiplicative combinations in the expectancy-value model of attitudes," *Journal of Applied Social Psychology*, vol. 38, no. 9, pp. 2222–2247, 2008.
- [5] A. Portes, "Social capital: its origins and applications in modern sociology," *Annual Review of Sociology*, vol. 24, no. 1, pp. 1–24, 1998.
- [6] M. Li, "Social network and social capital in leadership and management research: a review of causal methods," *The Leadership Quarterly*, vol. 24, no. 5, pp. 638–665, 2013.
- [7] J. Nahapiet and S. Ghoshal, "Social capital, intellectual capital, and the organizational advantage," *The Academy of Management Review*, vol. 23, no. 2, pp. 242–266, 1998.
- [8] J. S. Coleman, "Social capital in the creation of human capital," *The American Journal of Sociology*, vol. 94, no. suppl, pp. 95–120, 1988.
- [9] R. D. Putnam, "What makes democracy work?" *National Civic Review*, vol. 82, no. 2, pp. 101–107, 1993.
- [10] N. Lin, "Building a network theory of social capital," *Connections*, vol. 22, no. 1, pp. 28–51, 1999.
- [11] F. Fukuyama, "Social capital and civil society," in *Proceedings of the No. 00/74, IMF Working Papers, International Monetary Fund*, Washington, DC USA, March 2000, <https://EconPapers.repec.org/RePEc:imf:imfwpa:00/74>.
- [12] E. Ostrom, "Collective action and the evolution of social norms," *Journal of Economic Perspectives*, vol. 14, no. 3, pp. 137–158, 2000.
- [13] E. L. Glaeser, D. Laibson, and B. Sacerdote, "An economic approach to social capital," *The Economic Journal*, vol. 112, no. 483, pp. 437–458, 2003.
- [14] J. M. Roberts and F. Devine, "Some everyday experiences of voluntarism: social capital, pleasure, and the contingency of participation," *Social Politics: International Studies in Gender, State & Society*, vol. 11, no. 2, pp. 280–296, 2004.
- [15] Q. Fu, *Trust, Social Capital, and Organizational Effectiveness*, Master's Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 2004.
- [16] G. A. Brewer, "Building social capital: civic attitudes and behavior of public servants," *Journal of Public Administration Research and Theory*, vol. 13, no. 1, pp. 5–26, 2003.
- [17] M. M. Skoric, D. Ying, and Y. Ng, "Bowling online, not alone: online social capital and political participation in Singapore," *Journal of Computer-Mediated Communication*, vol. 14, no. 2, pp. 414–433, 2009.
- [18] M.-C. Torri, "Community-based enterprises: a promising basis towards an alternative entrepreneurial model for sustainability enhancing livelihoods and promoting socio-economic development in rural India," *Journal of Small Business & Entrepreneurship*, vol. 23, no. 2, pp. 237–248, 2010.
- [19] K. K. Fung and S. L. Hung, "Strengthening a community of poverty in an affluent society: strategies to build social capital in Tin Shui Wai North in Hong Kong," *Community Development Journal*, vol. 49, no. 3, pp. 441–457, 2014.
- [20] D. Asteria, A. Halimatussadiyah, fnm. Budidarmono, D. Utari, and R. Dwi Handayani, "Relation of social capital to women's proactive participation in the community for sustainability of river in urban areas," *E3s Web of Conferences*, vol. 73, Article ID 03003, 2018.
- [21] H. Suh and R. S. Heidi, "A contingent effect of trust? Interpersonal trust and social movement participation in political context," *Social Science Quarterly*, vol. 99, no. 4, pp. 1484–1495, 2018.
- [22] H. Zúñiga, N. Jung, and S. Valenzuela, "Social media use for news and individuals' social capital, civic engagement and political participation," *Journal of Computer-Mediated Communication*, vol. 17, no. 3, pp. 319–336, 2012.
- [23] Y. Hu, F. Xiong, S. Pan, X. Xiong, L. Wang, and H. Chen, "Bayesian personalized ranking based on multiple-layer neighborhoods," *Information Sciences*, vol. 542, pp. 156–176, 2021.
- [24] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, vol. 50, 2019.
- [25] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [26] P. S. Adler and S.-W. Kwon, "Social capital: prospects for a new concept," *Academy of Management Review*, vol. 27, no. 1, pp. 17–40, 2002.
- [27] C. Boix and D. N. Posner, "Social capital: explaining its origins and effects on government performance," *British Journal of Political Science*, vol. 28, no. 4, pp. 686–693, 1998.
- [28] D. Gambetta, *Trust: Making and Breaking Cooperative Relations*, Vol. 217, Basil Blackwell, New York, NY, USA, 1988.
- [29] S. E. S. Crawford and E. Ostrom, "A grammar of institutions," *American Political Science Review*, vol. 89, no. 3, pp. 582–600, 1995.
- [30] M. S. Granovetter, "Economic action and social structure: the problem of embeddedness," *The American Journal of Sociology*, vol. 91, no. 3, pp. 482–510, 1985.
- [31] K. L. Schlozman, N. Burns, and S. Verba, "Gender and the pathways to participation: the role of resources," *The Journal of Politics*, vol. 56, no. 4, pp. 963–990, 1994.
- [32] N. Roberts, "Public deliberation in an age of direct citizen participation," *The American Review of Public Administration*, vol. 34, no. 4, pp. 315–353, 2004.
- [33] J. Kim, "The Impacts of social capital on civic participative institution's effectiveness," *Korean Public Administration Quarterly*, vol. 23, no. 9, pp. 1007–1031, 2011.
- [34] P. Paxton, "Social capital and democracy: an interdependent relationship," *American Sociological Review*, vol. 67, no. 2, pp. 254–277, 2002.
- [35] R. D. Putnam, "The prosperous community: social capital and public life," *The American Prospect*, vol. 13, no. 4, pp. 35–42, 1993.
- [36] C. Grootaert and T. van Bastlaer, *Understand and Measuring Social Capital*, World Bank, Washington, DC, USA, 2002.

- [37] T. Harpham, "The measurement of community social capital through surveys," in *Social Capital and Health*, I. Kawachi, S. Subramanian, and D. Kim, Eds., Springer, New York, NY, USA, 2007.
- [38] F. Sabatini, "Social capital as social networks: a new framework for measurement and an empirical analysis of its determinants and consequences," *The Journal of Socio-Economics*, vol. 38, no. 3, pp. 429–442, 2009.
- [39] C. F. Baum, *An Introduction to Modern Econometrics Using Stata*, Stata Press books, College Station, TX, USA, 2006.
- [40] S. R. Arnstein, "A ladder of citizen participation," *Journal of the American Institute of Planners*, vol. 35, no. 4, pp. 216–224, 1969.
- [41] A. Ferrer-i-Carbonell and P. Frijters, "How important is methodology for the estimates of the determinants of happiness?" *The Economic Journal*, vol. 114, no. 497, pp. 641–659, 2004.
- [42] R. D. Putnam, *Making Democracy Work: Civic Tradition In Modern Italy*, Vol. 173, Princeton University Press, Princeton, NY, USA, 1993.
- [43] C. D. De and T. R. Tyler, "The effects of trust in authority and procedural fairness on cooperation," *Journal of Applied Psychology*, vol. 92, no. 3, pp. 639–649, 2007.
- [44] M. J. Hetherington, *Why Trust Matters: Declining Political Trust and the Demise of American Liberalism*, Princeton University Press, Princeton, NY, USA, 2005.
- [45] J. Delhey and K. Newton, "Who trusts? The origins of social trust in seven societies," *European Societies*, vol. 5, no. 2, pp. 93–137, 2003.
- [46] L. Guiso, P. Sapienza, and L. Zingales, "The role of social capital in financial development," *American Economic Review*, vol. 94, no. 3, pp. 526–556, 2004.
- [47] M. M. Conway, "Political participation in the United States," *Political Psychology*, vol. 7, no. 4, p. 807, 1991.

## Research Article

# Anatomy of Complex System Research

Yang Wang <sup>1</sup> and Xinyue Wang<sup>2</sup>

<sup>1</sup>*School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an 710049, China*

<sup>2</sup>*School of Humanities and Social Science, Xi'an Jiaotong University, Xi'an 710049, China*

Correspondence should be addressed to Yang Wang; yang.wang@xjtu.edu.cn

Received 25 June 2020; Accepted 23 July 2020; Published 9 October 2020

Guest Editor: Fei Xiong

Copyright © 2020 Yang Wang and Xinyue Wang. 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.

Complex systems play important roles in science and economy and have become one of the major intellectual and scientific challenges of the twenty-first century. However, the way complex system research connects to other academic fields is much less well understood, with only anecdotal evidence. In this work, we present an anatomy of complex system research by leveraging a large-scale dataset that contains more than 100 million digitalized publications. First, we find that complex system research shows a steady growth relative to the whole science in the last 60 years, with a sudden burst after 2000, which might be related to the development of computational technologies. Although early complex system study shows strong referencing behaviors to mathematics and physics, it couples significantly with computer science in the twenty-first century and affects engineering strongly. Moreover, we find empirical evidence that complex system research tends to have multidisciplinary nature, as it is often inspired by or affects a diverse set of disciplines. Finally, we find significant positive correlations between fields' reference broadness and future scientific impacts. Overall, our findings are consistent with several characteristics of complex system research, i.e., its multidisciplinary, quantitative, and computational nature, and may have broad policy implications for supporting and nurturing multidisciplinary research.

## 1. Introduction

"I think the next century will be the century of complexity" said Stephen Hawking, who is a famous theoretical physicist and cosmologist. Indeed, we are surrounded by systems that are unprecedentedly complex, containing many individual components together with interactions among them, possibly nonlinear ones [1]. Consider, for example, the survival of human society needs cooperation between billions of individuals or communication infrastructures that connect and integrate human efforts. Another example is science. As a complex system, science contains millions of digitalized papers, patents, grant proposals, white papers, together with citation or referencing relationships among them, catalyzing the rapid development of modern science and technology [2, 3]. The above examples vividly illustrate that today's complex system research has implications for our society, from human mobilities [4], to economics [5], world trades [6], and human interactions or collaborations [7, 8].

Particularly with the emergence of network science at the dawn of the twenty-first century, complex system study generates a substantial amount of original research [9]. Also, it is related to many revolutionary technologies of the twenty-first century, including Google, Facebook, and Twitter [10].

The applied values and substantial implications of complex system study raise very interesting and important questions: where does complex system research come from? How tightly connected are other academic areas and the cutting-edge complex system research? Anecdotal evidence indicates that complex system study may be rooted in chaos theory, and it is also believed to be tightly connected to self-organization, agent-based modeling, and artificial intelligence. Despite of that, we lack a systematic and empirical understanding on how complex system research affects or is affected by other academic fields. Thanks to the availability of large-scale bibliometric datasets and the development of science of science research, the above-mentioned questions

can be empirically answered for now. Science of science offers us unprecedented opportunities to explore and uncover the citation dynamics [11–13], knowledge space [14], collaboration networks [7, 8, 15], and their association with societal impact [16, 17]. Indeed, recent empirical work regarding knowledge flow and scientist mobility among academic fields have attracted much attention, including the field of artificial intelligence [18], subfields of physics [19–21], and science in general [22, 23], together with the diffusion model [24].

Here we leverage the Microsoft Academic Graph (MAG) dataset to study the reference and citation patterns between complex system research and other academic fields, exploring their relationships. Consistent with the anecdotal evidence, we find that complex system research is inspired by mathematics and physics before 2000. Moreover, due to the rapid development of the computational technologies and the availability of large-scale datasets, complex system study has become largely connected with computer science afterwards. Finally, we show empirically that complex system study is inspired by a diverse set of disciplines compared with many other fields such as nuclear physics, and this broadness in the reference list tends to be associated with field’s future scientific impact. Together, our study provides among the first empirical evidences that complex system research has multidisciplinary and computational nature. With the very nature of complex system research that may transfer to future scientific impact, our findings may have broad policy implications.

## 2. Results

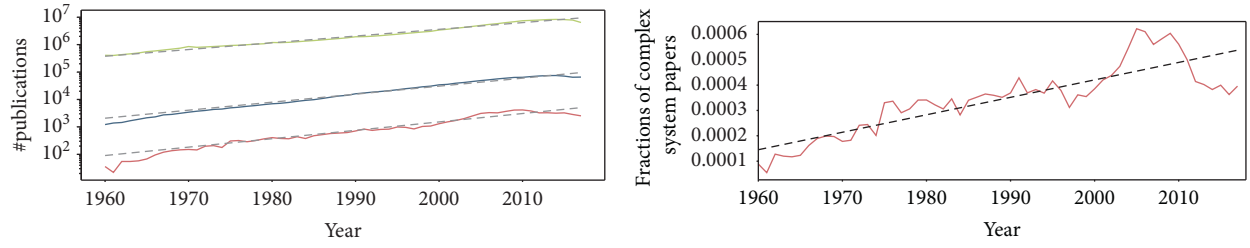
The MAG dataset is from Microsoft Academic Services [25, 26], which contains more than 100 million papers from 1800 to 2018 together with their citation relationships. Moreover, the MAG dataset uses the natural language processing to identify paper academic fields, resulting in a field tree with 19 coarse-grained (level 0) and more than 80,000 subfields (level 1 and level 2) spanning from biology to computer science. The citation relationships and paper field information, together with the sufficient time span, allow us to understand the evolution of scientific fields. Finally, as will be shown below, the MAG data is also useful to our study since it contains sufficient computer science papers, which are well related to complex systems (there are 70,335 complex system papers in the MAG dataset).

*2.1. Knowledge Production and Citation Networks.* Throughout the past 60 years, major technological changes and paradigm shifts such as the development of chaos theory and computer science result in a substantial knowledge production in the field of complex systems, including network science, data science, and computational social science [3, 9, 27]. To quantify the knowledge production, Figure 1(a) shows the yearly number of publications from 1960 to 2018. Consistent with prior studies [19, 23], we find that science grows at an exponential rate, doubling every 12.3 years (Figure 1(a)). Complex system research also shows an

exponential growth trend (Figure 1(a)). Comparing the knowledge production of complex systems and the whole science, we find that complex system research shows a steady growth relative to science. Specifically, the percentage of complex system paper grows from less than 0.01% to 0.06% from 1970 to the end of 2010 with a sudden burst between 2000 and 2010, representing the golden age of complex systems (Figure 1(b)). Hence, the field’s exponential growth is not only driven by social needs, but also by paradigm changes, such as the development of computer science and artificial intelligence. The results also agree with the Stephen Hawking’s famous quote mentioned in the beginning of the paper.

Following the substantial growth of the core complex system field, it is also interesting to study the growth of its related fields. The next question is as follows: How to detect fields that are related to complex systems? Sam Edwards remarked the definition of physics as “*what physicists do*” [28]. Following in his definition, we expand the field of complex systems to a broad perspective by including papers that cite or were cited by complex system papers. We find academic papers related to complex system research also showing a similar exponential growth pattern with a doubling time of 10.33 years (Figure 1(a)). One may argue that papers that cite or were cited by complex system papers may not belong to its related fields, since such citation relationships might come from noise or random connections in the whole citation network. To address this issue, we compare the citation patterns of all MAG papers to a null model in which each paper’s references are assigned randomly given their publication time, regardless of a paper’s journal or research field [16]. Specifically, in the randomized MAG citation network, we switch all citation links between papers but preserve the total number of references of each paper and the year of referencing and referenced papers [16]. After that, we define papers related to complex systems if its relationships (references or citations) to complex system field are significantly higher than expected by chance [19]. Followed by the whole procedure, we end up with more than 1.2 million papers that are related to complex systems, showing consistent results with Figure 1(a). The steady growth of complex systems and its related fields raises an interesting question: Where does complex system research locate in the scientific knowledge space?

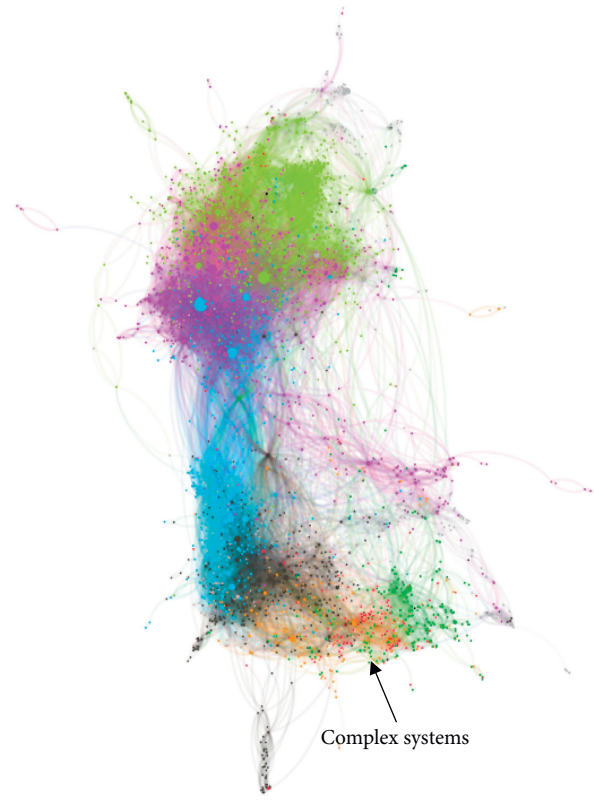
To visualize the knowledge graph, we extracted citation relationships among different fields to construct a field citation network (here, we use the MAG level-2 fields since complex system research belongs to this level). In the network, each node represents an academic field, and links between different fields are referencing or citation behaviors. Additionally, we consider the number of citations/references between two fields as the link weight in the citation network. To eliminate noise and insignificant links, we applied a backbone extraction method to the whole field citation network [29]. While the citation network often shows strong community structure, with subfields belonging to the same ancestor cite each other much more frequently than those with different ancestor fields (Figures 1(c)–1(e)), many fields show strong



— #Complex systems  
 — #Related  
 — #MAG

(a)

(b)



■ Biology  
 ■ Medicine  
 ■ Chemistry  
 ■ Physics

■ Mathematics  
 ■ Engineering  
 ■ Computer science

■ Biology  
 ■ Medicine  
 ■ Chemistry  
 ■ Physics

■ Mathematics  
 ■ Engineering  
 ■ Computer science

(c)

(d)

FIGURE 1: Continued.

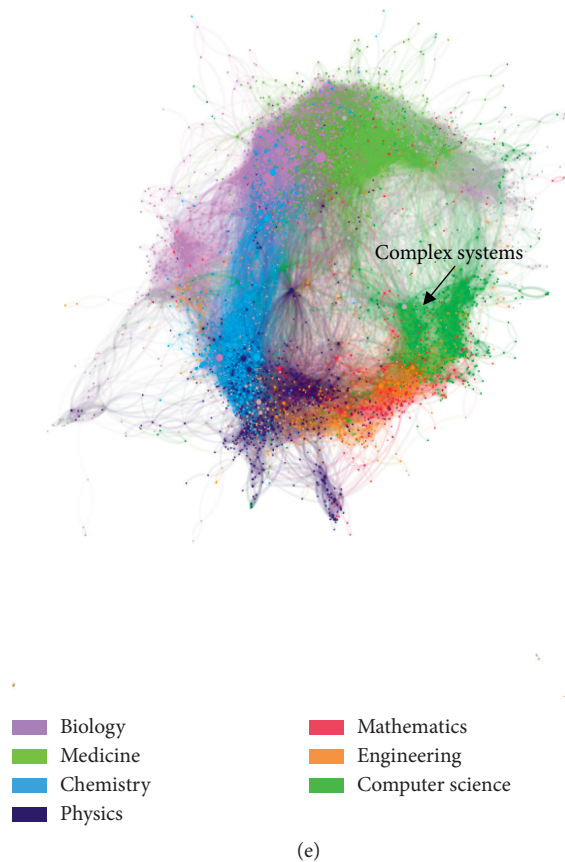


FIGURE 1: The knowledge production and citation networks. (a) The yearly number of publications as a function of time from 1960 to 2018, for the whole Microsoft Academic Graph (green) and the field of complex systems (red) and its related fields (blue). We fit each growth pattern using an exponential function, shown in dashed lines. (b) The fraction of yearly number of papers of complex system relative to the whole MAG, together with a linear fit. Field citation networks in the (c) 1980s, (d) 1990s, and (e) 2000s, with nodes being MAG level-2 fields and links of the citation/reference relationships between two fields. In order to eliminate insignificant links, we applied a backbone extraction method. Different colors represent ancestor fields of these level-2 fields. We also annotate the field of complex system in the citation network.

interdisciplinary features. For example, chemistry lies between physics and biology, while engineering stands in the intersection of physics, mathematics, and computer science. Looking across different periods, we find complex system research moved from mathematics in the 1980s to computer science in the 2000s (Figures 1(c)–1(e)), indicating a substantial paradigm shift towards computational research [30]. Our findings are consistent with the fact that the development of computer science as well as the availability of large-scale datasets has catalyzed complex system research, possibly the emergence of network science as well. The knowledge production by the field of complex system, together with its position in the whole knowledge graph, prompts us to ask the following: Which specific field is closely related to complex systems and how to quantify the relationships between various fields rather than inspections?

*2.2. Knowledge Flow and Its Related Fields.* What are the origins of complex system research? Where does the knowledge created by complex system research go? The “linear model” of science demonstrates the importance of basic science to the development of applied research and technological development [31]. From the history of complex system field, its study of nonlinear dynamics or chaos originated from mathematics, atmospheric science, and physics deepens our understanding of its underlying mechanisms. Also, the development of complex system study contributes to the emergence of network science, data science, and computational social science. Therefore, we hope to understand the relevant knowledge flow. Does complex system theory emerge from mathematics or physics? Do computer science and engineering apply methodologies developed by the field of complex systems?

To investigate, we study the association between complex system research and various scientific fields through the relationship of references and citations. Specifically, we follow the definition of related papers to complex system research, and Figure 2(a) shows the distribution of their academic fields. We find that computer science and mathematics are the top two most frequent fields related to complex system research, followed by the fields of biology, chemistry, and physics. Other fields like philosophy or art are less likely to have connections with complex system study (Figure 2(a)).

To eliminate the effect of exponential growth of the paper production (Figure 1(a)), we calculate the share of references made from complex system study to other academic fields, as well as papers in other fields to complex system study. Following previous research [18, 19], the reference share from field  $A$  to field  $B$  is defined as

$$\text{share}_t(A, B) = \frac{\#\text{refs from A papers to B papers in year } t}{\#\text{refs made by A papers in year } t}. \quad (1)$$

The reference share measurement controls the number of references made by field  $A$ , thus controlling for its knowledge production (Figures 2(b) and 2(d)). Additionally, in order to control the paper production in field  $B$  that is the referenced field, we also compare the real number of references between fields  $A$  and  $B$  with the same quantity in the randomized citation network as defined in the previous section. Specifically, we focus on the  $z$ -score between the two fields as follows:

$$z_{A,B} = \frac{N_{\text{real}} - \mu_{A,B}}{\sigma_{A,B}}, \quad (2)$$

where  $\mu_{A,B}$  and  $\sigma_{A,B}$  represent the average number and the standard deviation of connections between fields  $A$  and  $B$  generated from multiple randomizations, respectively (Figures 2(c) and 2(e)). A large  $z$ -score means fields  $A$  and  $B$  are more likely to be connected with respect to the null model.

Before 2000, complex system research frequently cites papers from mathematics and physics (Figure 2(b)); we find similar results after controlling for the paper production in referenced field (Figure 2(c)), suggesting early complex system study was shaped by mathematics and physics. The result is also consistent with our intuition that complex system study is closely related to the discovery of dynamical system theory or nonlinear system and nonequilibrium thermodynamics [32]. After 2000, however, complex system research strongly relies on computer science, indicating the paradigm shifts to computational research, such as data science and computational social science [33]. Interestingly, the references from complex system to chemistry and biology are largely driven by the dominance fraction of chemistry/biology papers with respect to the whole science. In fact, complex system research is less likely to cite biology and chemistry compared to the null model (Figure 2(c)).

How does complex system research affect other fields? We repeat the above analysis by calculating the reference share and  $z$ -score from other fields to complex system

research. Several insightful patterns emerge. First, complex system research also affects mathematics and physics significantly before 2000, while computer science substantially cites complex system research after 2000 with steady growth pattern (Figures 2(d) and 2(e)). The results suggest that, in terms of knowledge flow, complex system research does not only act as a target but also an important input. Interestingly, engineering depends strongly on the development of complex system research after 2005 (Figure 2(e)), suggesting the applied value of complex system research.

The above analysis allows us to uncover the relationships between complex system study and other academic fields, prompting us to aggregate multiple years in order to investigate whether complex system study affects other fields more or vice versa. Figure 3 shows the  $z$ -score defined above for three different periods, i.e., 1980s, 1990s, and 2000s. By comparing  $z$ -score from complex system field to other fields (Figure 3, red bars) with the same quantity from other fields to complex system research (Figure 3, blue bars), we find the  $z$ -score from complex system to the field of mathematics is higher than the other way around, suggesting that complex system research consistently affects mathematics more from the year of 1980 (Figure 3(a)). Computer science, however, shows comparable knowledge flow from and to complex system research (Figure 3(b)), suggesting an extensive convolution structure between these two fields. Interestingly, engineering is more affected by complex system research since 2000, indicating the applied value of complex system research (Figure 3(e)). Overall, the framework of knowledge flow allows us to uncover the statistical regularities of interactions between various scientific fields.

**2.3. Reference and Impact Broadness.** Inspired by the results above, we discuss, in this section, the major characteristics of complex system research, and questions we would like to answer are as follows: Does complex system research make references from a diverse set of fields? What is the impact of complex system research?

In order to answer these questions, we quantify the broadness of disciplines reflected in the references and citations for every paper that belongs to the same academic field using the entropy measurement [34]. Specifically, given a paper with a reference or citation list, we embed each reference/citation to a vector of academic fields  $d_\alpha = (d_{\alpha 1}, d_{\alpha 2}, \dots, d_{\alpha k})$  based on the MAG level-1 fields. Then, the  $i$ th element of the paper is given by

$$a_i = \sum_{\alpha=1}^n \sum_{\beta=1}^k \frac{\delta(d_{\alpha\beta}, i)}{k}, \quad (3)$$

where  $\delta$  is the delta function with 1 when  $d_{\alpha\beta} = i$  and 0 otherwise,  $k$  is the number of MAG level-1 fields a reference/citation belongs to,  $n$  is the number of references for a given paper, and  $\alpha$  and  $\beta$  are the index looping from the paper's reference list and field information of each reference, respectively. After embedding every paper to a vector, we perform L1 normalization to each vector, denoted by  $p$ . We then calculate the normalized entropy for each paper as



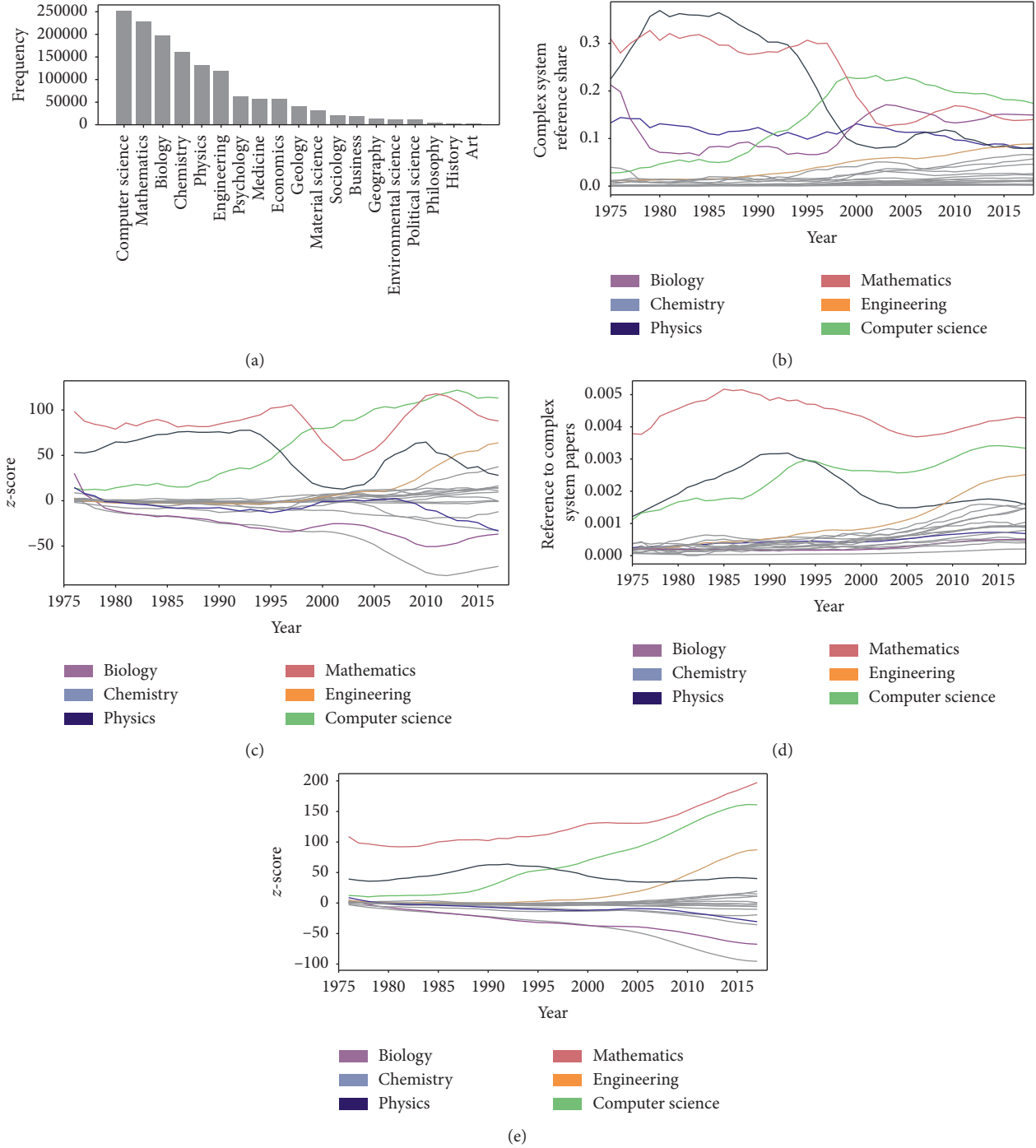


FIGURE 2: Knowledge between complex system study and other academic fields. (a) The distribution of academic fields for papers that cite or are cited by complex system research at least once. (b, c) The reference share and z-score (compared to the null model) from complex system research to other academic fields. (d, e) The same as (b) and (c) but from other academic fields to complex system research.

$$E(p) = -\frac{1}{\log_2 N} \sum_i^N p_i \log_2 p_i, \quad (4)$$

where  $N$  is the total number of level-1 fields in the MAG dataset. A paper with  $E = 1$  reflects its references/citations' equal distribution across all fields (broad referencing and impact); a paper with  $E = 0$  indicates the

paper's references/citations are solely from one field (deeply disciplinary). Finally, we average the entropy over all papers within the same academic field to get the field reference or impact broadness.

First, Figures 4(a) and 4(b) show the distribution of field reference and impact broadness at 1995, documenting the narrow range of both referencing and impact broadness that are in contrast to the fat-tail citations distribution [35].

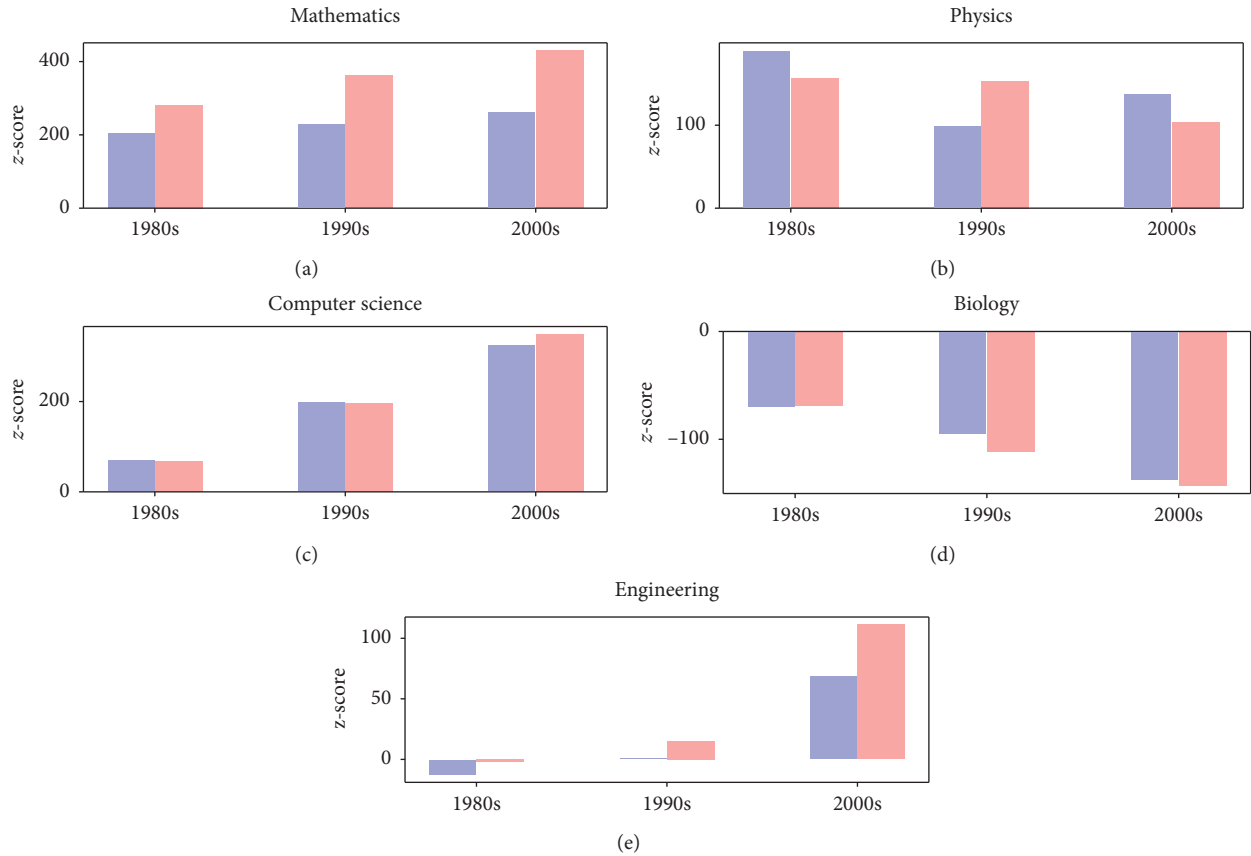


FIGURE 3: The aggregated z-score between complex system research and several selected academic fields. (a) Mathematics, (b) physics, (c) computer science, (d) biology, and (e) engineering. Specifically, blue bars mean how complex system research cites other fields, and red bars represent how other fields cite complex system research.

Interestingly, complex system research shows a relatively diverse feature among all academic fields, as shown in the dash line (Figures 4(a) and 4(b)). Moving beyond a single year, we repeat the same analysis for all academic fields from 1970 to 2018, finding similar results (Figures 4(c) and 4(d)). Comparing the reference or impact broadness of complex system research with the same quantities for nuclear physics, which is often considered as a deeply disciplinary field [19], we find that complex system research consistently shows higher value of reference or impact broadness (Figures 4(c) and 4(d)). Our results thus show a multidisciplinary nature of complex system research, suggesting that complex systems research may offer a language in which many academic fields can interact with each other without border.

**2.4. Reference Broadness and Scientific Impact.** The multidisciplinary nature of complex system research encourages us to explore the association between field's reference broadness and its scientific impact, which has policy implications for funding agencies. To answer this question, we calculate, for each field, its reference broadness and its impact, which is measured by the average paper citations within 8 years of publication for a specific field (C8). We find a significant positive correlation between field reference

broadness and the scientific impact (Figures 5(a) and 5(b)), suggesting that fields with more diverse reference lists are more likely to have more scientific impact in the future. To test the robustness of this association, we repeat the analysis from 1970 to 2008, finding consistent and significant results (Figure 5(c)).

Finally, one might argue that different scientific fields often show different citation patterns. For example, biology and chemistry often have, on average, more citations than mathematics and engineering [35]. Moreover, citation behaviors change over time [11]. In order to eliminate this concern, we use a simple linear regression by focusing on the effect of reference broadness on field impact, while controlling for time, ancestor field categories, the number of publications within each field, and the average number of references. We find consistent positive correlation between the reference broadness and field's scientific impact (Figure 5). To further eliminate the effect of the number of references, we did the same regression analysis for subsamples with similar number of references, finding consistent results. The results can be interpreted as follows: one standard deviation (SD) increase of reference broadness for a scientific field is associated with on average 13.9% increase of future scientific impact.

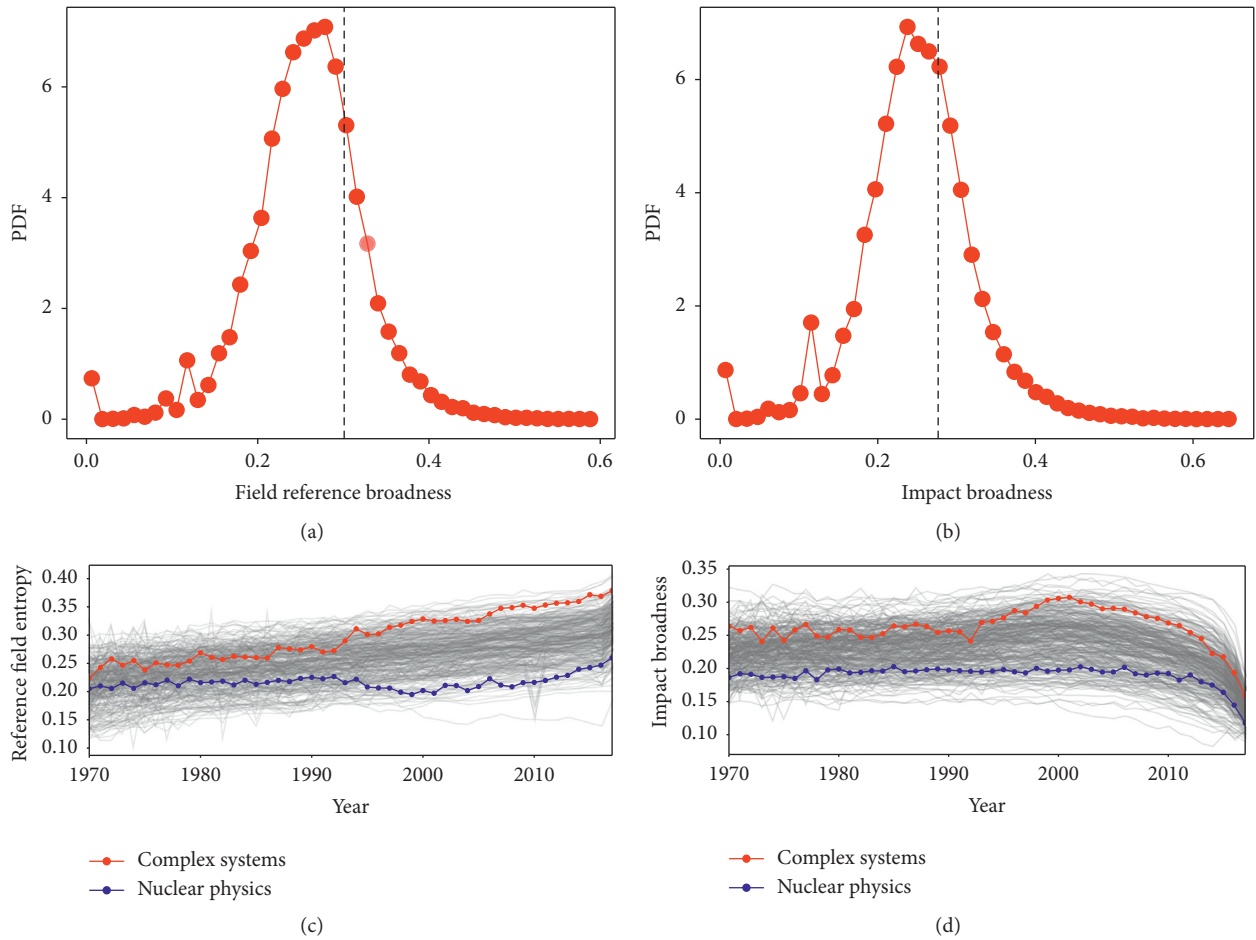


FIGURE 4: Field referencing and citation broadness. The distribution of field reference (a) and impact broadness (b) at 1995, with the dashed line representing complex system research. (c) Field’s reference broadness as a function of time, with red and blue lines showing the field of complex systems and nuclear physics, respectively. (d) The same as (c) but for impact broadness.

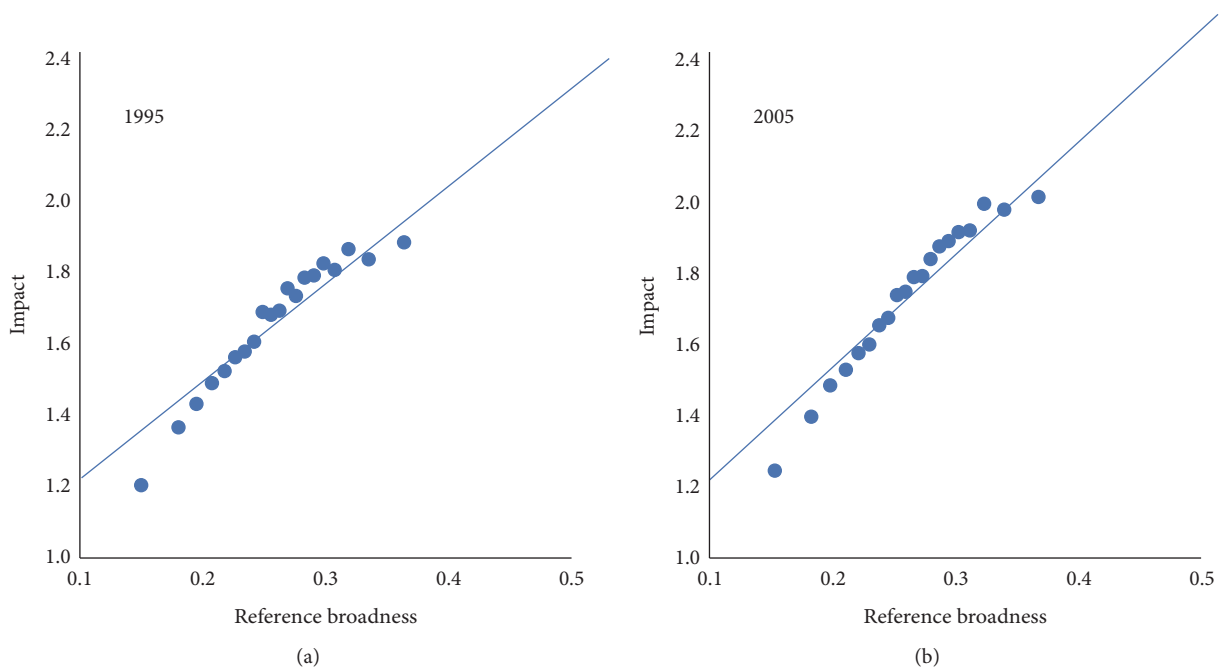


FIGURE 5: Continued.

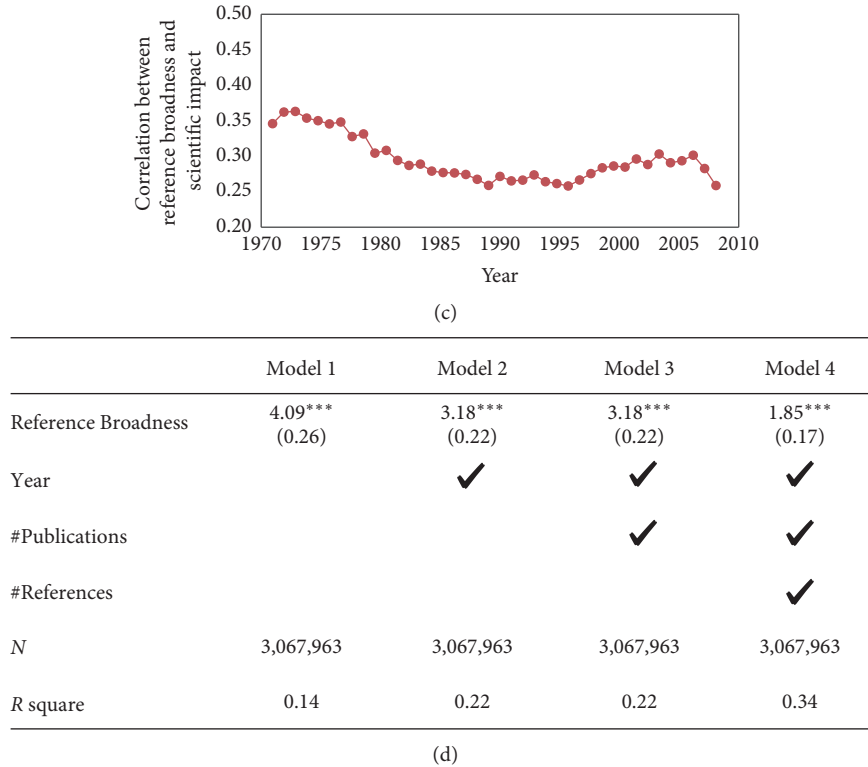


FIGURE 5: The association between field's reference broadness and its scientific impact. The correlation between field reference broadness and its impact for 1995 (a) and 2005 (b). (c) The correlation as a function of time. (d) Linear regression results with dependent variable the field's scientific impact and the focal independent variable as field's reference broadness. Standard errors are clustered at the MAG level-0 field level. \*\*\* $p < 0.001$ , \*\* $p < 0.05$ , and \* $p < 0.1$ .

### 3. Conclusions

Many of us involve with reviewers for funding agencies or publishers may confront with questions like: Can a particular work belong to the framework of complex system research? What is the nature of complex system study? In this paper, we present an anatomy of complex system research spanning from 1960 to 2018 by leveraging a large-scale bibliometric dataset. By investigating its knowledge production, connections with other academic fields, and its reference or impact broadness, our results show the very nature of complex system research, i.e., its multidisciplinary, computational or data driven, and applied nature. The empirical results are in line with many anecdotal evidences from textbooks or review articles [3, 9]. Our results also go beyond case studies, which only focus on referencing/citation behaviors within a specific field. By investigating the connections between fields in the whole knowledge space, we systematically quantify the evolution of the focal field with respect to science.

The significant positive association between fields' reference multidisciplinary and their future scientific impact may have policy implications, from individual department to funding agencies and government. In general, our results are consistent with prior findings that isolations of certain scientific fields bring significant impact penalties [19]. In order to foster and nurture multidisciplinary science that is

key to many real-world problems, funding agencies and government may support fields with multidisciplinary nature. From the perspective of individual scientist [20], our results provide actionable strategies of choosing future research areas [14, 36, 37] and mentors [38]. By choosing research areas that are able to link to their surrounding fields, individual scientists are able to show future scientific impact. Isolation to their environment, however, is destined to extinction [19].

Our understanding of the story that complex system research tends to be inspired by a diverse set of academic areas is not without limitation. There are still several remaining issues for future study. First, while this work considers large-scale citation relationships among scientific fields, there is a lack of other dimensions including scientist mobilities among various fields [20] or text similarities between different papers [39]. Such limitations offer a chance to go beyond our current understanding and ask further questions: how to integrate citations, individual scientists, as well as text feature among academic fields? Given that many important discoveries emerge from the intersections of various fields [16], the analysis of knowledge graph using large-scale dataset could help us uncover or predict emerging/promising research areas. Finally, with the intense connections between science and technology, future work is to investigate the association between fields' multidisciplinary and their roles in advancing technologies.

## Data Availability

This paper uses the MAG dataset, which is publicly available from <https://aka.ms/msracad>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The authors thank Haifeng Du and Qian Huang for their comments. This work was supported by the National Natural Science Foundation of China under Grant nos. L1924078 and 72004177.


## References

- [1] Y. Bar-Yam, "General features of complex systems," in *Encyclopedia of Life Support Systems (EOLSS)*, UNESCO, EOLSS Publishers, Oxford, UK, 2002.
- [2] P. Azoulay, J. Graff-Zivin, B. Uzzi et al., "Toward a more scientific science," *Science*, vol. 361, no. 6408, pp. 1194–1197, 2018.
- [3] A. Zeng, Z. Shen, J. Zhou et al., "The science of science: from the perspective of complex systems," *Physics Reports*, vol. 714–715, pp. 1–73, 2017.
- [4] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [5] C. A. Hidalgo and R. Hausmann, "The building blocks of economic complexity," *Proceedings of the National Academy of Sciences*, vol. 106, no. 26, pp. 10570–10575, 2009.
- [6] C. A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausmann, "The product space conditions the development of nations," *Science*, vol. 317, no. 5837, pp. 482–487, 2007.
- [7] M. E. J. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.
- [8] M. E. J. Newman, "Coauthorship networks and patterns of scientific collaboration," *Proceedings of the National Academy of Sciences*, vol. 101, no. 1, pp. 5200–5205, 2004.
- [9] A.-L. Barabási and M. Pósfai, *Network Science*, Cambridge University Press, Cambridge, UK, 2016.
- [10] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1–7, pp. 107–117, 1998.
- [11] D. Wang, C. Song, and A.-L. Barabási, "Quantifying long-term scientific impact," *Science*, vol. 342, no. 6154, pp. 127–132, 2013.
- [12] M. E. J. Newman, "The first-mover advantage in scientific publication," *EPL (Europhysics Letters)*, vol. 86, no. 6, 2009.
- [13] Q. Ke, E. Ferrara, F. Radicchi, and A. Flammini, "Defining and identifying sleeping beauties in science," *Proceedings of the National Academy of Sciences*, vol. 112, no. 24, pp. 7426–7431, 2015.
- [14] A. Rzhetsky, J. G. Foster, I. T. Foster, and J. A. Evans, "Choosing experiments to accelerate collective discovery," *Proceedings of the National Academy of Sciences*, vol. 112, no. 47, pp. 14569–14574, 2015.
- [15] A. M. Petersen, "Quantifying the impact of weak, strong, and super ties in scientific careers," *Proceedings of the National Academy of Sciences*, vol. 112, no. 34, pp. E4671–E4680, 2015.
- [16] B. Uzzi, S. Mukherjee, M. Stringer, and B. Jones, "Atypical combinations and scientific impact," *Science*, vol. 342, no. 6157, pp. 468–472, 2013.
- [17] D. Li, P. Azoulay, and B. N. Sampat, "The applied value of public investments in biomedical research," *Science*, vol. 356, no. 6333, pp. 78–81, 2017.
- [18] M. R. Frank, D. Wang, M. Cebrian, and I. Rahwan, "The evolution of citation graphs in artificial intelligence research," *Nature Machine Intelligence*, vol. 1, no. 2, pp. 79–85, 2019.
- [19] R. Sinatra, P. Deville, M. Szell, D. Wang, and A.-L. Barabási, "A century of physics," *Nature Physics*, vol. 11, no. 10, pp. 791–796, 2015.
- [20] F. Battiston, F. Musciotto, D. Wang, A.-L. Barabási, M. Szell, and R. Sinatra, "Taking census of physics," *Nature Reviews Physics*, vol. 1, no. 1, pp. 89–97, 2019.
- [21] Y. Sun and V. Latora, "The evolution of knowledge within and across fields in modern physics," *Scientific Reports*, vol. 10, Article ID 12097, 2020.
- [22] E. Yan, "Disciplinary knowledge production and diffusion in science," *Journal of the Association for Information Science and Technology*, vol. 67, no. 9, pp. 2223–2245, 2016.
- [23] Y. Dong, H. Ma, Z. Shen, and K. Wang, "A century of science: globalization of scientific collaborations, citations, and innovations," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1437–1446, Halifax, Canada, August 2017.
- [24] C. Jin, C. Song, J. Bjelland, G. Canright, and D. Wang, "Emergence of scaling in complex substitutive systems," *Nature Human Behaviour*, vol. 3, no. 8, pp. 837–846, 2019.
- [25] A. Sinha, Z. Shen, Y. Song et al., "An overview of microsoft academic service (MAS) and applications," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 243–246, Florence, Italy, May 2015.
- [26] K. Wang, Z. Shen, C. Huang et al., "A review of Microsoft academic services for science of science studies," *Frontiers in Big Data*, vol. 2, 2019.
- [27] S. Fortunato, C. T. Bergstrom, K. Börner et al., "Science of science," *Science*, vol. 359, no. 6397, Article ID eaao0185, 2018.
- [28] A. Donald, "Let's not forget plants," *Physical Biology*, vol. 11, Article ID 053008, 2014.
- [29] M. A. Serrano, M. Boguná, and A. Vespignani, "Extracting the multiscale backbone of complex weighted networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 16, pp. 6483–6488, 2009.
- [30] T. S. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, IL, USA, 1963.
- [31] M. Ahmadpoor and B. F. Jones, "The dual frontier: patented inventions and prior scientific advance," *Science*, vol. 357, no. 6351, pp. 583–587, 2017.
- [32] B. Castellani, *Map of the Complexity Sciences*, Art and Science Factory, Cleveland, OH, USA, 2018.
- [33] D. Lazer, A. Pentland, L. Adamic et al., "Social science: computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.
- [34] A. J. Gates, Q. Ke, O. Varol, and A.-L. Barabási, "Nature's reach: narrow work has broad impact," *Nature*, vol. 575, no. 7781, pp. 32–34, 2019.
- [35] F. Radicchi, S. Fortunato, and C. Castellano, "Universality of citation distributions: toward an objective measure of scientific impact," *Proceedings of the National Academy of Sciences*, vol. 105, no. 45, pp. 17268–17272, 2008.

- [36] T. Jia, D. Wang, and B. K. Szymanski, “Quantifying patterns of research-interest evolution,” *Nature Human Behaviour*, vol. 1, Article ID 0078, 2017.
- [37] A. Zeng, Z. Shen, J. Zhou et al., “Increasing trend of scientists to switch between topics,” *Nature Communications*, vol. 10, pp. 1–11, Article ID 3439, 2019.
- [38] Y. Ma, S. Mukherjee, and B. Uzzi, “Mentorship and protégé success in STEM fields,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 25, pp. 14077–14083, 2020.
- [39] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” 2013, <http://arxiv.org/abs/1301.3781>.

## Research Article

# A Tri-Attention Neural Network Model-Based Recommendation

Nanxin Wang,<sup>1</sup> Libin Yang ,<sup>1</sup> Yu Zheng,<sup>2</sup> Xiaoyan Cai,<sup>3</sup> Xin Mei,<sup>1</sup> and Hang Dai<sup>1</sup>

<sup>1</sup>School of Cyberspace Security, Northwestern Polytechnical University, Xi'an 7100072, China

<sup>2</sup>Faculty of Information Technology, Monash University, Wellington Road, Clayton, VIC 3800, Australia

<sup>3</sup>School of Automation, Northwestern Polytechnical University, Xi'an 710072, China

Correspondence should be addressed to Libin Yang; libiny@nwpu.edu.cn

Received 16 June 2020; Revised 13 August 2020; Accepted 22 September 2020; Published 9 October 2020

Academic Editor: Huizhi Liang

Copyright © 2020 Nanxin 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.

Heterogeneous information network (HIN), which contains various types of nodes and links, has been applied in recommender systems. Although HIN-based recommendation approaches perform better than the traditional recommendation approaches, they still have the following problems: for example, meta-paths are manually selected, not automatically; meta-path representations are rarely explicitly learned; and the global and local information of each node in HIN has not been simultaneously explored. To solve the above deficiencies, we propose a tri-attention neural network (TANN) model for recommendation task. The proposed TANN model applies the simulated genetic algorithm to automatically select meta-paths at first. Then, it learns global and local representations of each node, as well as the representations of meta-paths existing in HIN. After that, a tri-attention mechanism is proposed to enhance the mutual influence among users, items, and their related meta-paths. Finally, the encoded interaction information among the user, the item, and their related meta-paths, which contain more semantic information can be used for recommendation task. Extensive experiments on the Douban Movie, MovieLens, and Yelp datasets have demonstrated the outstanding performance of the proposed approach.

## 1. Introduction

With the increasing amount of data on the Internet, users find it difficult to obtain useful information. In recent years, recommender systems which can only retain relevant information have received increasing attention [1–3]. Researchers proposed many methods to solve the recommendation task, which can be classified into four classes: neighborhood-based methods [4–7], model-based methods [8–11], graph-based methods [12–16], and deep neural network based methods [17–21]. Neighborhood-based methods contain user-based collaborative filtering and item-based collaborative filtering, this kind of methods utilize neighbor information of users to make prediction [22]. Model-based methods first construct a descriptive model using the users' preferences, and the recommendations are generated based on the model [23]. Graph-based methods not only consider the neighborhood information of

each node but also consider the network structure [16, 20]. Inspired by the great success of deep neural networks in computer vision and natural language process, recent researchers have exploited deep neural networks in recommendation task.

Besides, since multiple types of auxiliary information become available, many methods propose to use this information to improve the performance of recommendation [24, 25]. As auxiliary information has heterogeneity and complexity, it is challenging to leverage this information in recommender systems. Heterogeneous information network (HIN), containing various kinds of nodes connected by multiple types of relations, that can model rich auxiliary data has been applied in recommender systems [26, 27]. Although the existing HIN-based recommendation approaches enhanced the performance than the traditional recommendation approaches, they still have the following problems: first, the meta-paths are manually selected, not

automatically; second, meta-path representations are rarely explicitly learned; third, the global and local information of each node in HIN has not been simultaneously explored.

To solve the above deficiencies, we develop a tri-attention neural network (TANN) model for recommendation task. The proposed TANN model applies the stud genetic algorithm to automatically select meta-paths at first. Then, it learns global and local representations of each node, as well as the representations of meta-paths existing in the HIN. After that, a tri-attention mechanism is proposed to enhance the mutual influence among users, items, and their related meta-paths. Finally, the encoded interaction information among the user, the item, and their related meta-paths, which contain much semantic information, can be used for the recommendation task. To summarize, our major contributions are

- (1) Meta-paths are automatically selected via the stud genetic algorithm.
- (2) A novel tri-attention neural network (TANN) model is developed to enhance the influence among the user, the item, and their related meta-paths in a mutually reinforced way.
- (3) A TANN model-based recommendation approach is developed, and the extensive experiments on the Douban Movie, MovieLens, and Yelp datasets demonstrate the outstanding performance of the approach.

We organize the rest of this paper as follows. Section 2 reviews related work. Section 3 presents the preliminaries and notations in the paper. Section 4 illustrates the proposed tri-attention neural network (TANN) model-based recommendation approach. Section 5 illustrates the experimental results and Section 6 concludes the paper.

## 2. Related Work

Accurately finding useful information in many e-resources becomes more and more difficult for users, due to the development of information technology and the increasing content of information. However, a recommendation system can overcome this obstacle.

Alqadah et al. [28] utilized each user's local biclustering neighborhood and developed a collaborative filtering method. Yao et al. [29] applied a clustering analysis and latent factor model to enhance the neighborhood-based recommendation's performance. Neighborhood-based systems use the stored ratings to make a recommendation, whereas model-based approaches learn a predictive model by using the ratings. Cremonesi et al. [30] presented a PureSVD-based matrix factorization method, which uses the user-item rating matrix's most principle singular vectors to describe users and items. Pan et al. proposed a consensus factorization based framework for coclustering networked data [31]. Sindhvani et al. [32] developed a weighted nonnegative matrix factorization method. Xiong et al. propose an information propagation-based social recommendation method (SoInp) and model the implicit user

influence from the perspective of information propagation [33]. Hofmann [34] utilized PLSA (Probabilistic Latent Semantic Analysis) to solve collaborative filtering and showed that the PLSA is equivalent to nonnegative matrix factorization.

Most graph-based recommendation approaches are based on a random walk [35]. Christoffel et al. [12] proposed a graph random walk based recommendation algorithm. Kang et al. [16] presented a graph-based top-n recommendation model, which not only considers the neighborhood information encoding by user graph and item graph, but also takes into account the data's hidden structure.

Since deep learning techniques have been successfully applied in speech recognition and computer vision, some researchers began to utilize deep learning techniques in recommendation task and found that deep learning based recommendation approaches achieve better results than the conventional recommendation approaches. Oord et al. [36] applied deep convolution neural networks to generate songs' latent factors. Wang and Yang [37] combined probabilistic graphical models and deep belief networks to simultaneously learn audio content's features and make personalized recommendations. Xue et al. [38] presented a deep matrix factorization approach to solve the top-n recommendation task. Kim et al. [39] integrated convolution neural network into probabilistic matrix factorization and proposed a context-aware hybrid model for the recommender system.

In recent years, researchers adopted the heterogeneous information network (HIN) which characterizes rich auxiliary data in recommender systems. Pham et al. [40] modeled the rich information based on the constructed heterogeneous graph to solve the recommendation task. Chen et al. [41] developed a heterogeneous information network based projected metric embedding approach for link prediction. Yu et al. [42] presented a recommendation model based on a constructed attribute-rich HIN. Jiang et al. [43] modeled the user preferences using a generalized random walk with restart model and developed a heterogeneous information network based personalized recommendation method. Hu et al. [44] incorporated meta-path based context and proposed a co-attention mechanism based deep neural network to solve the recommendation task.

Although HIN-based deep learning approaches have achieved good performance in recommendation task, they usually select meta-paths manually, ignoring how to automatically select meta-paths. Moreover, the existing approaches seldom consider the interaction between node information and meta-path information. Our work applies the stud genetic algorithm to automatically select meta-paths, proposes a tri-attention mechanism that considers interactions among user-item-meta-path triplets, and develops a recommendation approach to further enhance the recommendation performance.

## 3. Preliminaries and Notations

We use the definitions of heterogeneous information network (HIN) and meta-path in [45].



*Definition 1.* A heterogeneous information network is an information network, which contains various kinds of objects and various kinds of links. It is defined as  $G=(V, A, E, R, W)$ , where  $V$  is the set of different types of vertices,  $A$  is the object type set,  $E$  is the union of different types of links,  $R$  denotes the link type set, and  $W$  is the union of the weight on each link. An edge  $e \in E$  is defined as  $e_{ijr} = (v_i, r, v_j) v_i, v_j \in V, r \in R$ .

*Definition 2.* A meta-path  $P$  is defined as a path in the form of  $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} A_{l+1}$  (abbreviated as  $A_1, A_2, \dots, A_{l+1}$ ), which describes a composite relation  $R = R_1 \circ R_2 \circ \dots \circ R_l$  between  $A_1$  and  $A_{l+1}$ , where  $\circ$  denotes the composition operator on relations.

The meta-path  $P$ 's length is the number of relations contained in  $P$ . Taking the user-movie network as an example, we can use a 4-length meta-path to describe the user-movie relation such as user  $\xrightarrow{\text{have seen}}$  movie  $\xrightarrow{\text{seen by}}$  user  $\xrightarrow{\text{have seen}}$  movie, or short as UMUM.

*Definition 3.* A path instance is a sequence of entity nodes; it is an explicit path in a meta-path, that is,  $p \in P$ .

## 4. Tri-Attention Neural Network (TANN) Model-Based Recommendation

In this section, we present the proposed tri-attention neural network (TANN) model at first and then illustrate the TANN model-based recommendation approach.

*4.1. TANN Model.* The whole architecture of the proposed TANN is presented in Figure 1.

*4.1.1. Embeddings for Users and Items.* In order to make the users' and items' representations more meaningful, we propose global representation to represent coarse-grained features of users and items, and develop local representation to represent fine-grained features of users and items; then, we integrate global information and local information of each node in HIN.

(1) *Global Representations of Users and Items.* Following [46], we use a lookup layer to map the users' and items' one-hot representations to low-dimensional dense vectors. Given a user-item pair  $\langle \mathbf{u}, \mathbf{i} \rangle$ , let  $\mathbf{l}_u \in \mathbb{R}^{|\mathcal{U}| \times 1}$  and  $\mathbf{z}_i \in \mathbb{R}^{|\mathcal{I}| \times 1}$  denote their one-hot representations.  $\mathbf{L} \in \mathbb{R}^{|\mathcal{U}| \times d}$  and  $\mathbf{Z} \in \mathbb{R}^{|\mathcal{I}| \times d}$  represent the lookup layer's corresponding parameter matrices, which preserve users' and items' information; the user embedding's and item embedding's number of dimension is denoted as  $d$ ; and  $|\mathcal{U}|$  and  $|\mathcal{I}|$  denote the users' number and items' number, respectively. We apply HIN2VEC algorithm [47] to obtain the matrix  $\mathbf{L}$  and  $\mathbf{Z}$ . The user  $u$ 's and item  $i$ 's global representations are represented as

$$\mathbf{x}_u = \mathbf{L}^T \cdot \mathbf{l}_u, \quad (1)$$

$$\mathbf{y}_i = \mathbf{Z}^T \cdot \mathbf{z}_i. \quad (2)$$

(2) *Local Representations of Users and Items.* As each user can be represented as a sequence of item and each item can be represented as a sequence of user, we learn local representations of users (items) according to the corresponding item (user) sequence. Here, we use  $\mathbf{S}_{n(u)} \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{I}|}$  and  $\mathbf{S}_{n(i)} \in \mathbb{R}^{|\mathcal{U}| \times |\mathcal{U}|}$  to represent the sequence matrix of the user  $u$ 's and the item  $i$ 's neighbors, respectively. For each neighbor node in the sequence, we use one-hot representation to represent it;  $|\mathcal{I}_u|$  and  $|\mathcal{U}_i|$  are the number of  $u$ 's neighbors and the number of item  $i$ 's neighbors in HIN; and  $n(u)$  and  $n(i)$  denote the neighbor set of user  $u$  and item  $i$ , respectively. Then, we apply a lookup layer to obtain the low-dimensional vector of each node in the item (user) sequence of the user (item). After that, the local representation of the user (item) is obtained based on a neighbor attention mechanism, which can be described as follows:

$$\mathbf{x}'_u = (\gamma_{n(u)} \odot \mathbf{b}_{n(u)})^T \cdot \mathbf{1}, \quad (3)$$

$$\mathbf{y}'_i = (\gamma_{n(i)} \odot \mathbf{g}_{n(i)})^T \cdot \mathbf{1}, \quad (4)$$

where

$$\mathbf{b}_{n(u)} = \mathbf{S}_{n(u)} \cdot \mathbf{Z}, \quad (5)$$

$$\mathbf{g}_{n(i)} = \mathbf{S}_{n(i)} \cdot \mathbf{L}, \quad (6)$$

$$\gamma_{n(u)} = \text{softmax}(\mathbf{b}_{n(u)}), \quad (7)$$

$$\gamma_{n(i)} = \text{softmax}(\mathbf{g}_{n(i)}). \quad (8)$$

We concatenate the global representation vector and local representation vector of user  $u$  and item  $i$  and feed the concatenated vectors into MLP component to get the final representation, that is,

$$\tilde{\mathbf{x}}_u = \text{MLP}(\mathbf{x}'_u \oplus \mathbf{x}_u^T), \quad (9)$$

$$\tilde{\mathbf{y}}_i = \text{MLP}(\mathbf{y}'_i \oplus \mathbf{y}_i^T). \quad (10)$$

### 4.1.2. Meta-Path Embedding

(1) *Meta-Path Selection Items.* Assuming there exist  $M$  meta-paths in the heterogeneous information network  $G$ , we construct a phenotype matrix  $\mathbf{H}$  (the size of the matrix is  $C_M^X \times X$  ( $X \leq M$ )), representing all possible combinations of  $X$  meta-paths, where each row represents a meta-path. Then, we apply the stud genetic algorithm (SGA) [48] to automatically select optimal  $X$  meta-paths.

(2) *Meta-Path Instance Selection.* Traditional HIN embedding models mainly use simple random walk strategies to obtain path instances. However, the path instances obtained by such strategies are of low value and cannot be directly applied to the recommendation system. Therefore, we propose a weighted selecting strategy with priority. In each step, the walker considers that the next step should walk to a

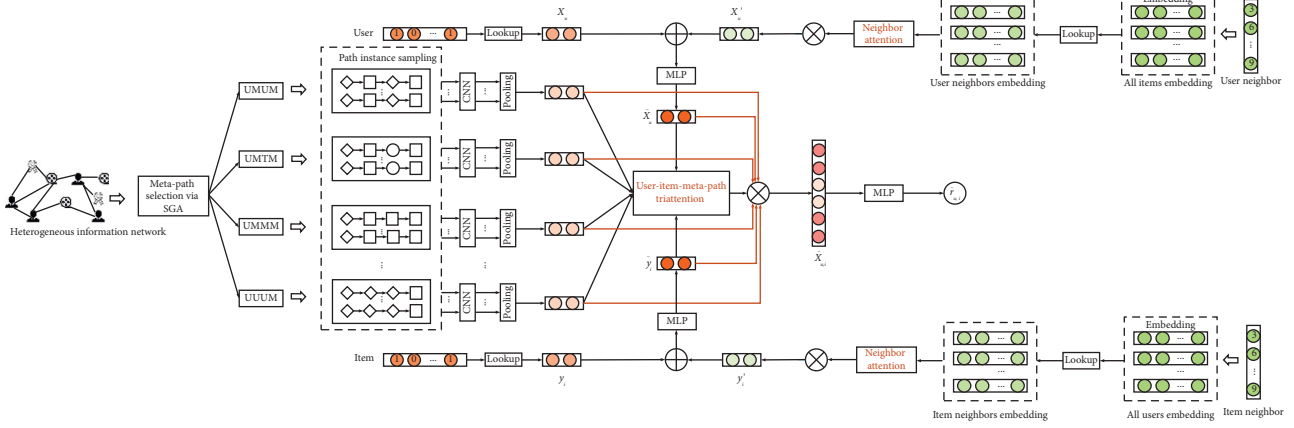


FIGURE 1: The whole architecture of the TANN model.

higher-priority neighbor, and using such walking strategy, a path instance which contains more semantic information can be obtained for recommendation task. Then, how to define the priority of each node in a sequence is a key problem.

Inspired by He et al. [46] and Hinton and Salakhutdinov [49], we use a similar pretraining technique to measure each candidate node's priority. The basic idea is to take the score between different nodes in the heterogeneous network as the weight allocation standard. For example, we define the score ranges from 1 to 5 in film evaluation; if the score of user  $u$  for movie  $i$  is 5, then we deem the weight value of the link between user  $u$  and movie  $i$  is the highest. For each node in HIN, it has a weight value for each node of its neighbors, and the similarity score between the node and the corresponding neighbor node can also be obtained. We measure the priority of each neighbor node by the product value between the weight value and the corresponding similarity score. Such score can reflect the correlation's degree between the two nodes. We use the above priority score strategy to construct each meta-path's instances.

Finally, for each meta-path, we obtain a different number of meta-path instances with a given length of  $L$  and then calculate scores of these meta-path instances as follows: for a path instance, we sum the product of weight and cosine similarity between adjacent nodes and then obtain that the sum value is divided by  $L$  as the score of the corresponding meta-path instance. So, we can obtain meta-path instances' scores of each meta-path and select the  $k$  path instances with high scores as the selected meta-path instance.

(3) *Meta-Path Instance Embedding.* Since a meta-path is a sequence of entity nodes, we apply the convolution neural network (CNN) to map a meta-path into a low-dimensional vector. For a meta-path  $P$ , we use  $\mathbf{X}^P \in \mathbb{R}^{L \times d}$  to represent the path embedding matrix, where  $p$  is a path instance,  $L$  represents the path instance's length, and  $d$  is the nodes' embedding dimension. The path instance  $p$ 's embedding is computed as follows:

$$\mathbf{h}_p = \text{CNN}(\mathbf{X}^P; \Theta), \quad (11)$$

where  $\Theta$  denotes CNN's related parameters.

(4) *Meta-Path Embedding.* As each meta-path contains many path instances, we first apply the proposed meta-path instance selection strategy to obtain each meta-path's top  $k$  path instances. Then, we use the maximum pooling operation to get important dimensional features from the selected path instances. Let  $\{\mathbf{h}_p\}_{p=1}^k$  denote  $k$  selected path instances' embedding. The embedding of the meta-path  $P$  is computed as

$$c_p = \max\text{-pooling}\left(\{\mathbf{h}_p\}_{p=1}^k\right). \quad (12)$$

(5) *Tri-Attention Mechanism Based Interaction Embedding.* As meta-paths contain rich semantic information, different users through different meta-paths show different preferences; even when the same user interacts with different items through the same meta-path, semantic information contained in the meta-path is also different. In order to better represent the semantic information existing among users, items, and meta-paths, we develop a tri-attention mechanism to assign different weights to different triplets of user-item-meta-path.

Given the user embedding  $\mathbf{x}_u$ , item embedding  $\mathbf{y}_i$ , meta-path embedding  $c_p$  which exists between the user  $u$  and item  $i$ , we use two full-connection layers to get the triattentive score as

$$\beta_{1P} = f(\mathbf{W}_1^u \tilde{\mathbf{x}}_u + \mathbf{W}_1^i \tilde{\mathbf{y}}_i + \mathbf{W}_1^P c_p + \mathbf{b}_1), \quad (13)$$

$$\beta_{2P} = f(\mathbf{w}_2 \beta_1 + b_2), \quad (14)$$

where  $\mathbf{W}_1^*$  is the first layer's weight matrix,  $\mathbf{b}_1$  is the first layer's bias vector,  $\mathbf{w}_2$  is the second layer's weight vector, and  $\mathbf{b}_2$  is the second layer's bias.  $f(\cdot)$  is set to the ReLU function.

We use softmax function to obtain the final interaction weights; that is,

$$\beta_p = \frac{\exp(\beta_{2P})}{\sum_{p' \in P_{u \rightarrow i}} \exp(\beta_{2P'})}, \quad (15)$$

where  $P_{u \rightarrow i}$  is the set of meta-paths existing between the user  $u$  and the item  $i$ . The interaction embedding among the user  $u$ , the item  $i$ , and their related meta-path  $P$  can be represented as

$$\tilde{x}_{u,i} = \sum_{p \in P_{u \rightarrow i}} \beta_p \cdot (\tilde{x}_u \oplus \tilde{y}_i \oplus \mathbf{c}_p), \quad (16)$$

where  $\oplus$  represents the vector concatenation operation.

**4.2. TANN Model-Based Recommendation Approach.** Once we obtain the interaction embedding, we apply an MLP component to model complicated interactions:

$$\tilde{r}_{u,i} = \text{MLP}(\tilde{x}_{u,i}), \quad (17)$$

in which the MLP contains two hidden layers with ReLU nonlinear activation functions and an output layer with sigmoid functions.  $\tilde{r}_{u,i}$  is interpreted as the relevance score between the user  $u$  and the item  $i$ , and  $\tilde{r}_{u,i}$  is used to generate a recommendation list for the user  $u$ .

Defining an appropriate objective function is a critical step for model optimization; following [14, 17], we use negative sampling to learn the model’s parameters:

$$\ell_{u,i} = -\log \tilde{r}_{u,i} - \sum_{j=1}^c \log(1 - \tilde{r}_{u,j}), \quad (18)$$

in which the first term is the observed interactions, and the second term indicates the negative feedback drawn from the noise distribution which is set to the uniform distribution (it can be set to other biased distributions), and  $c$  is the number of negative item sampling.

## 5. Experiment and Evaluation

**5.1. Datasets.** We use three datasets, that is, Douban Movie dataset (<https://github.com/librahu>), MovieLens movie dataset (<https://grouplens.org/datasets/movielens>), and Yelp business dataset (<https://www.yelp.com/dataset>). As a rating indicates whether a user has rated an item, we deem the rating as an interaction record [47, 50]. Table 1 lists the detailed description of the datasets. Each dataset’s first row shows the number of users, items, and their interactions, and the other rows list the other relations’ statistics. The manually selected meta-paths and SGA selected meta-paths of each dataset are listed in Table 2. Since long meta-paths may import noisy semantics [51], we set the length of meta-path to 4 and set the number of each meta-path’s selected path instances to 5.

**5.2. Evaluation Methods.** In order to evaluate the recommendation’s performance, each dataset’s user implicit feedback records are divided into a training set and test set, according to a certain proportion. For example, we use 90% feedback records to predict the remaining 10% feedback records. As it is a waste of time ranking each user’s items, especially for large datasets, for each user in the dataset, we randomly select 200 negative samples having no interaction records with the user at first. After that, we obtain each user’s

TABLE 1: Statistics of the three datasets.

Datasets	Relations (A-B)	#A	#B	#A-B
MovieLens	User-movie	943	1682	100,000
	User-user	943	943	47,150
	Movie-movie	1682	1682	82,798
	Movie-type	1682	18	2891
Douban movie	User-movie	1830	12337	697311
	User-user	1830	1830	9155
	Movie-movie	12337	12337	61690
	Movie-type	12337	38	12337
Yelp	User-business	300	450	15109
	User-user	300	300	1505
	Business-business	450	450	2255
	Business-category	450	47	450

recommendation list by ranking the positive items and negative items of the list. We use Pre@K, Recall@K, and NDCG@K to evaluate the experimental results.

When we apply the SGA algorithm to select the optimal 4 meta-paths, we follow the principle that “the smaller the objective function value is, the larger the fitness value is”; that is, we take negative evaluation scores. We implement the TANN model using TensorFlow with Keras (<https://keras.io/>). The batch size is set to 256, the regularization parameter is set to 0.0001, the learning rate is set to 0.001, and the dimension of user embedding and item embedding is set to 64. We apply Adaptive Moment Estimation (Adam) [52] to optimize the model.

**5.3. Performance of TANN-Based Recommendation.** To illustrate the benefits of applying the SGA algorithm, weighted random walk strategy, and integrating HIN’s global and local information of TANN model, we compare TANN against the following:

- (1) TANN w/o SGA, which applies manually selected meta-paths; we use the meta-paths in the second column of Table 2;
- (2) TANN w/o WRWS, which applies a random walk strategy (RWS) instead of weighted random walk strategy (WRWS) to select meta-path instances;
- (3) TANN w/o local, which considers global information of HIN only, ignoring local information, that is, it uses a global representation of users and items only, ignoring local representation of users and items.

Table 3 shows that the performance of TANN w/o SGA method is the poorest, which indicates that the selection of meta-paths has a great influence on the recommendation results. The performance of TANN w/o RWRS is better than the performance of TANN w/o SGA; it indicates that the meta-path selection strategy plays an important role compared to the weighted random walk strategy for recommendation task. In addition, as results illustrate, the performance of TANN w/o local is better than that of the above two methods. This can be mainly credited to the fact that the nodes and meta-paths in HIN contain rich implicit

TABLE 2: Manually selected meta-paths and the meta-paths selected by SGA in three datasets.

Datasets	Manually selected meta-paths	Meta-paths selected by SGA
MovieLens	UMTM, UUUM, UMUM, UMMM	UMTM, UUMM, UMUM, UMMM
Douban movie	UMMM, UMUM, UUUM, UUMM	UMTM, UMUM, UMMM, UUMM
Yelp	UUUB, UUBB, UBUB, UBCB	UUBB, UBUB, UBBB, UBCB

TABLE 3: Experimental results of TANN-based recommendation approach on the four datasets.

Model	MovieLens			Douban movie			Yelp		
	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10
TANN	<b>0.6968</b>	<b>0.3476</b>	<b>0.2250</b>	<b>0.9397</b>	<b>0.8082</b>	<b>0.1186</b>	<b>0.4285</b>	<b>0.1363</b>	<b>0.2554</b>
TANN <i>w/o</i> local	0.6957	0.3467	0.2242	0.9333	0.8042	0.1177	0.4184	0.1332	0.2488
TANN <i>w/o</i> RWRS	0.6901	0.3463	0.2235	0.9301	0.7878	0.1156	0.4120	0.1330	0.2442
TANN <i>w/o</i> SGA	0.6895	0.3451	0.2234	0.9087	0.7468	0.1094	0.4063	0.1293	0.2323

and effective information. TANN *w/o* local method applies SGA to automatically obtain meta-paths avoiding artificial interference, and it uses the weighted random walk strategy to get optimal meta-path instances that can represent the information of heterogeneous information network structure. TANN method, which not only considers information of users, items, and meta-paths but also considers the mutual influence among them, consistently outperforms the other three methods.

*5.4. Comparison with Other Recommendation Approaches.* Moreover, we compare our proposed TANN method with other four recommendation methods: (1) BPR (Bayesian Personalized Ranking) [53], which is a Bayesian posterior optimization based personalized ranking algorithm; (2) LRML (Latent Relational Metric Learning) [54], which employs an augmented memory model to construct latent relations between each user-item interaction; (3) CDAE (Collaborative Denoising Autoencoders) [55], which uses a Denoising Autoencoder structure to learn users and items’ distributed representations; and (4) MCRc (Meta-path based Context for REcommendation) [44], which leverages rich meta-paths and co-attention mechanism. Table 4 compares the experimental results.

From Table 4, we can see that LRML performs the poorest, as it only learns relations that describe each user-item interaction. CDAE learns the corrupted user-item preferences’ latent representations that can best reconstruct the full input, so it performs better than LRML. Both LRML and CADE concentrate on explicit feedback. BPR uses not only explicit feedback but also implicit feedback; thus, it can obtain better results than LRML and CADE. MCRc learns representations of users, items, and meta-path based context, as it encodes much more information than the previous methods; its performance is better than the above three methods. MCRc selects meta-path manually and only utilizes global information of users and items, while our proposed TANN method selects meta-path automatically, uses local and global information of users and items, applies a tri-attention mechanism to enhance the users,

items, and meta-paths’ representations, and its performance consistently outperforms the other four recommendation methods.

*5.5. Impact of Meta-Paths Selection.* In this set of experiments, we examine whether the automatically selected meta-paths can produce better recommendation performance than manually selected meta-paths. The experiments are conducted on the MovieLens dataset. The optimal meta-path set selected by SGA is UMTM, UUMM, UMUM, and UMMM (P1). We randomly selected three different meta-path sets: UMTM, UMUM, UUUM, and UMMM (P2); UMUM, UUMM, UUUM, and UMMM (P3); and UUMM, UUUM, UMTM, and UMMM (P4). The recommendation performance with different meta-path sets is shown in Figure 2.

We can observe in Figure 2 that the recommendation performance based on the meta-path set which is selected by SGA algorithm is better than the performance based on the three manually selected meta-path sets. The experimental results show that the interference of human factors should be avoided in the construction of heterogeneous information network.

*5.6. Impact of Users’ and Items’ Local Information.* We study whether incorporating users’ and items’ local information can further enhance the recommendation performance. The experiments are conducted on the Yelp dataset. We randomly select a user from the user list and obtain the recommendation list with/without local information of users and items. In Figure 3, the ground-truth of the user’s preference movie ids is listed in the middle column, the movie ids using TANN *w/o* local information method are listed in the left column, and the movie ids using TANN *w/i* local information method are listed in the right column. The bold ids in the left and right columns indicate the matched results with the ground-truth ids. We can see that when integrating local information, the recommendation method can find more accurate results than without local

TABLE 4: Comparison with other recommendation approaches on the three datasets.

Model	MovieLens			Douban movie			Yelp		
	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10
TANN	<b>0.6968</b>	<b>0.3467</b>	<b>0.2235</b>	<b>0.9397</b>	<b>0.8082</b>	<b>0.1186</b>	<b>0.4285</b>	<b>0.1363</b>	<b>0.2554</b>
MCRec	0.6909	0.3411	0.2192	0.9365	0.8080	0.1194	0.4212	0.1343	0.2435
BPR	0.6905	0.3386	0.2176	0.6129	0.2811	0.0393	0.4192	0.1297	0.2342
CDAE	0.6303	0.3039	0.1906	0.6008	0.2716	0.0374	0.3948	0.1167	0.2241
LRML	0.5381	0.2242	0.1418	0.4855	0.2649	0.0256	0.3934	0.1147	0.2124

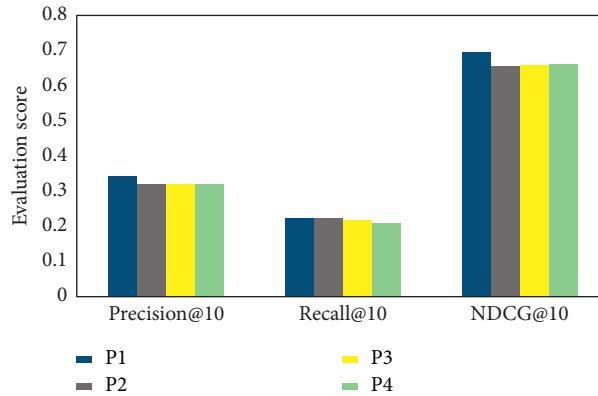


FIGURE 2: Recommendation performance with different meta-path sets.

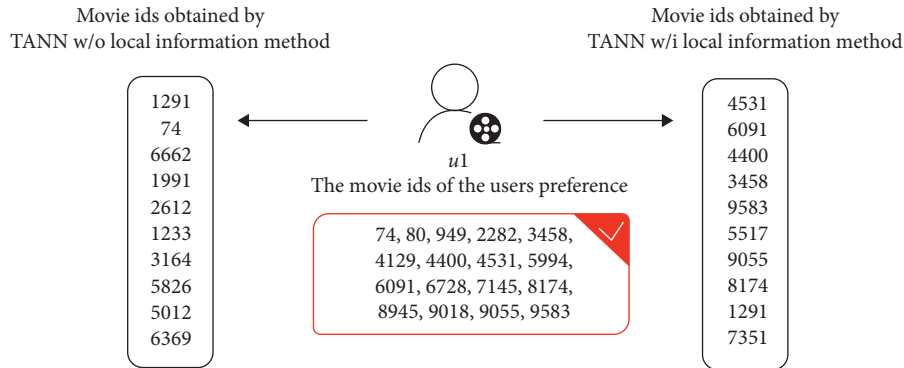


FIGURE 3: An example of removing local information in the TANN method.

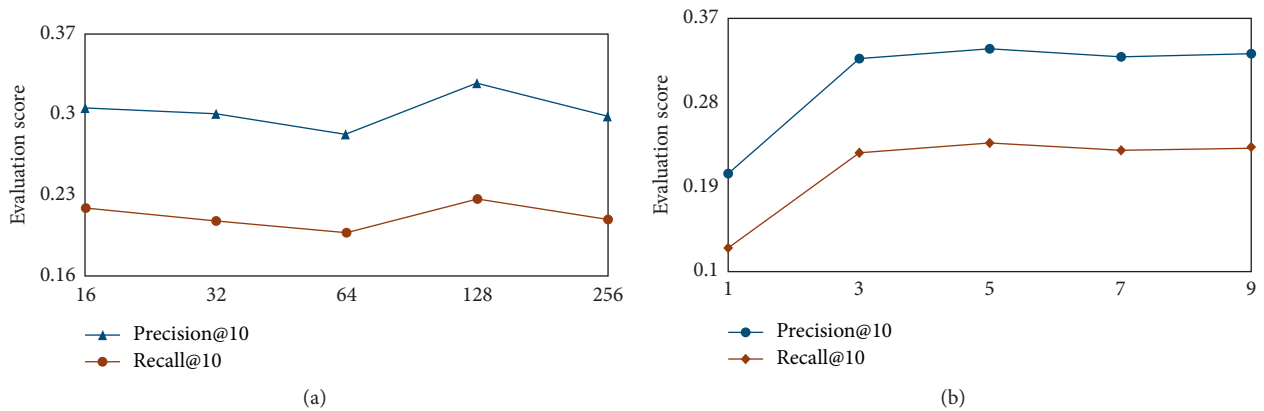


FIGURE 4: Continued.

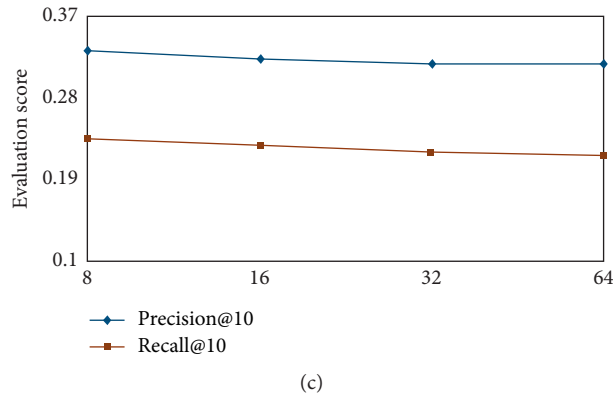


FIGURE 4: Parameter tuning of TANN on MovieLens dataset. (a) The embedding size for the weighting vector of attentive scores. (b) The number of negative samples. (c) The embedding size for the output layer.

information. It illustrates that local information has an impact on improving the recommendation result. Compared with the global information, the local information can reflect the neighborhood characteristics of the nodes centrally.

**5.7. Parameter Tuning.** Our model includes a few important parameters to tune. In this section, we examine three parameters' performance effect on MovieLens dataset, that is, the embedding size for the weighting vector of attentive scores (i.e., the embedding size of  $w_2$  in equation (14)), the negative samples' number (in equation (18)), and the output layer's embedding size.

For the embedding size of the tri-attention mechanism, we vary it in the set of {16, 32, 64, 128, 256}. As shown in Figure 4(a), our method achieves the best performance when it is 128. For the negative samples' number, we vary it in the set of {1, 3, 5, 7, 9}. We can find from Figure 4(b) that when 5 negative items are taken for each positive item, the evaluation result is the best. For the embedding size of the output layer, we vary it in the set of {8, 16, 32, 64}, and the best result can be obtained when the embedding size of the output layer is 8. The optimal performance is obtained with 128-dimension of the tri-attention mechanism, 5 negative samples, and 8-dimension of the output layer.

## 6. Conclusion

We present a tri-attention neural network (TANN) model for the recommendation in this paper. We first apply the stud genetic algorithm to automatically select meta-paths and propose a tri-attention mechanism to enhance the mutual influence among users, items, and their related meta-paths. Then, we encode the interaction information among the above three objects which contain more semantic information and can be used for the recommendation task. Extensive experiments on the Douban Movie, MovieLens, and Yelp datasets demonstrated the outstanding performance of the proposed approach. In the future, we will explore other auxiliary

information in the heterogeneous information network for the recommendation task.

## Data Availability

The authors have presented the website of the three datasets in Section 5.1. Readers can download the datasets from the corresponding link, or can contact the corresponding author.'

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (nos. 61872296 and 61772429), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (18YJC870001), and the Fundamental Research Funds for the Central Universities under Grant 3102019ZDHKY04.

## References

- [1] L. Sharma and A. Gera, "A survey of recommendation system: Research challenges," *International Journal of Engineering Trends and Technology*, vol. 4, no. 5, pp. 1989–1992, 2013.
- [2] M. Balabanović and Y. Shoham, "Fab: content-based, collaborative recommendation," *Communications of the ACM*, vol. 40, no. 3, pp. 66–72, 1997.
- [3] KG. SarwarBM and RJ. KonstanJA, "Application of dimensionality reduction in recommender system — a case study," in *Proceedings of the ACM WebKDD Web Mining for E-Commerce Workshop*, vol. 82, p. 90, Boston, MA, USA, 2000.
- [4] X. Ning, C. Desrosiers, and G. Karypis, *A Comprehensive Survey of Neighborhood-Based Recommendation Methods Recommender Systems Handbook*, pp. 37–76, Springer, Boston, MA, USA, 2015.
- [5] Z. Li, H. Chen, F. Xiong, X. Wang, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.

- [6] X. Ning and G. K. Slim, "Sparse linear methods for top-n recommender systems," in *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, pp. 497–506, Vancouver, Canada, 2011.
- [7] Y. C. Hu, "Recommendation using neighborhood methods with preference-relation-based similarity," *Information Sciences*, vol. 284, pp. 18–30, 2014.
- [8] C. Paolo, Y. Koren, and R. Turrin, "Performance of recommender algorithms on top-n recommendation tasks," in *Proceedings of the Fourth ACM Conference on Recommender Systems*, ACM, Barcelona, Spain, pp. 39–46, 2010.
- [9] S. Pan, X. Wang, H. Yang et al., "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018.
- [10] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *Proceedings of the International Conference on Data Mining*, pp. 263–272, Pisa, Italy, December 2008.
- [11] A. N. Nikolakopoulos, V. Kalantzis, E. Gallopoulos, and J. D. Garofalakis, "EigenRec: generalizing PureSVD for effective and efficient top-N recommendations," *Knowledge and Information Systems*, vol. 58, no. 1, pp. 59–81, 2019.
- [12] F. Christoffel, B. Paudel, C. Newell, and A. Bernstein, "Blockbusters and wallflowers: accurate, diverse, and scalable recommendations with random walks," in *Proceedings of the 9th ACM Conference on Recommender Systems*, pp. 163–170, Vienna, Austria, September 2015.
- [13] S. Pan, R. Hu, S. Fung et al., "Learning graph embedding with adversarial training methods," *IEEE Transactions on Cybernetics*, vol. 50, 2019.
- [14] C. Cooper, S. H. Lee, T. Radzik, and Y. Siantos, "Random walks in recommender systems: exact computation and simulations," in *Proceedings of the 23rd International Conference on World Wide Web*, ACM, New York, NY, USA, pp. 811–816, 2014.
- [15] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: representation, acquisition and applications," 2020, <http://arxiv.org/abs/2002.00388>.
- [16] Z. Kang, C. Peng, M. Yang, and Q. Cheng, "Top-N recommendation on graphs," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 2101–2106, Toronto, Canada, 2016.
- [17] A. M. Elkahky, Y. Song, and X. He, "A multi-view deep learning approach for cross domain user modeling in recommendation systems," in *Proceedings of the Twenty-Fourth International Conference on World Wide Web*, ACM, New York, NY, USA, pp. 278–288, 2015.
- [18] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, "Session-based recommendations with recurrent neural networks," in *Proceedings of the International Conference on Learning Representations*, Vancouver, Canada, 2016.
- [19] A. V. Den Oord, S. Dieleman, and B. Schrauwen, "Deep content-based music recommendation," *Neural Information Processing Systems*, pp. 2643–2651, 2013.
- [20] R. Hu, G. Long, C. Zhang et al., "Learning graph embedding with adversarial training methods," *IEEE Transactions On Cybernetics*, vol. 50, 2019.
- [21] H. Wang, N. Wang, and D. Yeung, "Collaborative deep learning for recommender systems," in *Proceedings of the Twenty-First ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, pp. 1235–1244, 2015.
- [22] X. Yao, B. Tan, C. Hu, W. Li, Z. Xu, and Z. Zhang, "Recommend algorithm combined user-user neighborhood approach with latent factor model," in *Proceedings of the International Conference on Mechatronics and Intelligent Robotics*, Springer, Cham, Switzerland, pp. 275–280, 2017.
- [23] H. Zhang, F. Min, and S. Wang, "A random forest approach to model-based recommendation," *Journal of Information and Computational Science*, vol. 11, no. 15, pp. 5341–5348, 2014.
- [24] G. mavicius and A. Tuzhilin, "Context-aware recommender systems," in *Proceedings of the 2008 ACM Conference on Recommender Systems*, RecSys 2008, Lausanne, Switzerland, pp. 191–226, October 2008.
- [25] R. Ronen, E. Yomtov, and G. Lavee, "Recommendations meet web browsing: enhancing collaborative filtering using internet browsing logs," in *Proceedings of the IEEE 32nd International Conference on Data Engineering*, pp. 1230–1238, Helsinki, Finland, 2016.
- [26] X. Yu, X. Ren, Y. Sun et al., "Personalized entity recommendation: a heterogeneous information network approach," in *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, pp. 283–292, New York City, NY, USA, 2014.
- [27] H. Zhao, Q. Yao, J. Li, Y. Song, and D. L. Lee, "Metagraph based recommendation fusion over heterogeneous information networks," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 635–644, Halifax, Nova Scotia, Canada, 2017.
- [28] F. Alqadah, C. K. Reddy, J. Hu, and H. F. Alqadah, "Biclustering neighborhood-based collaborative filtering method for top-n recommender systems," *Knowledge and Information Systems*, vol. 44, no. 2, pp. 475–491, 2015.
- [29] Z. Yao, J. Wang, and Y. Han, "An improved neighborhood-based recommendation algorithm optimization with clustering analysis and latent factor model," in *Proceedings of the 38th Chinese Control Conference*, pp. 27–30, Guangzhou, China, 2019.
- [30] P. Cremonesi, Y. Koren, and R. Turrin, "Performance of recommender algorithms on top-n recommendation tasks," in *Proceedings of the Fourth ACM Conference on Recommender Systems*, ACM, New York, NY, USA, pp. 39–46, 2010.
- [31] T. Guo, S. Pan, X. Zhu, and C. Zhang, "CFOND: consensus factorization for co-clustering networked data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 4, pp. 706–719, 2018.
- [32] V. Sindhwani, S. S. Bucak, J. Hu, and A. Mojsilovic, "One-class matrix completion with low-density factorizations," in *Proceedings of the 2010 IEEE International Conference on Data Mining, Ser. ICDM '10*, IEEE Computer Society, Washington, DC, USA, pp. 1055–1060, 2010.
- [33] F. Xiong, W. Shen, H. Chen et al., "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, 2019.
- [34] T. Hofmann, "Latent semantic models for collaborative filtering," *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 89–115, 2004.
- [35] S. Baluja, R. Seth, D. Sivakumar et al., "Video suggestion and discovery for youtube: taking random walks through the view graph," in *Proceedings of the 17th International Conference on World Wide Web*, pp. 895–904, Beijing, China, 2008.
- [36] A. Van den Oord, S. Dieleman, and B. Schrauwen, "Deep content-based music recommendation," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 2643–2651, Lake Tahoe, NE, USA, 2013.
- [37] X. Wang and Y. Wang, *Improving Content-Based and Hybrid Music Recommendation Using Deep Learning*, pp. 627–636, ACM Multimedia, Seattle, WA, USA, 2014.

- [38] H. J. Xue, X. Dai, J. Zhang, S. Huang, and J. Chen, "Deep matrix factorization models for recommender systems," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 3203–3209, Melbourne, Australia, August 2017.
- [39] D. Kim, C. Park, J. Oh, and H. Yu, "Deep hybrid recommender systems via exploiting document context and statistics of items," *Information Sciences*, vol. 417, pp. 72–87, 2017.
- [40] T.-A. N. Pham, X. Li, G. Cong, and Z. Zhang, "A general recommendation model for heterogeneous networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 12, pp. 3140–3153, 2016.
- [41] H. Chen, H. Yin, W. Wang, H. Wang, Q. V. H. Nguyen, and X. Li, "PME: projected metric embedding on heterogeneous networks for link prediction," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1177–1186, London, UK, August 2018.
- [42] X. Yu, X. Ren, Y. Sun et al., "Recommendation in heterogeneous information networks with implicit user feedback," in *Proceedings of the 7th ACM Conference on Recommender Systems*, pp. 347–350, Hong Kong, China, 2013 October.
- [43] Z. Jiang, H. Liu, B. Fu, Z. Wu, and T. Zhang, "Recommendation in heterogeneous information networks based on generalized random walk model and bayesian personalized ranking," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pp. 288–296, Los Angeles, CA, USA, October 2018.
- [44] B. Hu, C. Shi, W. Zhao, and P. Yu, "Leveraging meta-path based context for top- N recommendation with a neural co-attention model," *Knowledge Discovery and Data Mining*, pp. 1531–1540, 2018.
- [45] Y. Sun, J. Han, X. Yan, S. P. Yu, and T. Wu, "PathSim: meta path-based top-K similarity search in heterogeneous information networks," in *Proceedings of the 37th International Conference on Very Large Data Base*, Seattle, Washington, September 2011.
- [46] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. Chua, "Neural collaborative filtering," in *Proceedings of the 26th International Conference on World Wide Web*, pp. 173–182, Perth, Australia, April 2017.
- [47] C. Desrosiers and G. Karypis, "A comprehensive survey of neighborhood-based recommendation methods," in *Recommender Systems and Book*, pp. 107–144, Springer, Boston, MA, USA, 2011.
- [48] W. Khatib and P. J. Fleming, "The stud ga: a mini revolution?" in *Proceedings of the 5th International Conference on Parallel Problem Solving from Nature*, Springer, New York, NY, USA, pp. 683–691, 1998.
- [49] G. E. Hinton and R. Salakhutdinov, "A better way to pretrain deep Boltzmann machines," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 2447–2455, Lake Tahoe, NE, USA, 2012.
- [50] Y. Koren, "Factor in the neighbors," *ACM Transactions on Knowledge Discovery from Data*, vol. 4, no. 1, pp. 1–24, 2010.
- [51] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in *Proceedings of the 2008 ACM Conference on Recommender Systems*, pp. 23–25, Lausanne, Switzerland, October 2008.
- [52] D. Kingma and J. Ba. Adam, "A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, CA, USA, 2015.
- [53] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: bayesian personalized ranking from implicit feedback," 2012, <http://arxiv.org/abs/1205.2618>.
- [54] Y. Tay, L. Anh Tuan, and S. C. Hui, "Latent relational metric learning via memory-based attention for collaborative ranking," in *Proceedings of the 2018 World Wide Web Conference*, pp. 729–739, Lyon, France, April 2018.
- [55] Y. Wu, C. DuBois, A. X. Zheng, and M. Ester, "Collaborative denoising auto-encoders for top-n recommender systems," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 153–162, San Francisco, CA, USA, 2016.



## Research Article

# An Adaptive Network Model to Simulate Consensus Formation Driven by Social Identity Recognition

Kaiqi Zhang <sup>1,2</sup>, Zinan Lv,<sup>3</sup> Hai Feng Du <sup>2,4</sup> and Honghui Zou<sup>3</sup>

<sup>1</sup>School of Economy and Management, Chang'an University, Xi'an, China

<sup>2</sup>Center for Public Management and Complexity Science Research, Xi'an Jiaotong University, Xi'an, China

<sup>3</sup>College of Transportation Engineering, Chang'an University, Xi'an, China

<sup>4</sup>School of Public Policy and Management, Xi'an Jiaotong University, Xi'an, China

Correspondence should be addressed to Hai Feng Du; haifengdu@mail.xjtu.edu.cn

Received 1 May 2020; Revised 9 July 2020; Accepted 11 September 2020; Published 8 October 2020

Academic Editor: Hongshu Chen

Copyright © 2020 Kaiqi 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.

Models of the consensus of the individual state in social systems have been the subject of recent research studies in the physics literature. We investigate how network structures coevolve with the individual state under the framework of social identity theory. Also, we propose an adaptive network model to achieve state consensus or local structural adjustment of individuals by evaluating the homogeneity among them. Specifically, the similarity threshold significantly affects the evolution of the network with different initial conditions, and thus there emerges obvious community structure and polarization. More importantly, there exists a critical point of phase transition, at which the network may evolve into a significant community structure and state-consistent group.

## 1. Introduction

With the development of the Internet, the method of information spreading mainly depends on online social media, which promotes the formation of consensus in the complex network [1–4]. Understanding and analyzing the consensus formation can help the explanation and governance of social phenomena, such as public state polarization [5], rumor spreading [6] or riots, segregation, and interactions in cities [7]. The latest research on complex networks shows that consensus is the result of the coevolution of network structure and individual state. In addition, the formation process of consensus can be revealed by an adaptive network [8, 9].

An adaptive network model is a dynamically changing model that can self-organize network connections and update individual states. Models of the consensus in social systems have been the subject of a considerable amount of recent research studies in the physics literature [10]. These models can be categorized into two classes. One focuses on the role of network structure or individual state in the dynamic evolution process. Holme and Newman [10]

(hereinafter referred to as the Holme model) proposed a model based on the well-known Voter model [11, 12]. Their model combines state dynamics with assortative network formation, revealing an apparent phase transition between regimes in which one process or the other dominates the dynamics [10]. The other tries to grasp the dynamic interaction of the individual state. A popular model was proposed by Kozma and Barrat [4] (hereinafter referred to as the Kozma model), who investigated how the coevolution of an adaptive network of interacting individuals and the individual's states influences each other and how the final state of the system depends on this coevolution. Their model is based on the Deffuant model [13] for which a large number of states can coexist (and not only 2 as in the Voter model) [4].

It is observed in these models that network structure, coevolution rule, and distribution of state among social individuals significantly affect the final structure of the network and distribution of state. Furthermore, we find that these models undergo a continuous phase transition, from a regime in which the state is arbitrarily diverse to one in which most individuals hold the same state. Also, it can be controlled by adjusting the relevant parameters of the model [2, 3, 14–16].

It can be found that most models try to reveal a general law of consensus by constructing an adaptive network. State convergence can be observed in social media where social networks tend to divide into several groups or communities of individuals with similar states. In addition, adaptive rewiring of links can happen in real-life systems such as acquaintance networks where people are more likely to maintain a social connection if their views and values are similar. An obvious question to ask is why individuals want to change their states or rewire their social connections. In another phrase: what are their initial needs and underlying motivations to change state or adjust structure? However, the Holme model cannot reflect the mutual interaction between individuals in social media. It requires prior knowledge, which depends on certain network structure and individual state information, as a part of coevolution rules. Improved models like the Kozma model focus on simple evolution rules that do not require prior knowledge of the states of individuals to which new links are established. But in their models, all individuals have the same behavior choices in response to different group effects. In other words, individuals are homogeneous, and their heterogeneity is not well reflected.

Up to now, few studies however considered the fact that individuals have different responses when they are in different groups or even when they are in the same group, they may still make different choices because they are heterogeneous. In addition, in the individual's cognition of the message they received, the "herding effect" is particularly considered. It is believed that individuals will not only make judgments based on their rational expectations but also make decisions based on public behavior. Therefore, we propose a model that focuses on the realistic needs of individuals in a specific environment and the impact of behavioral motivation. To better understand the formation of consensus, social norms and individual psychology are considered in our model. Specifically, we introduce the social identity theory explaining social comparison and self-categorization when individuals face behavior choice. Social comparison and self-categorization both can give rise to realistic levels of agreement between acquaintances. Here, we combine both processes with a single parameter—*similarity threshold*—controlling the balance of the two processes. We investigate the role of various parameters such as the modularity and the rate of the largest group. We show that the network structure and evolution rules have important consequences on the evolution mechanisms of consensus formation.

The contributions of our work are threefold. (1) There is only one free parameter (similarity threshold), which can highlight the heterogeneity of individuals in the coevolution, in our model. This parameter has a clear meaning in social governance, which is a practice like disclosing and dispelling false information on social media, policy guidance, and behavioral intervention. (2) We make coevolution rules based on social identity theory, individuals' behavior choices are highly influenced by local relationships, and majority state in a group has a significant effect on individuals. (3) Based on the BA network, we use similarity threshold to

explain the polarization and formation of states in social media.

The organization of the paper is as follows. In Section 2, the relevant issues about the basic rules of the existing adaptive network model and social identity theory are discussed. To give a full description of the model, we start our study with the improved evolution rule and methods in Section 3. Then, we discuss the evolution process and experiment results in Section 4. Finally, Section 5 gives the conclusions of this paper.

## 2. Related Work

### 2.1. Adaptive Network Model Based on Structural Coevolution.

The adaptive network addresses the self-organization of complex network structure and its implications for system behavior [17–19]. Over the past decades, adaptive networks have been developed and applied to various subjects, ranging from physical, biological to social and engineered systems. Applications of adaptive networks also include the evolution of an organizational network, information/knowledge/culture sharing, and trust formation within a group or corporation [20, 21]. Our study of adaptive networks is in the field of consensus of which models proposed by Holme et al. [10] and Zanette et al. [22] are seminal works.

According to the setting of the Holme model, consider a network of  $N$  nodes, representing individuals, joined in pairs by  $M$  links, representing social connections between individuals. Each individual is assumed to hold one of  $K$  possible states on some topic of interest. The state of individual  $i$  is denoted as  $k_i$ . In each iteration of the model simulation, nodes will perform two basic operations according to the rules: (1) the links are randomly selected and placed between nodes with the same or similar state or (2) the node changes its state to be consistent with the state of the surrounding nodes. Specific rules are shown in Figure 1.

Figure 1(a) represents the formation of new social connections between people of similar states. Figure 1(b) represents the impact of acquaintances on one another. However, the rewiring operation in Step (a) requires certain prior knowledge to guide it. Meanwhile, the one-way influence operation in Step (b) cannot reflect the process of interaction between nodes. Therefore, Kozma et al. [4] proposed an adaptive network model based on the continuous distribution view.

Based on the Holme model and affected by the Deffuant model [13], the Kozma model is proposed. At each time step  $t$ , two neighboring nodes are selected, and they communicate if their states are close enough, i.e., if  $|o(i, t) - o(j, t)| \leq d$  where  $d$  defines the tolerance threshold. In this case, the local communication tends to bring states even closer, according to the following equation:

$$\begin{aligned} o(i, t+1) &= o(i, t) - \mu[o(j, t) - o(i, t)], \\ o(j, t+1) &= o(j, t) + \mu[o(j, t) - o(i, t)], \end{aligned} \quad (1)$$

or an attempt to break the connection between  $i$  and  $j$  is made: if  $|o(i, t) - o(j, t)| > d$ , a new node  $l$  is chosen at random and the link  $(i, j)$  is rewired to  $(i, l)$ . The tolerance

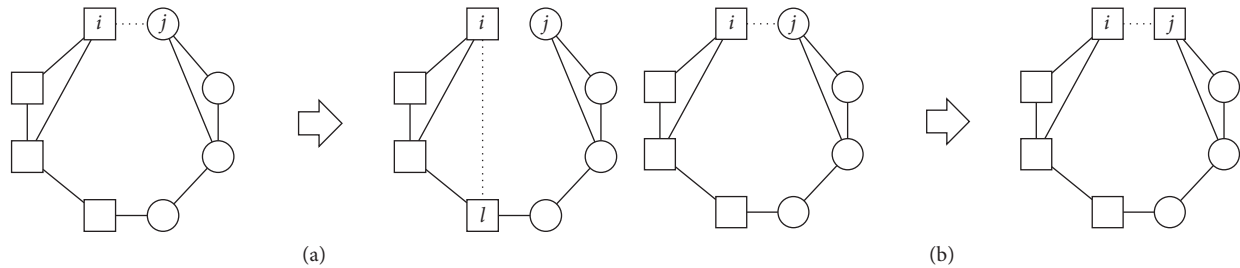


FIGURE 1: An illustration of the Holme model, with node shapes representing states. At each time step, the system is updated according to the process illustrated in (a) with probability  $\xi$  or (b) with probability  $1 - \xi$ . In (a), a node  $i$  is selected at random and one of its links (in this case the link  $(i, j)$ ) is rewired to a new node  $l$  holding the same state as  $i$ . In (b), node  $i$  affects one of its neighbor  $j$  and makes  $j$  adopt the state of it.

threshold  $d$  has an important impact on network regulation, and individual states tend to be homogeneous as  $d$  increases. Compared to the Holme model, the Kozma model emphasizes the microinteraction between nodes and controls the state convergence through tolerance threshold  $d$ . Also, the Kozma model emphasizes the randomness in the evolution to reduce the model's dependence on prior knowledge.

It can be found that interaction rule between nodes is the core part of model construction, and the interaction rules in all existing models contain both adjustments of node state and structure. Such adjustment of nodes would affect the whole network structure. However, in most models, the adjustment of one node is affected by another node, and the effect of local structure to the node is completely ignored. Taking the node  $j$  in Figure 1(b) as an example,  $j$  is consistent with the state of node  $i$  with probability  $1 - \xi$  in the Holme model. But according to the social norm, the group state has a great impact on individuals. So, the impact of  $i$  on  $j$  may be less than that of the other two neighbor nodes. With such a local structure, abstracting the behavior choice of  $j$  only based on the probability  $1 - \xi$  does not conform to reality. In the same case, in the Kozma model, the interaction between  $i$  and  $j$  can be influenced by other neighbor nodes. Moreover, the heterogeneity of nodes is also ignored. In real life, every individual is independent, and his/her behavior choice is affected by various factors such as personality, willpower, education, and so on. But the rewiring probability  $\xi$  in the Holme model or the tolerance threshold  $d$  in the Kozma model applies to all nodes, which means all nodes in their models are homogeneous. Therefore, existing models are not very similar to what we may observe in real life. The heterogeneity of individuals should be considered, and the interaction between individuals and groups affected by social norms needs to be designed with corresponding rules. So, related theories in sociology need to be introduced to improve related rules.

## 2.2. The Formation of Consensus in Social Identity Theory.

Social identity theory stems from the description of group formation and group relationship by ethnocentrism and realistic conflict theory [23]. Tajfel and Turner proposed social identity theory in 1986 to theorize how people conceptualize themselves in intergroup contexts and how a

system of self-categorizations “creates and defines an individual's place in society” [24, 25]. Tajfel defined *social identity* as “the individual's knowledge that he belongs to certain social groups together with some emotional and value significance to him of this group membership.” Social identity can not only help individuals to identify the same social attributes and integrate into specific social groups but also enable individuals to confirm their value realization. This value realization includes both the value of physical capital and the satisfaction of psychological cognition. Figure 2 shows the two stages of social identity: self-categorization and social comparison.

Self-categorization simplifies an individual's cognition of human society. Individuals usually use simple and independent attributes such as ethnicity, race, class, occupation, gender, and so on to describe things. They categorize themselves into a group with the same or similar attribute to reduce cognition of the complex world. Different attributes and state disagreements create group boundaries. Also, the group boundary formed in self-categorization is the precondition of the feature comparison which triggers the process of social comparison. The differences between groups are magnified in social comparison, thus forming positive heterogeneity. Pursuing positive heterogeneity is also a key part for individuals to realize their value and complete the evolution from individual behavior to consensus.

Overall, social identity enhances individual cognition and the positive heterogeneity formed in social identity is also the internal psychological driving force of individuals' tendency to group behavior. In the interaction between individuals and groups, the differences between individuals will gradually be weakened through the formation of social identity, which will also help to modify individual behavior choices, thereby forming a consensus. But existing models lack interaction rules between individuals and their groups, so we introduce social identity theory as an interaction rule of individual behavior. We design specific rules based on self-categorization and social comparison to better understand the internal mechanism of consensus.

## 3. The Proposed Model

In this section, we propose an adaptive network model describing the formation of consensus, which is based on social identity theory.

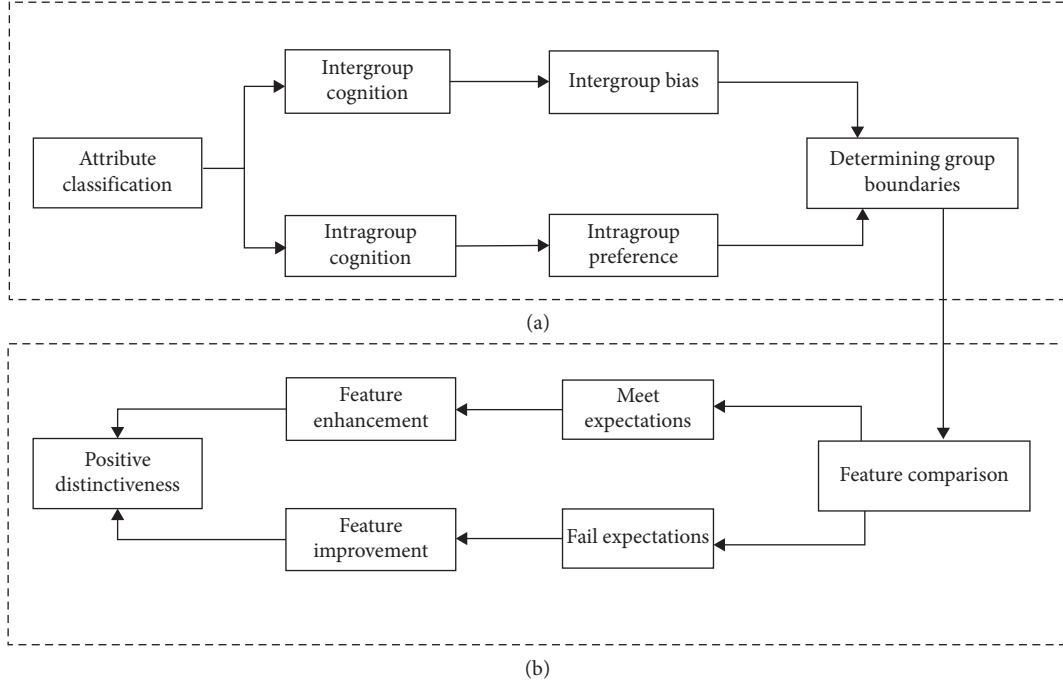


FIGURE 2: Formation of social identity. (a) The self-categorization process. Everyone can be subordinated to one or more social groups according to different attributes. Also, each group has the same attribute. At the same time, to distinguish the subordinate groups from other groups, an individual's cognition of the in-group similarity will continue to increase, and the out-group differences will continue to expand, thus forming obvious attributes and behavioral characteristics belonging to a certain group. (b) The social comparison process. It focuses on establishing positive heterogeneity for one's group. Positive heterogeneity is a precondition for the formation of social identity.

**3.1. Model Definition.** According to the above description of the classical social identity theory, social identity leads to consensus. There, we incorporate the social norm into the model rules and define it as a social state of individuals with similar states or attributes. Here, we set an  $r$ -dimensional vector abstracting the individual multiple states or attributes and represent it as  $\omega_i = \{a_1^i, a_2^i, \dots, a_r^i\}$ , where  $a_k^i$  represents the  $k$ th state of the individual  $i$ . The state difference between individuals can be expressed by the Euclidean distance:

$$d(i, j) = \sqrt{\sum_k^r (a_k^i - a_k^j)^2}. \quad (2)$$

To normalize equation (2) ( $d(i, j)$ ), there are

$$\text{sim}_{i,j} = \frac{1}{1 + d(i, j)}, \quad (3)$$

where  $\text{sim}_{i,j}$  represents the similarity between  $i$  and  $j$ . Formally,  $\text{sim}_{i,j}$  close to 1 indicates a higher similarity between individuals; otherwise, it indicates a lower similarity. From the perspective of social identity theory, the self-categorization of social individuals needs to judge the similarity between individuals according to specific social standards. A similarity threshold is introduced here to help with the self-categorization of individuals. The function of self-categorization is given by

$$\delta(i, j) = \begin{cases} 1, & \text{sim}_{i,j} \geq \theta, \\ 0, & \text{sim}_{i,j} < \theta, \end{cases} \quad (4)$$

where  $\theta$  is the similarity threshold. When  $\text{sim}_{i,j} \geq \theta$ , it is considered that individuals are homogeneous, and if there is a social interaction between  $j$  and  $i$ , it is determined that  $j$  is a homogeneous neighbor individual of  $i$ ; otherwise, it means that there is heterogeneity among individuals. Like the rewiring mechanism in the Holme model and the Kozma model, we abstract self-categorization as a relationship modification process. The establishment of homogeneous relationships is the goal of each individual, and we call this process the structural update mechanism (SU mechanism).

While updating the structure of social relations, strengthening individuals' cognition is an important part in the formation of consensus. Inspired by the Kozma model's interaction rules, we design a collective adaptive mechanism (CA mechanism) to abstract the strengthening of individuals' cognition. Specifically, it is assumed that set  $T_i$  represents a set of neighbor individuals of  $i$ . According to equation (4),  $T_i$  can be subdivided into a set of homogeneous individuals  $s_i = \{m \mid \delta(i, m) = 1\}$  and a set of heterogeneous individuals  $h_i = \{n \mid \delta(i, n) = 0\}$  ( $s_i \cup h_i = T_i$ ). In particular, individuals only update their states with neighbors in  $s_i$ , but not in  $h_i$ . In addition, the individual will update states according to the local average state of  $s_i$ , and the local average state is given by

$$\bar{\omega}_i = \left\{ \bar{a}_1, \bar{a}_2, \dots, \bar{a}_r \right\}; \bar{a}_k = \frac{\sum_{m \in s_i} a_k^m}{\|s_i\|}, \quad (5)$$

where  $\bar{a}_k$  is the average state of the  $k$ th state. Similar to the interaction rules in the Kozma model, individuals update their states one by one and update  $a_k^i$  to a new state  $a_k^{i'}$ :

$$a_k^{i'} = a_k^i + \mu [\bar{a}_k - a_k^i], \quad (6)$$

where  $\mu$  is the update parameter of the state; usually,  $\mu \in (0, 0.5]$ . Equation (6) shows that individual will update its states to make sure the convergence to the average state of its homogeneous neighbors.

The above two mechanisms restore the description of the self-categorization and social comparison process and present individual cognitive changes in the formation of consensus. Both mechanisms are influenced by the similarity threshold  $\theta$ . In social media, the greater the similarity threshold  $\theta$ , the more heterogeneous the relationships among individuals, and they are more likely to adjust with existing structures, such as unfollowing those who disagree with them. The smaller the similarity threshold  $\theta$ , the more homogeneous the relationships between individuals, and they are more likely to strengthen the individual's convergence to the average state.

Compared with the Holme model and the Kozma model, our model is based on social psychology and pays more attention to group pressure on individuals. Also, there are three differences in our mechanism. Firstly, we introduce the similarity threshold in the SU mechanism, where a local structure is updated by the homogeneity between individuals. It is more in line with the interaction of individuals in the real environment. Secondly, in the CA mechanism, we do not introduce relevant parameters such as the tolerance threshold  $d$  which has been talked in the Kozma model. Instead, the individual's cognitive update is realized by the average state of homogeneous neighbors, emphasizing the interaction between individuals and groups, reflecting the individual's convergence in the formation of consensus. Thirdly, different from the obedience principles in existing models, the convergence in our mechanism emphasizes the impact of local homogeneous individuals, and update rule about the heterogeneous relationships between individuals highlights the impact of social interaction on individual behavior choice. In summary, the mechanism proposed in this paper is derived from the general description of social identity theory. It abstracts the process of individuals classifying themselves into a group by different states and achieving consensus by defining group boundaries. With this process, we can analyze the possible behavioral strategies of individuals based on the specificity of consensus among different groups.

Table 1 presents a brief comparison of the three models. The evolution rules of our model will be introduced in detail below.

**3.2. Construction of Adaptive Network Model.** Based on the above theoretical modeling, we construct an adaptive network model to simulate the coevolution of the network

based on individuals' structures and states. Specifically, set network  $G = (V, E)$  represents social relationships between individuals, where  $V$  represents a set of nodes formed by  $X$  individuals, and  $L$  represents a set of connected links formed by  $Y$  social relationships. Also, the state of each node is a set of  $r$ -dimensional vectors  $\omega$ , and each state of the node  $a_k^i$  obeys a continuous distribution within a certain interval. There is only one controllable parameter, which is the similarity threshold  $\theta$ . Referring to the relevant settings of the existing model, the state update parameter  $\mu$  in equation (6) is set to a certain value. The update process of the model at time  $t$  is divided into three steps as shown in Figure 3:

Step 1: randomly select node  $v_i$  and calculate the similarity between  $v_i$  and its neighbors based on equation (3). At the same time, referring to the similarity threshold  $\theta$ , the neighbor nodes of  $v_i$  are divided into the homogeneous node-set  $s_i$  and the heterogeneous node-set  $h_i$ .

Step 2: when  $\|s_i\| \geq \|h_i\|$ , as shown in Figure 3(a), the number of homogeneous neighbor nodes of  $v_i$  is greater than that of heterogeneous neighbor nodes. The homogeneous neighbor nodes of  $v_i$  will calculate their average state based on equation (5); here, it is the average state of  $v_j$  and  $v_k$ . Nodes  $v_j$  and  $v_k$  will influence the state of  $v_i$  in the direction of the arrow. Node  $v_i$  updates its state based on equation (6).

Step 3: when  $\|s_i\| < \|h_i\|$ , as shown in Figure 3(b), the number of homogeneous neighbor nodes of  $v_i$  is smaller than that of heterogeneous neighbor nodes. It indicates that  $v_i$  is outside the group at the moment. According to the similarity  $s_i$  between nodes, use the roulette method to randomly select a node and achieve a rewiring process. It should be pointed out that the target node of the rewiring link is randomly selected.

Based on the above steps, the overall implementation of the model is shown in Algorithm 1.

Combined with the information dissemination process in social media, we can think that the impact operation in step 2 is actually the forwarding and dissemination of information in social media and the rewiring operation in step 3 is the shielding and attention of information.

## 4. Experiments and Results

In this section, we analyze the evolution process and results of the model based on a computer simulation experiment and discuss the impact of parameters.

**4.1. Experiment Settings and Parameters.** The simulation experiment is divided into two parts: (1) structure evolution experiments of the generated network model, focusing on the impact of parameters on structure evolution under different network structures; (2) state distribution evolution experiments of the generated network model, focusing on the impact of parameters on the distribution of node states under different network structures. For each set of

TABLE 1: Model comparison.

	Holme model	Kozma model	Our model
Theoretical basis	Behavior impact theory	Behavior impact theory; herd effect	Social identity theory
Steps and rules	Edges are randomly selected, and nodes perform rewiring operation with probability $\xi$ and influence operation with probability $1 - \xi$	With a probability $\omega$ , an attempt to break the connection between $i$ and $j$ is made if $ o(i, t) - o(j, t)  > d$	If the number of homogeneous neighbor nodes of a node is greater than that of heterogeneous nodes, do CA mechanism; otherwise, do SU mechanism
Decisive parameter	Rewiring probability: $\xi$	Tolerance threshold: $d$	Similarity threshold: $\theta$

This table compares the Holme model, Kozma model, and our model from three aspects: theoretical basis, steps and rules, and decisive parameters.

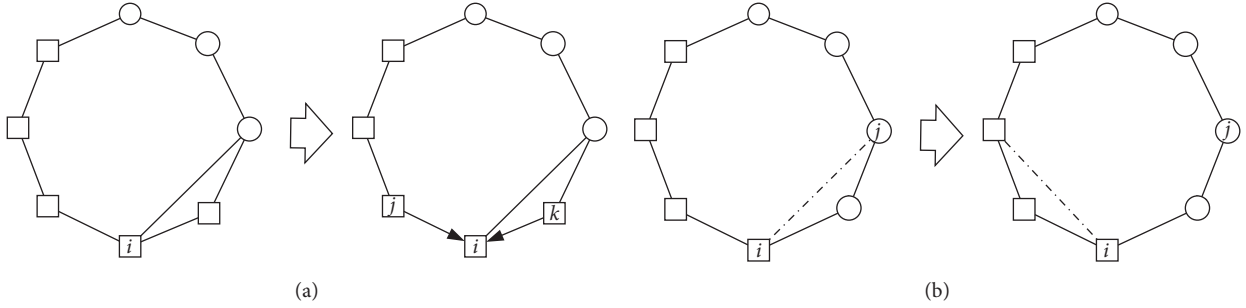


FIGURE 3: CA mechanism and SU mechanism at time  $t$  with node shapes representing states. At each time step, the system is updated according to the process illustrated in (a) or (b) with a precondition of the number of homogeneous (heterogeneous) neighbor nodes. When the number of homogeneous neighbor nodes of  $v_i$  is greater than that of heterogeneous neighbor nodes, as shown in (a), the state of  $v_i$  will be affected by its two homogeneous neighbor nodes. The direction of the arrow indicates the direction in which neighbor nodes influence it. Otherwise, as shown in (b),  $v_i$  will randomly select a node and achieve a rewiring process.

```

Input: Network:  $G = (V, E)$ , Node Status:  $\{\omega_i\}$ , Number of Iteration Times:  $it_{\max}$ , Similarity Threshold:  $\theta$ 
Output: Network:  $G' = (V, E)$ , Node Status:  $\{\omega_i\}$ 
(1) While  $t < it_{\max}$ 
(2)   For each node  $u$  in  $V$ :
(3)     Obtain the neighbor nodes set  $T$ 
(4)     For each node  $v$  in  $T$ :
(5)       Calculate the similarity between node  $u$  and node  $v$ :  $\text{sim}_{u,v}$ ;
(6)       If  $\text{sim}_{u,v} \geq \theta$ :
(7)          $s_u \leftarrow \text{node } v$ ;
(8)       Else:
(9)          $h_u \leftarrow \text{node } v$ ;
(10)    End For
(11)    If  $\|s_u\| \geq \|h_u\|$ :
(12)      Do Collective Adaptive Mechanism
(13)      For each  $a_k^{\text{mean}} \leftarrow \sum a_k^n / \|s_u\|$ ,  $n \in s_u$ ;  $a_k^n \in \omega_n$ 
(14)      For each  $a_k^u \leftarrow a_k^u + \mu [a_k - a_k^u]$ 
(15)    Else:
(16)      Do Structural Update Mechanism
(17)      Random Select node  $n \in h_u$  by Roulette
(18)      Rewiring link  $l(u, n) \rightarrow l(u, l)$ 
(19)    End If
(20)  End For
(21)   $t++$ 
(22) End While

```

ALGORITHM 1: Simulated social identity algorithm.

experiments, there are two different initial networks, which are the ER network and the BA network. The reason we use the BA network as an experimental benchmark network is that the relationship structure of most social media platforms such as Twitter and Facebook is a scale-free network [1]. Specifically, the parameters of the generated network model are shown in Table 2:

Based on the generated network, the experiment randomly assigns the initial state of nodes. Then, simulation experiments are carried out based on the model parameters. The initial parameters are shown in Table 3:

The state update parameter  $\mu$  ( $\mu \in [0, 0.5]$ ) is a convergence parameter. In order to simplify the experiment, we make  $\mu = 0.5$ , which also shows that nodes  $v_i$  and  $v_j$  will adopt the same intermediate opinion after communication. This setting is consistent with the Kozma model. The parameters related to the network structure, such as the average degree or network density, and the parameters related to the state of the node, such as state dimension  $r$  or state distribution, will definitely affect the experimental results, but they are essentially independent of the evolution rules.

The similarity threshold  $\theta$  is the only controllable parameter in the model. We focus on it to discuss the simulation results. Specifically, let the similarity threshold change in the range of  $[0.00, 1.00]$  in the step of 0.05. According to the description in social identity theory, the formation of consensus is also accompanied by the formation of groups, and there are obvious boundaries between groups. To distinguish the group boundaries in network structure and node state, our experiments will measure the groups that may generate from the structure and state, respectively. For the structure groups, it is expressed as community structure [26], and modularity [27] is used to measure the community structure; the calculation is given by

$$Q = \frac{1}{2L} \sum_{ij} \left( a_{ij} - \frac{g_i g_j}{2L} \right) \sigma(C_i, C_j), \quad (7)$$

where  $a_{ij}$  is adjacency relations between  $v_i$  and  $v_j$ ;  $g_i$  and  $g_j$  represent the degree of the two nodes;  $L$  is the number of edges in the network; and  $\sigma(C_i, C_j)$  indicates whether nodes  $v_i$  and  $v_j$  belong to the same community. If so,  $\sigma(C_i, C_j) = 1$ ; otherwise,  $\sigma(C_i, C_j) = 0$ . According to Newman's description of modularity, when  $Q \geq 0.3$ , it is assumed that the network has a community structure. There are many methods to detect community structure. In this paper, we use the method proposed in [28] to detect the community structure. It can detect the community structure based on the local attributes of nodes, which is suitable for the design in our rules about node states and local structures.

For the state groups, we distinguish them by the Euclidean distance  $d_0^i$  between its states' point and the coordinate origin based on the equation (2). By statistical sorting of  $d_0^i$ , a set of nodes within one standard deviation is regarded as a group, and the characteristics of state group is described by the ratio of the largest group's node size  $S_{\max}$  to the network node size  $N$ . Formally,  $S_{\max}/N$  is closer to 1 which indicates that there is only one unique state group in the network; otherwise, there may be multiple state groups.

TABLE 2: Parameters of the generated network model.

	ER network	BA network
Node size	$N = 1000$	$N = 1000$
Edge size	$M = 5018$	$M = 4975$
Model parameter	$p_c = 0.01$	$m_i = 5$

$p_c$  in the ER network represents the probability of connecting links in the network.  $M$  is the average link size of multiple experiments. In the BA network,  $m_i$  represents the number of connected links of the new node.

TABLE 3: The initial parameters of the model.

Parameters	Value
Node states: $a_k^i$	$U [1, 10]$
Node states dimension: $r$	4
Model iterations: $it$	2000
Update parameter of state: $\mu$	0.5

To highlight the difference between node states, the value of the node state  $a_k^i$  is set to a uniform continuous distribution in the range of  $[1, 10]$ . At the same time, the dimension of the node state is limited to 4 to effectively distinguish states between nodes. The time complexity of the model implementation is affected by the size of the link  $M$ . Therefore, the number of iterations is set to 2000 to ensure the effect of simulation. Also, update parameters of state  $\mu$  are set to 0.5 along with the existing model's setting.

The experimental environment is Python 3.7, and the operating environment is Intel(R) Core(TM) i7-4790 CPU @ 3.60 GHz with 24.0 GB RAM. Since the model contains random distribution parameters, all the results are the average results of 30 runs under the same parameter setting.

## 4.2. Experiment Results of Structure Evolution

**4.2.1. Analysis of the General Evolution Process.** We conduct experiments on the modularity of the network during model evolution. Figure 4 shows the modularity changes during the evolution of the networks.

In general, with the iteration in our model, the modularity of the two networks continues to increase. It indicates that our model can realize the evolution of the network structure and promote the formation of community structure, which is a structure group. Specifically, the SU mechanism in our model can effectively adjust the local structure of nodes, thereby continuously strengthening the community structure and realizing the structure aggregation of individuals. It can be found that the modularity of the ER network is bigger than that of the BA network. We think it is related to the initial network structure. Although the modularity values differ under different similarity thresholds, the evolution law of the model is relatively consistent.

**4.2.2. Impact of Similarity Threshold on Structure Evolution.** We further investigate the impact of the similarity threshold on the structure evolution.

As shown in Figure 5, the modularity value is generally affected by similarity threshold  $\theta$ , and there exists an obvious phase change characteristic. For the two networks, when the similarity threshold is around 0.35, the community structure of the network is most obvious, and the structure group is

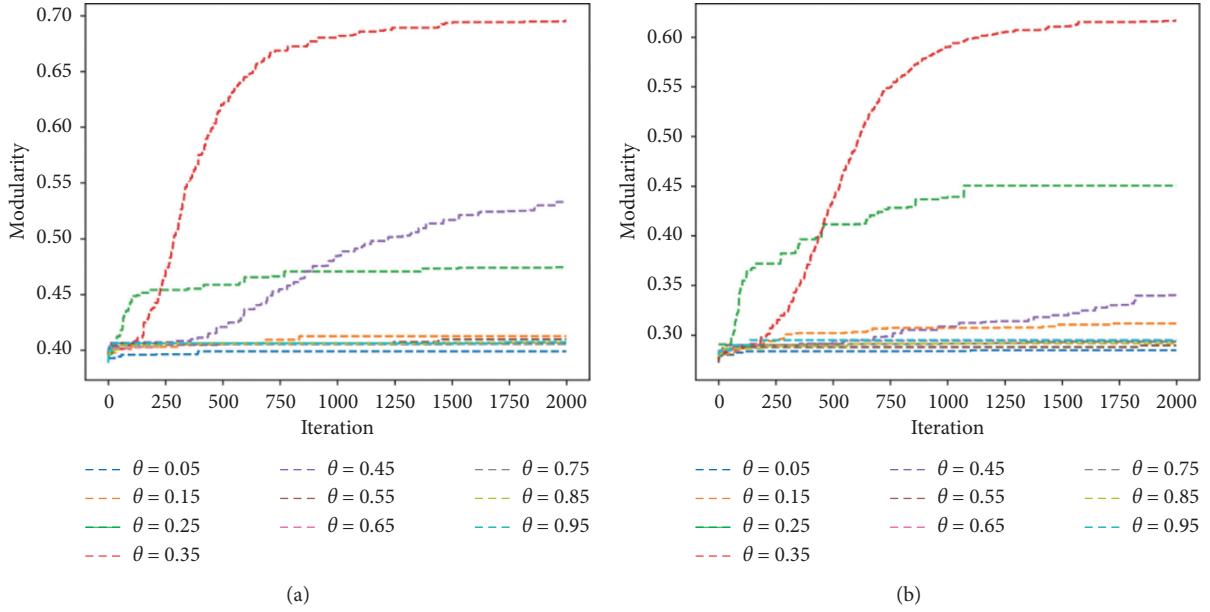


FIGURE 4: Evolution results of two different networks under different similarity thresholds. (a) Evolution process of BA network. (b) Evolution process of ER network.

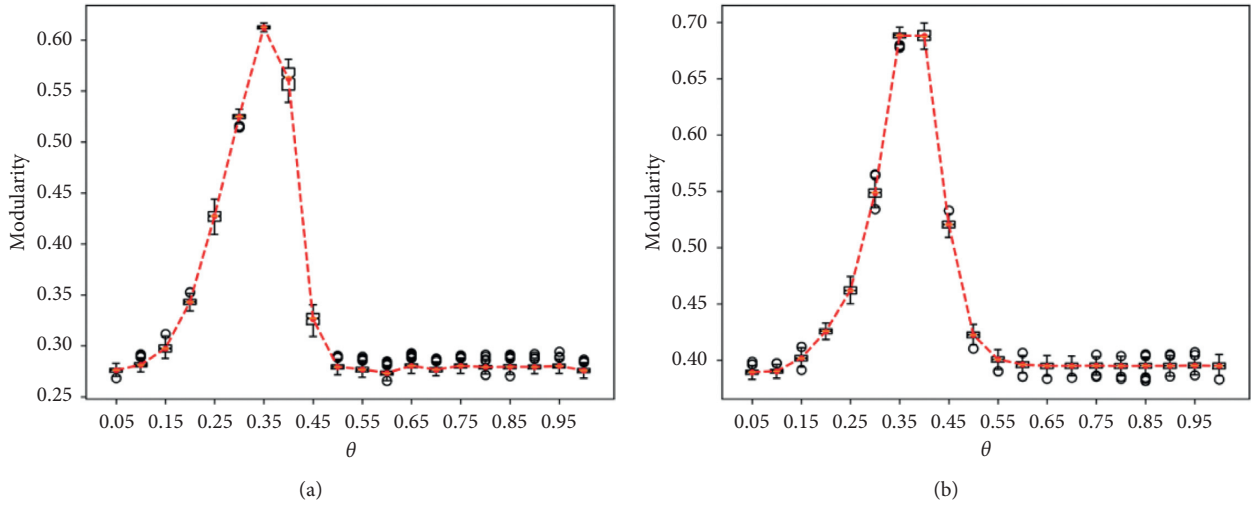


FIGURE 5: The modularity of the final network structure under different similarity thresholds. The modularity of the (a) ER network and (b) BA network.

relatively prominent. In general, the modularity of the evolved networks exhibits evolution characteristic which increases at first, then falls back, and tends to be stable as similarity threshold  $\theta$  increases. At the same time,  $\theta = 0.35$  and  $\theta = 0.45$  are two obvious phase transition points.

It can be found that the similarity threshold has a significant impact on the evolution of the structure group. The modularity phase change is concentrated at the interval of  $[0.15, 0.45]$ , and when the similarity threshold is outside the interval, the modularity is stable. It indicates that random initial network structure and random initial node state have less impact on the structure group of the evolved network. In

other words, via the model, the final structure of the network can be well controlled and will form a stable structure group. Combining equation (4) which is the definition of similarity threshold, similarity threshold controls the number of homogeneous nodes and heterogeneous nodes in the local structure. The larger the similarity threshold, the more heterogeneous nodes in the local structure, and thus the node is more likely to change its current structure. Otherwise, the node has more homogeneous nodes and it is easier to form the convergence of state. However, from the results in Figure 5, larger similarity threshold does not lead to a more obvious community structure or structure group. On



the contrary, that situation occurs when the similarity threshold is more moderate (i.e., around  $\theta=0.35$  in the experiment). From the perspective of social identity theory, the similarity threshold can be understood as the evaluation criteria of individual differences in society. The higher the evaluation criteria (the greater the similarity threshold), the greater the social distance between individuals, and individuals are more likely to form stable social relationships and social structures. The lower the evaluation criteria, the smaller the social distance between individuals, and individuals are also likely to form stable social relationships. It can be found that there may be a phase change threshold in the evaluation criteria. When the threshold is exceeded, individuals' need for consensus may increase rapidly, and they will form independent groups in structure, which may trigger consensus.

**4.2.3. Impact of Similarity Threshold on the Network Degree Distribution.** We further investigate the impact of the similarity threshold on the degree distribution of the BA network.

It can be seen in Figure 6 that without the intervention of the similarity threshold, the degree distribution of the BA network shows a power-law distribution. Under the logarithmic coordinate system, the degree distribution of the BA network will eventually show a lognormal distribution. When the similarity threshold is small ( $\theta < 0.15$ ), the degree distribution is close to a power-law distribution. As the similarity threshold increases to two phase transition points ( $\theta=0.15$  and  $\theta=0.35$ ), the power-law structure is broken and gradually forms a lognormal distribution; with the similarity threshold exceeding 0.35, the structure of the degree distribution tends to be stable.

From the perspective of social identity theory, everyone has equal opportunities to reach consensus with others, thus forming a group. The similarity threshold can be understood as a social barrier. When the similarity threshold is low, social barriers are relatively low; as long as there is a communication, people can reach a consensus, thus forming a whole. But the prerequisite is that people have equal opportunities to communicate. Therefore, it is difficult for them to form large groups, and they can only exist in society as a large number of independent small groups. This phenomenon appears in the area where we can find that the degree distribution of this segment appears as a power-law distribution. As the evaluation standard improves, due to differences in similarity, people begin to selectively combine into a union. Under such a situation, the power-law distribution is destroyed, and the large group splits into several small groups, gradually forming a lognormal distribution. From Figure 5, we can see that the modularity of this region is a growing trend, and there are still a large number of structure groups in the network, so the score of the fitting function will fluctuate at this time. When the similarity threshold exceeds the phase transition point ( $\theta=0.35$ ), modularity declines and stabilizes, and the degree distribution also presents a stable lognormal distribution.

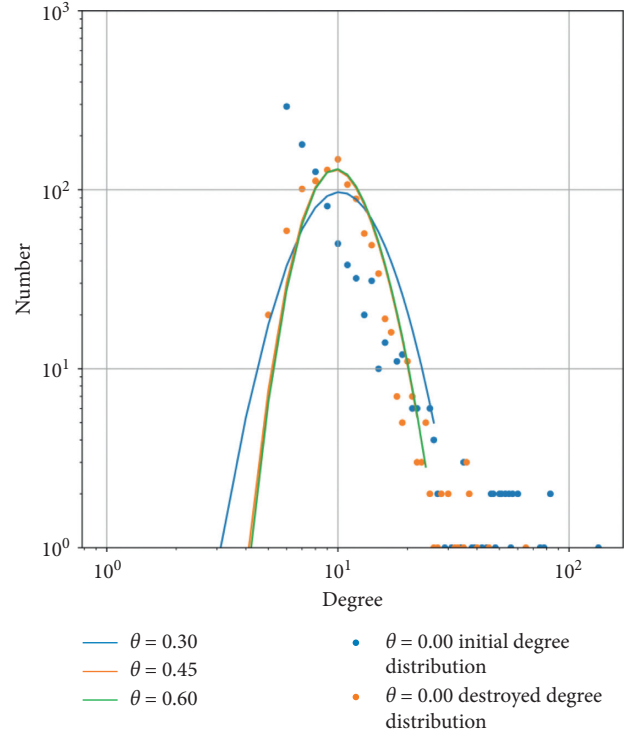


FIGURE 6: Degree distribution of the BA network in different similarity thresholds. We conducted many experiments on the BA network and obtained its degree distribution. We mainly selected the evolution results of 2000 iterations when  $\theta$  is equal to 0, 0.15, 0.3, 0.45, and 0.6. When  $\theta$  is greater than 0.6, the evolution result is similar to the result of  $\theta=0.6$ . The blue scatter plot shows the degree distribution when  $\theta=0$ , which is the degree distribution of the initial network. The green scatter plot is the degree distribution when  $\theta=0.15$ . A scatter plot indicates that it follows a power-law distribution. When  $\theta > 0.15$ , the original distribution is destroyed and gradually becomes a lognormal distribution, and we use a curve to represent the changing trend of this lognormal distribution. The blue, orange, and green curves, respectively, represent the degree distribution when  $\theta$  is equal to 0.3, 0.45, and 0.6.

### 4.3. Experiment Results of State Evolution

**4.3.1. Impact of Similarity Threshold on Node State.** Corresponding to the evolution of the structure group, the impact of the similarity threshold on node state has also been investigated. The  $S_{\max}/N$  values of the two networks under different similarity thresholds are shown in Figure 7. Results are the average of 30 independent experiments.

As shown in Figure 7, the similarity threshold has a significant impact on the evolution of node state and causes a phase change in the  $S_{\max}/N$  value. It can be found that the phase change point is concentrated near  $[0.15, 0.35]$ . In general, as the similarity threshold increases, the  $S_{\max}/N$  values of the two networks tend to be stable after large fluctuations. Figure 7(a) shows the state evolution of the ER network. With the increase of the similarity threshold,  $S_{\max}$  shows the change characteristics of increasing and then decreasing and gradually stabilizing. When the similarity threshold interval is  $[0.15, 0.35]$ , the size of the largest state group decreases from 0.97 to around 0.3, and with the

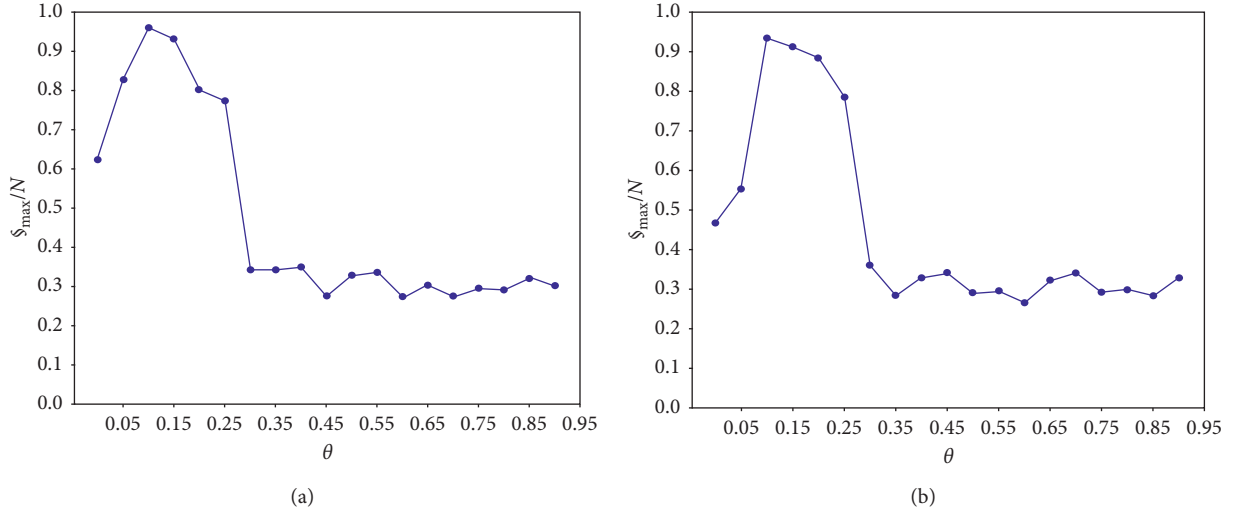


FIGURE 7: Distribution of  $S_{\max}/N$  under different similarity thresholds. The horizontal axis is the value of the similarity threshold, and the vertical axis is the  $S_{\max}/N$  value of the network at the last iteration. (a) ER network. (b) BA network.

increase of similarity threshold, the ratio is stabilized at around 0.3 after  $\theta \geq 0.35$ . It indicates that the initial network has a greater impact on the model when  $\theta$  is very small. For the ER network, the similarity threshold  $\theta$  has a significant impact on the state group at the interval of (0.15, 0.35). Influenced by the ER network structure, the size and structure of the state group are stable at other intervals. Figure 7(b) shows the state evolution of the BA network. The largest state population size  $S_{\max}$  in the network stabilizes first and has a sudden rise, then slumps, and gradually stabilizes with the increase of similarity threshold. When the similarity threshold  $\theta$  is at the interval of [0.10, 0.15], the size of the largest state group increases from 0.56 to around 0.93. With the continuous increase of similarity threshold,  $S_{\max}$  tends to decrease at the interval of [0.15, 0.35] and stabilizes around 0.3 after  $\theta \geq 0.35$ . For the BA network, as the similarity threshold increases, the node state in the network also converges and reaches the highest point at  $\theta = 0.15$ . However, a higher similarity threshold does not lead to convergence of node states, and only 30% of nodes are stable and will form a consistent state.

Based on the results of the above two networks, it can be found that the evolution law is generally consistent. All networks have extreme values at the interval of [0.15, 0.35], which means there exist obvious phase changes in this interval. It can be known from equation (6) that node state is determined by the average state of its local homogeneous neighbor nodes, and the number of homogeneous neighbor nodes is determined by equation (4). Therefore, the similarity threshold also affects the node state. Similar to the experimental result of the structure group, a higher or lower similarity threshold does not cause nodes to converge on the state, and there also exists a critical value to cause state convergence. From the perspective of social identity theory, it can be found that the evaluation criteria of individual differences will affect the scope of consensus if we take consensus as a state of social individuals, and there exist clear critical points in the evaluation criteria of differences, which

may lead to the formation of large-scale consensus. From the description of positive heterogeneity in social identity, the formation of large-scale consensus may lead to identity heterogeneity among different groups. When heterogeneity does not have positive characteristics, such as prejudice discrimination, popular and singular style, and so on, groups may adopt behavior strategies to form a consensus. In conclusion, the evaluation criteria of individual differences can be used as a risk indicator to estimate the occurrence of consensus.

*4.3.2. Analysis of the Evolution Results of Node States.* In Figure 7, when  $\theta = 0.15$ , the node state difference before and after evolution is the most significant. We specifically choose the evolution of the node state similarity matrix with  $\theta = 0.15$  as an illustration. As a comparison, we also show the experimental results of  $\theta = 0.35$ . Figure 8 shows the evolution of the state similarity matrix among the two sets of network nodes under different iteration steps.

Figures 8(a) and 8(b) show the evolution results of the state similarity matrix when  $\theta = 0.15$ . In general, these networks will generate a very large state group, which will diverge as the model evolves. Specifically, the evolution law of the ER network is shown in Figure 8(a). When the number of iteration steps is close to 400, a very large group of similar states has been formed, that is, large-scale consensus. With the further evolution of the model, the similarity within the group will increase, but at the same time, there will be differentiation. In the final stage of the model, the network will form a large-scale state group with similar overall states but partial differences. It illustrates that the homogeneity of individual states promotes the formation of large-scale state groups, but the heterogeneity of individual structures leads to overlap states within groups. The evolution law of the BA network is similar to the ER network. The network will form a very large state group in the final state of the model, but there will be differentiation within the group. It can be seen

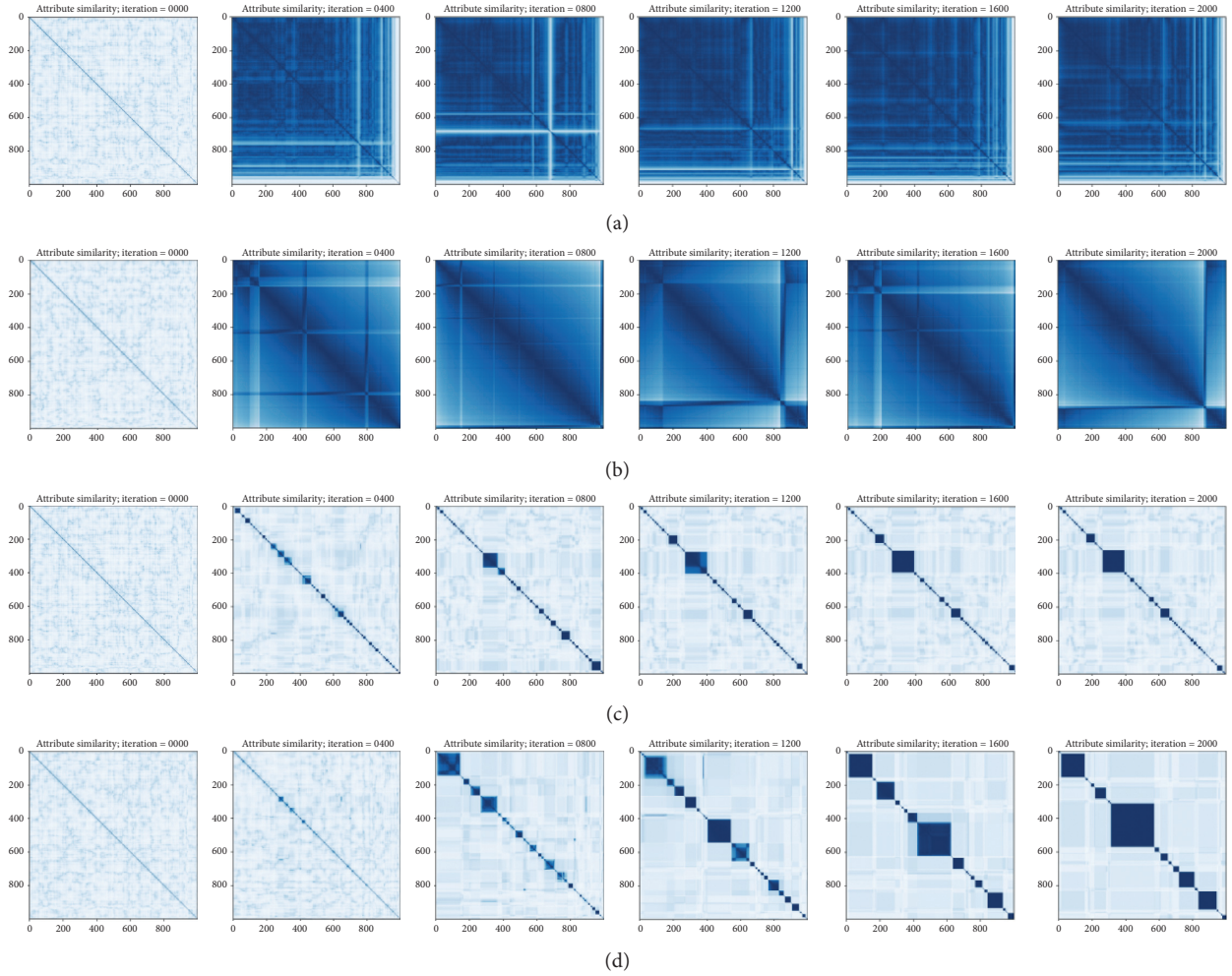


FIGURE 8: Evolution results of the state similarity matrix. The similarity matrix element  $\text{sim}_{i,j}$  calculates the state similarity between all nodes based on equation (2) and uses the color to represent the size of  $\text{sim}_{i,j}$ . The figure shows the state similarity matrix of the two networks when the number of iterations is 0, 400, 800, 1200, 1600, and 2000. (a) ER network ( $\theta = 0.15$ ). (b) BA network ( $\theta = 0.15$ ). (c) ER network ( $\theta = 0.35$ ). (d) BA network ( $\theta = 0.35$ ).

from Figure 7 that when  $\theta = 0.15$ , the proportions of the largest state group after the evolution of the two networks are close to 1. The large-scale state groups formed in Figures 8(a) and 8(b) have further verified the result of Figure 7.

Figures 8(c) and 8(d) show the evolution results of the state similarity matrix when  $\theta = 0.35$ . In general, as the model evolves, both networks will generate several state groups and eventually form several state groups with strong state consistency. Specifically, the evolution law of the ER network is shown in Figure 8(c). As the model evolves, several decreasing state groups will appear in the network. With the evolution of the model, some of the state groups will become larger, and some will merge. The final state of the model is to form multiple state groups of different sizes. It shows that when  $\theta = 0.35$ , there will not be a large-scale state consistency in the network. The evolution law of the BA network is similar to the ER network, and large-scale group consensus will not appear. It can be seen from the connection with Figure 7 that when  $\theta = 0.35$ , the proportion of

the largest state group in the network is only about 0.3, so there will be no large-scale group consensus in the final state of the model. It can be seen from Figure 8 that the network is divided into several state groups. At the same time, it can be seen from Figure 5 that it is divided into several structure groups. This shows that independent opinions are formed within each group and are structurally isolated from each other.

In summary, the similarity threshold in social media can be considered as an individual's tolerance for other people's opinions. The overall environmental tolerance in the virtual space is different from that of reality. Users can decide to follow or unfollow another user at any time, which makes the evolution of individual relationships more convenient. Individuals are more likely to form groups and exchange opinions within the group. From the existing research [1], we can know that the interaction between users in today's social media platforms such as Twitter and Facebook is very unequal. In the real world, the characteristics of relationships between celebrities and fans are similar to that of the

scale-free structure in the BA network. From the results of our experiments, we can know that the opinion polarization that currently appears in social media is essentially the evolution result of the network structure. Our experimental results verify this argument, and we find that the similarity threshold has a strong control effect on this process.

For the network degree distribution in the above process, it may evolve into a structure with the lognormal distribution. This shows that under the control of the similarity threshold, a number of parallel community structures appear among polarized groups, and this community structure changes the distribution of network degrees. To cite an instance, celebrities feed cultural products to people, and people come together as fanatics to negotiate against celebrities. This kind of tight organization brings together individuals who were unrelated before through an idol. Because they can form a community, the fan group's voice and status have been significantly improved. They are no longer ordinary audiences who were unable to speak in the dual relationship of "idol-followers," but become elders who can actively participate in the career planning of idols. In other words, the unequal power-law distribution in the original social media has changed after the emergence of the community, and it has become a lognormal distribution, which is also easy to trigger group isolation. Therefore, judging the evolution trend of the network degree distribution has certain guiding significance.

For the state distribution characteristics, it can be seen that the structure and state show different aggregation characteristics under different similarity thresholds. Generally speaking, when the threshold is around 0.15, there are obvious state groups in the network, and when the threshold is near 0.35, there are obvious structure groups. If there is a possibility that the similarity threshold may change dynamically, it indicates the inheritance relationship between the state group and the structure group. In social media, this means that if an individual's tolerance for other people's opinions slowly increases from a very low value, state groups will appear first and then structure groups.

## 5. Conclusions

In this paper, an adaptive network model is constructed based on social identity theory. Based on the Holme model and Kozma model, we improve the model evolution rules and parameters to make it more in line with reality. The formation of a structure group and state group has been analyzed. Also, we discussed the possible mechanism of formation and development of consensus in complex networks. It can be found that whether in the structure group or state group, the similarity evaluation criteria (similarity threshold) determine the final evolution result of the network, and the absolute high similarity or low similarity does not lead to differences between groups. On the contrary, when individuals tend to adjust their social relationship or state, there is often a clear threshold. Near this critical value, the network structure will be modularized or socialized, and individual states will form an obvious consensus. It indicates that the formation of consensus is often accompanied by

differentiation or isolation in social groups. It also shows that the formation and development of consensus may be affected by the similarity between individuals. Moreover, through the evolution of the similarity matrix, we found that after the emergence of several large state groups, a larger-scale consensus will be formed, and then disagreement will occur, which will lead to re-differentiation in the large group. We also found that affected by the similarity threshold, the degree distribution of the BA network will change from a power-law distribution to a lognormal distribution. In summary, our model realized the adaptation of network structure and node state by restoring the social identity process of individuals. The individual similarity may determine in which way individuals interact with groups and the evolutionary direction of consensus. In the real social environment, individuals' similarity belongs to a cognitive category, and its cognitive results can often be controlled through information disclosure, policy guidance, and behavioral intervention. In turn, it realizes the governance and regulation of public opinion and demonstrations. In addition, we find that when isolation occurs in the structure, opinions will also generate isolation. Therefore, whether the isolation of structure causes the differentiation of opinions or the differentiation of opinions causes structure isolation, we will study it in future research.

This study is still far from completion; our model simplified the general process of an individual's pursuit of consensus because of the complexity of individual behavior, and we did not consider the direction of interaction between social individuals and the sequence of individuals updating their states. This will be supplemented and improved in the subsequent work.

## Data Availability

The data used in our experiments are all generated using code in the Python 3.7 environment. The experiment requires two types of initial networks, ER network and BA network. In Section 4.1, we introduced the initial parameters of the generated networks, and readers can get the networks according to these parameter settings. If you have any further requirements, the data generation code 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 supported by the Natural Science Foundation, Shaanxi (grant no. 2019JQ-531), the Fundamental Research Funds for the Central Universities, CHD (grant no. 300102238103), and the Social Science Foundation, Shaanxi (grant no. 2019S044).

## References

- [1] H. Kwak, C. Lee, H. Park et al., "What is twitter, a social network or a news media?" in *Proceedings of the 19th International Conference on World Wide Web*, Raleigh, NC, USA, April 2010.
- [2] T. Evans and F. Fu, "Opinion formation on dynamic networks: identifying conditions for the emergence of partisan echo chambers," *Royal Society Open Science*, vol. 5, no. 10, Article ID 181122, 2018.
- [3] F. Fu and L. Wang, "Coevolutionary dynamics of opinions and networks: from diversity to uniformity," *Physical Review E-Statistical, Nonlinear, and Soft Matter Physics*, vol. 78, no. 1, 2008.
- [4] B. Kozma and A. Barrat, "Consensus formation on adaptive networks," *Physical Review E-Statistical, Nonlinear, and Soft Matter Physics*, vol. 77, no. 1, pp. 1–11, 2008.
- [5] T.-M. Kleiner, "Public opinion polarisation and protest behavior," *European Journal of Political Research*, vol. 57, no. 4, 2018.
- [6] S. Vosoughi, M. N. Mohsenvand, and D. Roy, "Rumor gauge," *ACM Transactions on Knowledge Discovery from Data*, vol. 11, no. 4, pp. 1–36, 2017.
- [7] R. Li, L. Dong, J. Zhang et al., "Simple spatial scaling rules behind complex cities," *Nature Communications*, vol. 8, no. 1, p. 1841, 2017.
- [8] T. Gross and H. Sayama, *Adaptive Networks Theory, Models and applications*, Springer, Berlin, Heidelberg, Germany, 2009.
- [9] H. Sayama, I. Pestov, J. Schmidt et al., "Modeling complex systems with adaptive networks," *Computers and Mathematics with Applications*, vol. 65, no. 10, pp. 1645–1664, 2013.
- [10] P. Holme and M. E. J. Newman, "Nonequilibrium phase transition in the coevolution of networks and opinions," *Physical Review E-Statistical, Nonlinear, and Soft Matter Physics*, vol. 74, no. 5, pp. 1–5, 2006.
- [11] C. Castellano, D. Vilone, and A. Vespignani, "Incomplete ordering of the Voter model on small-world networks," vol. 63, no. 1, pp. 153–158, 2003.
- [12] V. Sood and S. Redner, "Voter model on heterogeneous graphs," *Physical Review Letters*, vol. 94, no. 17, pp. 1–4, 2005.
- [13] J.-P. Nadal, G. Weisbuch, G. Deffuant et al., "Meet, discuss, and segregate!" *Complexity*, vol. 7, no. 3, pp. 55–63, 2002.
- [14] Y. Yu, G. Xiao, G. Li et al., "Opinion diversity and community formation in adaptive networks," *Chaos*, vol. 27, no. 10, pp. 1–12, 2017.
- [15] K. Bhawalkar, S. Gollapudi, and K. Munagala, "Coevolutionary opinion formation games," in *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing-STOC'13*, Palo Alto, CA, USA, June 2013.
- [16] R. Basu and A. Sly, "Evolving voter model on dense random graphs," *The Annals of Applied Probability*, vol. 27, no. 2, pp. 1235–1288, 2017.
- [17] H. Sayama, I. Pestov, J. Schmidt et al., "Modeling complex systems with adaptive networks," *Computers and Mathematics with Applications*, vol. 65, no. 10, pp. 1645–1664, 2013.
- [18] T. Gross and H. Sayama, *Adaptive Networks: Theory, Models and applications*, Springer, Berlin, Heidelberg, Germany, 2009.
- [19] T. Gross and B. Blasius, "Adaptive coevolutionary networks: a review," *Journal of the Royal Society Interface*, vol. 5, no. 20, pp. 259–271, 2008.
- [20] R. S. Burt, *Structural Holes: The Social Structure of Competition*, Harvard University Press, Cambridge, MA, USA, 1995.
- [21] V. Buskens and A. Van de Rijdt, "Dynamics of networks if everyone strives for structural holes," *American Journal of Sociology*, vol. 114, no. 2, pp. 371–407, 2008.
- [22] C. McGarty, V. Y. Yzerbyt, and R. Spears, "Social, cultural and cognitive factors in stereotype formation," *Stereotypes as Explanations*, Cambridge University Press, Cambridge, UK, 2002.
- [23] D. H. Zanette and S. Gil, "Opinion spreading and agent segregation on evolving networks," *Physica D: Nonlinear Phenomena*, vol. 224, no. 1-2, pp. 156–165, 2006.
- [24] H. Tajfel, *Social Identity and Intergroup Relations*, Cambridge University Press, Cambridge, UK, 2010.
- [25] M. A. Hogg, "A social identity theory of leadership," *Personality and Social Psychology Review*, Sage Publications, Los Angeles, CA, USA, 2001.
- [26] 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.
- [27] M. E. J. Newman, "Equivalence between modularity optimization and maximum likelihood methods for community detection," *Physical Review E*, vol. 94, no. 5, Article ID 052315, 2016.
- [28] K. Q. Zhang, H. F. Du, M. Cai et al., "Improved community structure detection algorithm based on the node's property," *Systems Engineering-Theory & Practice*, vol. 33, no. 11, pp. 2879–2886, 2013.

## Research Article

# A Bichannel Transformer with Context Encoding for Document-Driven Conversation Generation in Social Media

Yuanyuan Cai,<sup>1</sup> Min Zuo ,<sup>1</sup> Qingchuan Zhang ,<sup>1</sup> Haitao Xiong,<sup>1</sup> and Ke Li<sup>2</sup>

<sup>1</sup>National Engineering Laboratory for Agri-Product Quality Traceability and Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China

<sup>2</sup>Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Haidian District, Beijing, China

Correspondence should be addressed to Min Zuo; [zuomin1234@163.com](mailto:zuomin1234@163.com)

Received 7 June 2020; Accepted 30 July 2020; Published 17 September 2020

Guest Editor: Liang Wang

Copyright © 2020 Yuanyuan Cai 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.

Along with the development of social media on the internet, dialogue systems are becoming more and more intelligent to meet users' needs for communication, emotion, and social intercourse. Previous studies usually use sequence-to-sequence learning with recurrent neural networks for response generation. However, recurrent-based learning models heavily suffer from the problem of long-distance dependencies in sequences. Moreover, some models neglect crucial information in the dialogue contexts, which leads to uninformative and inflexible responses. To address these issues, we present a bichannel transformer with context encoding (BCTCE) for document-driven conversation. This conversational generator consists of a context encoder, an utterance encoder, and a decoder with attention mechanism. The encoders aim to learn the distributed representation of input texts. The multihop attention mechanism is used in BCTCE to capture the interaction between documents and dialogues. We evaluate the proposed BCTCE by both automatic evaluation and human judgment. The experimental results on the dataset CMU\_DoG indicate that the proposed model yields significant improvements over the state-of-the-art baselines on most of the evaluation metrics, and the generated responses of BCTCE are more informative and more relevant to dialogues than baselines.

## 1. Introduction

Dialogue systems such as Siri, Cortana, and Duer have been widely used to facilitate interactions between humans and intelligence devices as virtual assistants and social Chatbots. For example, people can conveniently make airline reservations with the help of an intelligent agent in social media. Conversational response generation, as a challenging task in natural language processing, plays a critical role in conversational systems.

Conversational generation aims to produce grammatical, coherent, and plausible responses in accordance with the input from users. Previous studies on dialogue generation mainly focus on either one-round conversation [1] or multiturn conversation [2]. One-round conversation tasks commonly determine responses on the basis of a single current query, while a multiturn conversation that consists of context-message-response triples commonly builds context-sensitive generators according to the dialogue history [2, 3]. Multiturn

conversation tasks tend to generate a variety of correlative responses in either goal-driven customer services [4–6] or chit-chat without predefined goals [7].

In the previous studies, the sequence-to-sequence (seq2seq) framework [8] with the attention mechanism has commonly been used to generate conversational responses and has achieved remarkable success in various domains [9–14]. The seq2seq models map a type of sequential syntactic structure to another without explicitly defining structural features by building an end-to-end neural network [2, 15]. Most seq2seq models use a recurrent neural network [16, 17] as the encoder and decoder to capture the sequential dependency. However, hierarchical recurrent neural networks, which suffer from time-consuming training, have difficulty in solving the problem of long-distance textual semantic dependency.

Since lacking of a knowledge background, the previous studies on conversational generation may suffer from the safe and generic responses such as “That’s all right” and “Yes.” The

uninformative responses are hard to match the relative content in the given document and satisfy the demand of users. To address this challenge, some knowledge-based methods have been proposed in recent works for conversational response generation. In these works, external knowledge is leveraged to facilitate conversation understanding and generation [18, 19], which includes structured data such as knowledge graphs [20, 21], unstructured textual knowledge [22], and visual knowledge [23]. With the development of the internet and big data, unstructured knowledge is more accessible than structured knowledge, which is constructed manually and depends heavily on the experience of experts. Therefore, some recent works take conversation-related documents and texts as the background knowledge to enrich useful information in conversations to generate more informative and interesting responses [24].

Our work is inspired by the recent success of the transformer framework [25], which is entirely based on attention mechanisms in end-to-end natural language processing tasks and eliminates complex recurrent and convolution network architectures [26]. We propose a transformer-based model for multiturn document-driven conversation. The proposed model encodes the conversational context and the current utterance, respectively. It also incorporates the multihop attention mechanism into the encoder and decoder to capture the correlative content for response generation, which draws global dependencies between documents, utterances, and responses. We conduct experiments on the conversation generation task regarding many metrics, including BLEU [27], METEOR [28], NW [24], and perplexity [19]. The experimental results indicate that the proposed model significantly outperforms the state-of-the-art methods. We also conduct ablation experiments to indicate the effects of the input elements fed into the encoder. The human judgment on various ablation models shows that the responses generated by the BCTCE model are more relevant to the context (document and dialogue history), more informative, and fluent than its several variants.

The contributions of this work are as follows:

We propose a novel BCTCE based on the transformer framework to build an encoder-decoder generator for document-driven conversation. The experimental results show that our model achieves new state-of-the-art performance.

The BCTCE learns the distributed representation of conversational context by encoding the document and dialogue utterances in parallel and integrating them within the interattention mechanism.

The BCTCE leverages layer-wise multihop attention mechanisms to gradually enhance the interaction between inputs, where the dialogue utterances and the document which can provide supplementary knowledge are used to generate the context-aware and dialogue-consistent responses.

The BCTCE can reduce the time of training and inference compared to the recurrent network-based response generators.

We review the related work in Section 2 and present the details of the proposed model in Section 3. Section 4 shows the experimental process, including datasets and evaluation criteria. The result analysis is also given in Section 4. Finally, we conclude this work and present future work with a brief summary in Section 5.

## 2. Related Work

Previous models for conversation are generally divided into rule-based, retrieval-based, and generation-based models. The rule-based and retrieval-based models depend on handcrafted rules or existing knowledge bases to match the correct answer, while the generation-based models require less manual effort by leveraging data-driven training of the algorithm on a noisy but large-scale corpus.

Recently, deep neural networks have been widely used for both response retrieval [29, 30] and response generation [31]. Some retrieval-based works determine the correct responses by the semantic similarity between the representations of a require and its candidate answers learned by neural networks. Sequence-to-sequence (seq2seq) frameworks [8] that have achieved success in many domains, such as machine translation [9, 32], have been commonly used for response generation [2, 15, 32]. In particular, seq2seq-based models play an important role in studies on multiturn conversation [4], which commonly build encoder-decoder networks for response generation. They map a sequential syntactic structure to another without explicitly defining features, where recurrent neural networks (RNNs) such as long short-term memory (LSTM) and gated recurrent units (GRUs) [33] are commonly employed as the kernel unit. Vinyals and Quoc explored the LSTM network to produce sequential responses end-to-end for the multiturn conversation [4]. Shang et al. combined global and local context information on the basis of the original RNN for a one-round conversation. Sordani et al. encoded the semantic information of the context and message by a multilayered nonlinear forward network and took RNN as a decoder to generate responses [3]. Chen et al. utilized a memory network to preserve more historical information in a multiturn dialogue [34]. RNNs are commonly used to sequentially encode each word in the input context and produce the response word-by-word during decoding. However, they were limited by the long time required for sequential training resulting from exploding or vanishing gradient. In addition, the model may suffer from information loss due to hardly capturing long-term semantic dependencies between utterances.

Attention mechanisms have become an integral part of sequence models in response generation, modeling the textual dependencies in the input or output sequences without regard to the position information [35, 36]. In previous works on neural response generation, the attention mechanism was incorporated into the encoder-decoder framework to preserve the key semantic information in sentences [1, 37]. Vaswani et al. proposed a neural network with self-attention and multihead attention to emphasize

different positions of a single sequence and the correlation between each word and its context [25]. Wu et al. proposed an utterance-level attention network combined with a word-level attention network to draw different important parts for the response [38]. Wu et al. proposed a multihop attention mechanism to learn a single context vector by repeatedly computing attention scores [37].

However, previous works merely produce general, rigid, and stylized responses without the natural variation in the language [39]. To address this issue, some studies have proposed context-aware conversational generators to produce more diverse and meaningful responses. Li et al. improved the LSTM-based generator by simply taking maximum mutual information as the objective function [15]. Xing et al. proposed a topic-aware neural generator that leverages topic information to simulate prior knowledge of humans by a joint attention mechanism and a biased generation probability [40]. More works focus on employing extra knowledge to guide the generation and hence tend to generate meaning and context-related responses. Liu et al. presented a neural knowledge diffusion model to introduce knowledge into dialogue generation [41]. Young et al. incorporated common sense knowledge about the concepts covered in utterances into end-to-end conversational models [30]. Madotto et al. used a multihop attention mechanism over memories with pointer networks to effectively incorporate knowledge base information in generative dialogue systems [42]. Moon et al. combined a knowledge graph with conversational utterances to infer the correct entity as the output response [43]. Lian et al. focused on the selection of knowledge for conversational response generation [21]. Both Li et al. [44] and Li et al. [26] proposed document-grounded dialogue generation models to form informative and interesting multiturn responses.

### 3. Model Architecture

This work proposes a novel transformer-based model which leverages joint encoding of a given document and dialogue for response generation, as shown in Figure 1. It follows the encoder-decoder framework by only using stacked layers, each of which consists of a multihead attention mechanism and position-wise connection network. Encoder-decoder neural networks with attention functions have been widely leveraged for solving sequential language generation [45].

In a multiturn conversational generation task, a dialogue is commonly considered as a sequence of  $K$  utterances  $\{u_1, u_2, \dots, u_k\}$ , which contains the dialogue history  $\{u_1, u_2, \dots, u_{k-1}\}$  and the current utterance  $u_k$ , where  $u_i = \{w_1^u, w_2^u, \dots, w_{|u_i|}^u\}$  denotes the  $i$ -th utterance in the multiturn dialogue and  $w_j^u$  denotes the  $j$ -th word in the  $i$ -th utterance. In this work, we denote the dialogue as a token-level sequence  $U = \{w_1^u, w_2^u, \dots, w_{l_u}^u\}$ , where  $l_u$  denotes the length of the dialogue sequence. The given document for response generation is denoted as  $D = \{w_1^d, w_2^d, \dots, w_{l_d}^d\}$ , where  $l_d$  is the length of the document sequence. The dialogue utterances and document are fed into the encoders to learn their distributed representation, which is illustrated in

Section 3.2. The output of the decoder in our model is a generated response  $R = \{w_1^r, w_2^r, \dots, w_T^r\}$ , where  $T$  is the length of the response.

*3.1. Attention Mechanism and Multihead Attention.* We take advantage of the attention mechanism to capture the interactions between the document and the dialogue, which allows the model to attend the useful information for response generation. We assume that there are  $n_1$  queries and  $n_2$  key-value pairs. Then, we use an attention function to obtain a weighted sum of the values for each query, where the query, key, and value are all vectors of dimension  $d_k$ . The weight  $\alpha$  assigned to each value is computed by a scaled dot-product function of the query with the corresponding key, shown as the following equation:

$$\text{Attention}(Q, K, V) = \alpha \cdot V, \quad (1)$$

$$\alpha = \text{softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right),$$

where  $Q$  is the matrix packing with a set of  $d_k$ -dimensional vectors of queries,  $Q \in R^{n_1 \times d_k}$ ,  $K$  and  $V$  are also matrices,  $K \in R^{n_2 \times d_k}$  and  $V \in R^{n_2 \times d_k}$ . The weight  $\alpha \in R^{n_1 \times n_2}$ .

Moreover, our model implements multihead attention [25] in all the attention computations to jointly collect information from different representation subspaces at different positions. We define the multihead attention function with  $M$  heads for projecting the queries, keys, and values  $M$  times with different learned linear projections. The result of the multihead attention function MultiHead is a vector that concatenates all the output vectors across  $M$  heads, shown as follows:

$$\text{MultiHead}(Q, K, V) = [\text{head}_1; \dots; \text{head}_m; \dots; \text{head}_M], \quad (2)$$

where  $\text{head}_m$  denotes the  $m$ -th weighted vector calculated as the following formula:

$$\begin{aligned} \text{head}_m &= \text{Attention}(Q_m, K_m, V_m), \\ Q_m &= \tilde{Q} \cdot W_m^Q, \\ K_m &= \tilde{K} \cdot W_m^K, \\ V_m &= \tilde{V} \cdot W_m^V, \end{aligned} \quad (3)$$

where  $\tilde{Q}$ ,  $\tilde{K}$ , and  $\tilde{V}$  indicate the vectors of the input query, key, and value with the same dimension  $d_{\text{model}}$ , respectively.  $W_m^Q, W_m^K, W_m^V \in R^{d_{\text{model}} \times d_{\text{att}}}$  are trainable parameter matrices for the  $m$ -th head, and  $d_{\text{att}} = d_{\text{model}}/M$ .

*3.2. Encoder.* The encoder of the proposed model consists of an utterance encoder and a context encoder. The former aims to learn the representation of the current utterance, and the latter aims to learn the representation of the conversational context (document and dialogue utterances). The encoder is inspired by the reading behavior of human beings. Generally, a basic process of reading comprehension is that



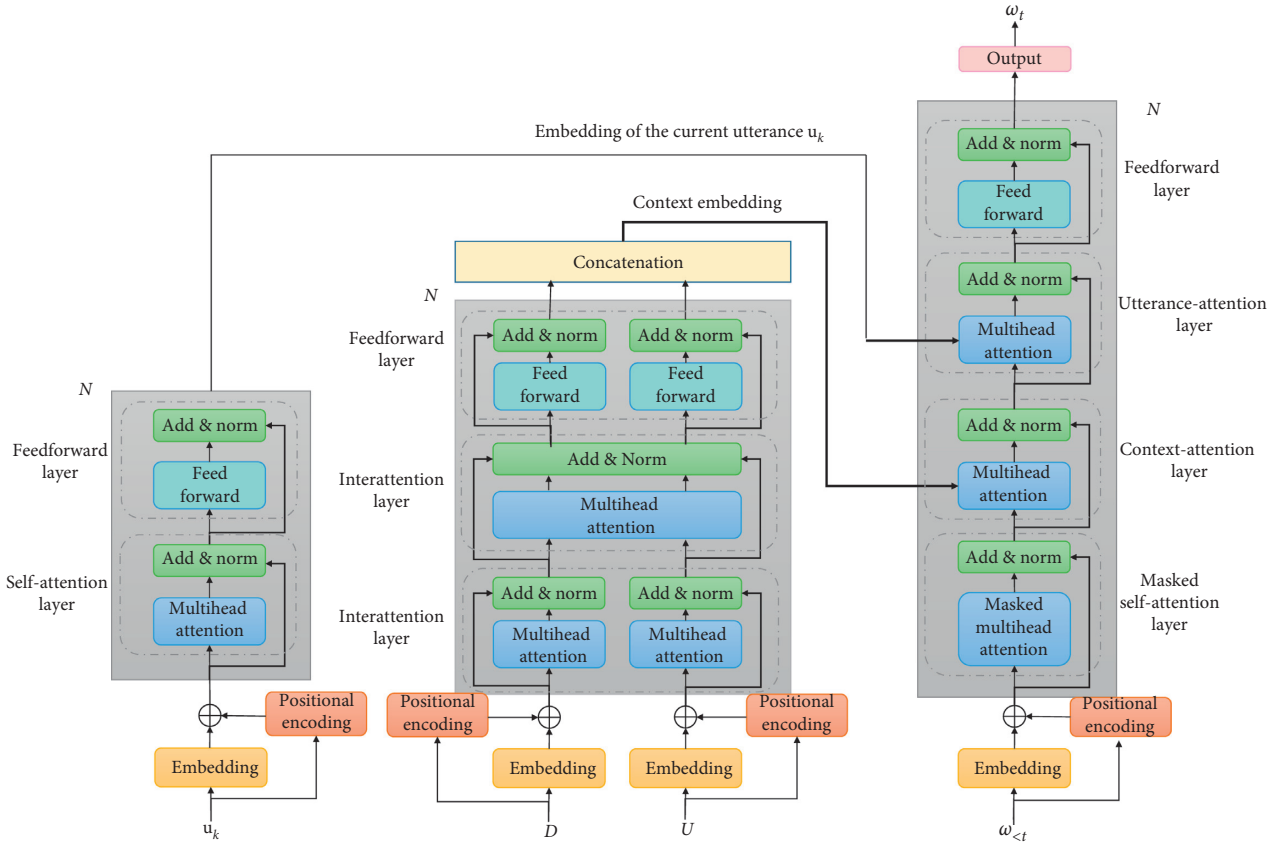


FIGURE 1: The framework of the BCTCE (bichannel transformer with context encoding). The left is the utterance encoder. The center is the context encoder which has binary channels. The right is the decoder for response generation.

firstly reading through the given document and dialogue to understand the theme and capture the key information from the context. Then, we focus on the current utterance for generating the answer. In the context encoder, the given document and dialogue utterances are encoded in parallel by intraattention interaction and interattention interaction. This parallel learning process aims to better represent the context and fuse the information of the document and utterances.

We first map the symbolic representations of input sequences to distribution representations. The tokens of the current utterances, document, and dialogue utterances are fed one by one into the encoder. Moreover, it has been widely accepted that position information is critical to indicate the order of the sequential input. However, the self-attention mechanism itself cannot distinguish between different positions. So, we introduce an additional position embedding to encode position information of the input into the word vectors, shown as the “positional encodings” module in Figure 1. The sum of the original word embedding and the position embedding is defined as the distribution representation  $e(w)$  of word  $w$ :

$$e(w) = \text{embed}(w) + \text{PE}(w), \quad (4)$$

where  $\text{embed}(\cdot)$  denotes an embedding lookup function; the position embedding  $\text{PE}(\cdot)$  is defined as in [25]:

$$\text{PE}(w)_{\text{pos},d} = \begin{cases} \sin(\text{pos}/10000^{d/d_{\text{model}}}), & \text{if } d \text{ is even,} \\ \cos(\text{pos}/10000^{(d-1)/d_{\text{model}}}), & \text{otherwise,} \end{cases} \quad (5)$$

where  $\text{pos}$  is the position of  $w$  in the dialogue sequence or document sequence,  $d$  denotes the  $d$ -th dimension of the representation, and  $d_{\text{model}}$  is the dimension of the input embedding.

The utterance encoder is the same as that of the original transformer [25] with  $N$  stacks, shown in the left of Figure 1. It outputs the embedding of the current utterance and has two sublayers in each stack. The first sublayer is constructed by the multihead self-attention mechanism, and the second is a position-wise fully connected feedforward network.

The context encoder is a variant of the transformer encoder with  $N$  stacks, shown in the center of Figure 1. Differing from the original transformer that only deals with a single channel in the encoder, it builds binary channels to separately encode both dialogue and document in the first step. The context encoder is composed of a stack of  $N$  identical layers, and each stack contains three sublayers:

- (i) Intra-attention layer: this layer is employed to encode two individual input sequences using the multihead

self-attention mechanism separately. The representation of dialogue utterances is

$$\bar{h}_n^U = \text{MultiHead}(h_{n-1}^U, h_{n-1}^U, h_{n-1}^U). \quad (6)$$

The representation of the document is

$$\bar{h}_n^D = \text{MultiHead}(h_{n-1}^D, h_{n-1}^D, h_{n-1}^D), \quad (7)$$

where  $h_{n-1}^U$  and  $h_{n-1}^D$  are the outputs of the last stack.  $h_0^U$  and  $h_0^D$  are the embedding representation of the given dialogues and document.

- (ii) Interattention layer: this layer is proposed to fuse the semantic and syntactic features of the dialogue and document. It helps the model to focus on the relevant contents of the document and dialogue since the dialogue-related segment in the document may supply useful information for answering the current require. This layer is also implemented by a multihead attention function, and the query, key, and value are all  $[\bar{h}_n^U; \bar{h}_n^D]$ . The output of this layer is  $\tilde{h}_n^U$  and  $\tilde{h}_n^D$ .
- (iii) Feedforward layer: it uses a fully connected position-wise network to nonlinearly map tokens across different positions separately and identically. The nonlinear function used in this layer is

$$\begin{aligned} h_n^U &= \text{Relu}\left(\tilde{h}_n^U \cdot W_{n1}^U + b_{n1}^U\right) \cdot W_{n2}^U + b_{n2}^U, \\ h_n^D &= \text{Relu}\left(\tilde{h}_n^D \cdot W_{n1}^D + b_{n1}^D\right) \cdot W_{n2}^D + b_{n2}^D, \end{aligned} \quad (8)$$

where  $W_{n1}^U, W_{n1}^D \in R^{d_{\text{model}} \times d_{\text{inner}}}$ ,  $W_{n2}^U, W_{n2}^D \in R^{d_{\text{inner}} \times d_{\text{model}}}$ ,  $b_{n1}^U, b_{n1}^D \in R^{d_{\text{inner}}}$ , and  $b_{n2}^U, b_{n2}^D \in R^{d_{\text{model}}}$  are trainable parameters.  $d_{\text{inner}}$  is the size of the hidden layer in the feedforward network.

In addition, each sublayer has an ‘‘Add & Norm’’ operation, which is defined in the original transformer framework. The output of the last stack is matrix  $h_N^U$  and matrix  $h_N^D$ , which are concatenated as the output of the context encoder.

**3.3. Decoder.** The decoder also has  $N$  stacks and contains three layers per stack, as shown in the right of Figure 1. At time step  $t$ , the previous  $t - 1$  tokens and the output of the encoder are fed into the decoder to predict the  $t$ -th token in the response illustrated as the output of the  $N$ -th stack  $h_t^R \in R^{d_{\text{model}}}$ .

- (i) Masked self-attention layer: this layer is similar to the intra-attention layer in the encoder. The difference is that we mask the subsequent positions of each token to ensure that the consequent utterance only depends on the previous tokens.

- (ii) Context-attention layer: to enrich the context information in generated responses, the output of the context encoder is fed into the decoder and integrated with the previous response tokens by a multihead attention function, where the query of the function is the output of the masked self-attention layer, and the key and the value are the output of the context encoder.

- (iii) Utterance-attention layer: generally, the generated response must be relevant to the current utterance. Thus, we also use a multihead attention function to introduce the key information of the current utterance for generating dialogue-related responses in this layer. The query of the function is the output of the context-attention layer, and the key and the value are the output of the utterance encoder.

- (iv) Feedforward layer: this layer is the same as the feedforward layer in the encoder.

We select the token derived from an external vocabulary  $V_o$ ; the probability of each candidate token being chosen is

$$P^v(w_t) = \text{softmax}(h_t^R \cdot e(w_t)). \quad (9)$$

Then, we define the probability of generating the  $t$ -th token as  $P(w_t)$ :

$$P(w_t) = P^v(w_t). \quad (10)$$

At each time step, we select the token that has the highest probability as the generated token:

$$\hat{w}_t = \text{argmax}(P(w_t)). \quad (11)$$

During the training process, the loss for time step  $t$  is defined as the negative log-likelihood of the target word  $w_t^*$ :

$$l_t = -\log P(w_t^*). \quad (12)$$

The final loss is

$$l = \frac{1}{T} \sum_t l_t. \quad (13)$$

**3.4. Copying Mechanism.** In this work, we tend to generate more imaginative and context-aware responses. However, some tokens in the ground truth may not be included in the vocabulary (OOV, out of vocabulary). As such, we propose a variant of our transformer model that incorporates the copying mechanism [46, 47] into the decoder to generate tokens that appear in the document and dialogue in addition to the external vocabulary. The tokens in generated responses may be chosen from the input or an external vocabulary according to a computed probability.

At each time step  $t$ , according to the multihead attention weights resulted from the context-attention layer, we determine the probability that the generated tokens are derived from the input document and dialogue utterances as the average of all the attention weights  $\alpha_t^1, \alpha_t^2, \dots, \alpha_t^M$  derived from  $M$  heads:

$$\alpha_t = \frac{1}{M} \sum_{m=1}^M \alpha_t^m, \quad (14)$$

where  $\alpha_t \in R^{l_u+l_d}$  and  $\alpha_t^m$  indicates the attention weight of the  $m$ -th head.

According to the copying mechanism, the probability of tokens being chosen from the vocabulary is  $p_t^g \in [0, 1]$ , while the probability of tokens being chosen from the input sequences is  $1 - p_t^g$ :

$$P(w_t) = \begin{cases} p_t^g * P^v(w_t) + (1 - p_t^g) * \sum_{i:w_i=w_t} \alpha_t(i), & \text{if } w_t \in V_o \text{ and } w_t \in U \cup D, \\ (1 - p_t^g) * \sum_{i:w_i=w_t} \alpha_t(i), & \text{if } w_t \notin V_o \text{ and } w_t \in U \cup D, \\ p_t^g * P^v(w_t), & \text{if } w_t \in V_o \text{ and } w_t \notin U \cup D, \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

## 4. Experiments

**4.1. Experiment Settings.** We conduct the experiments, and the stacks of both encoder and decoder are set to 4. The number of attention heads is set to 8. The dimension of input embedding  $d_{\text{model}}$  is set to 512, and the hidden size of the feedforward network  $d_{\text{inner}}$  is set to 2,048. In the process of encoding, we take the previous four utterances and the given document as the input. We use the Adam algorithm [48] with learning rate 0.0001 for optimization. The batch size is set to 64, and the dropout rate is set to 0.1. In addition, we train the model for 50 epochs.

**4.2. Dataset.** We evaluate the proposed model on the dataset CMU\_DoG (CMU document-grounded conversations) for document-driven conversations [24]. This dataset consists of a set of documents and a spectrum of dialogues corresponding to each document, which may contain movie names, ratings, introduction, and some other scenes. The documents present conversation-related information that may help generate context-aware responses in a multiturn conversation task.

The dataset has a total of 4,112 conversations with an average of 21.43 turns. The dialogue utterances are derived from two different scenarios, both of which involve two participants. In the first scenario, only one participant has access to the given document, while both participants have access to the same given document in the second scenario. The number of conversations for scenario ONE is 2,128, and for scenario TWO, it is 1,984. This high-quality dataset explicitly presents the corresponding relationship between each section of a document and the conversation turns. The average length of documents is approximately 200. There are 72,922 utterances for training, 3,626 utterances for validation, and 11,577 utterances for testing.

$$p_t^g = \text{sigmoid}(W^T \cdot h_t^R), \quad (15)$$

where  $W$  is a trainable parameter.

The probability of the  $t$ -th generated token  $w_t$  is calculated according to the source it is derived from as follows:

### 4.3. Evaluation

**4.3.1. Baselines.** We take the transformer [25], incremental transformer [44], D3G [26], SEQ, and SEQs [24] as the baseline models, which are proposed for document-grounded conversational generation.

**4.3.2. Quantitative Evaluation.** To measure the performance of the proposed model and the baselines, we take BLEU [27], METEOR [28], NW [24], and perplexity (PPL) [19] as evaluation criteria to perform automatic evaluation.

- (1) BLEU: BLEU is known to correlate reasonably well with human evaluation on the task of conversational response generation. It measures  $n$ -gram overlap between generated responses and the ground truth, which is defined as BLEU- $n$ . We calculate various BLEU scores between the golden responses and the generated responses. Moreover, we calculate the unigram overlap between the given document and the generated responses to further compare models in terms of the correlation between the responses and the document. Therefore, we only use the BLEU-1 score (called as Doc\_BLEU) and ignore the brevity penalty factor in the BLEU computation.
- (2) METEOR: we also compare our proposed model with state-of-the-art baselines in terms of the METEOR metric under the full mode (this mode contains the exact matching between words and phrase matching between stems, synonyms, and paraphrases). METEOR, which focuses on the recall rate, has more relevance with human judgment in comparison to BLEU.
- (3) NW: we explore the set operation (NW) to evaluate the relevance between documents and the conversations generated by the models. Let the set of tokens in the generated response be  $N$ , the set of tokens in

the document be  $M$ , the set of tokens in the previous three utterances be  $H$ , and the set of stop words be  $S$ . We calculate the set operation (NW) as  $|(N \cap M)/H|/S$ . A higher NW score indicates that more tokens that appear in the document are used to expand the information in responses.

- (4) Perplexity: in addition to the previous three criteria, we use perplexity to automatically evaluate the fluency of the response. Lower perplexity indicates better performance of the models and higher quality of the generated sentences.

**4.3.3. Human Judgment.** Manual evaluations are essential for dialogue generation. So, we augment the automatic evaluation with the human judgment of fluency, dialogue coherence, and lexical diversity. All the three evaluation metrics are scored 0/1/2. We randomly sample multiple conversations containing 822 utterances from the test set. We used a crowdsourcing service that asks annotators to score these utterances given its previous utterances and related documents. The final score of each utterance is the average of the scores rated by three annotators.

- (1) Fluency: whether the response is natural and fluent. Score 0 represents the response is not fluent and incomprehensible; 1 represents the response is partially fluent but still comprehensible; and 2 represents the response is sufficiently fluent.
- (2) Dialogue coherence: whether the response is logically coherent with the dialogue. Score 0 represents the response is irrelevant with the previous utterances; 1 represents the response matches the topic of the previous utterances; and 2 represents the response is exactly coherent with the previous utterances.
- (3) Lexical diversity: whether the response is vivid and diverse. Score 0 represents the safe response which is applicable to almost all conversations, e.g., “i think so” and “i agree with you”; 1 represents the response suitable to limited conversations but plain and uninformative; and 2 represents the response is evidently vivid, diverse, and informative.

**4.4. Results and Discussion.** Tables 1–3 show the comparison of our model with other models on the CMU\_DoG dataset. As shown in Table 1, our model outperforms all baselines in terms of BLEU- $n$  scores and METEOR score. Our model achieves a new state-of-the-art performance, which indicates that the responses generated by our model are more similar to the ground-truth responses. Our model significantly outperforms all the baselines by at least 1.43, 0.68, 0.39, 0.47, and 0.01 in terms of BLEU-1, BLEU-2, BLEU-3, BLEU-4, and METEOR, respectively.

Table 2 shows the comparison of document relevance and response quality in terms of Doc\_BLEU score, NW, the average length of responses (avg\_len), and the PPL score. Our model outperforms the baselines by 4.5%–11.9% in

terms of BELU-1, while it has lower NW score than D3G and incremental transformer (our impl). These results indicate that our model can more effectively use the shared information between the document and the dialogue to produce responses than D3G and incremental transformer (our impl). Moreover, the average length of the responses generated by our model is higher than that of the baselines, which shows that our model may generate more informative responses. In addition, our model achieves a competitive PPL score with others.

As the results of human judgment shown in Table 3, our transformer-based model outperforms all the baselines in terms of the dialogue coherence and diversity. However, the performance of our model is slightly worse than incremental transformer (our impl) [44] on fluency.

**4.5. Training Time.** Figure 2 presents the training time for one epoch of our BCTCE model and some baselines (As the official source code of “Incremental Transformer,” the process of training for each step is followed by the evaluation of the generated responses. Therefore, it is hard to get the actual training time for one epoch of “Incremental Transformer” model.). For a fair comparison, all models use the same batch size, max length of the document, max length of the dialogue sequence, and max length of the response.

As shown in the figure, the training time for our BCTCE model is much less than D3G, while it is higher than other models. The reasons are as follows:

- (1) The SEQ model is a simple sequence-to-sequence RNN model with attention mechanism. It only uses the dialogue as the input of the encoder and discards the document, thus requiring considerably less time for model training than our model.
- (2) The SEQS model feeds the concatenation of the encoded utterances and document into the decoder, and the original transformer directly takes the concatenation of the utterances and document as the input of the model. Although they have lower training time, our model extends the original transformer by introducing an extra context encoder and a “context-attention layer” in the decoder, where the interaction between utterances and document is time-consuming but effective.
- (3) Our transformer-based model significantly reduces the training time in comparison with the D3G model which also uses the interaction between document and dialogue in its RNN structure.

**4.6. Ablation Study.** To validate the effectiveness of each module of the BCTCE, we conduct ablation experiments on the CMU\_DoG dataset:

- (1) +copy: we introduce the copy mechanism into the BCTCE model to generate the response from the document and dialogue utterances in addition to external vocabulary.

TABLE 1: Quantitative evaluation results for baselines and the proposed models in terms of BLEU scores and METEOR score.

Models	BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR
SEQ [24] <sup>§</sup>	6.12	1.52	0.59	0.30	4.18
SEQS [24] <sup>§</sup>	6.57	1.65	0.67	0.35	4.30
D3G [26] <sup>§</sup>	6.32	1.71	0.71	0.41	4.17
Transformer [25] <sup>¶</sup>	8.55	2.49	1.12	0.60	4.53
Incremental transformer [44] <sup>†</sup>	—	—	—	0.95	—
Incremental transformer (our impl) <sup>‡</sup>	8.19	2.88	1.66	0.85	5.21
BCTCE	9.98	3.56	2.05	1.42	5.22

Results marked with <sup>§</sup> are trained and evaluated with the source code from [26], results marked with <sup>¶</sup> are trained and evaluated with our implemented code, results marked with <sup>†</sup> are from [44], and results marked with <sup>‡</sup> are trained and evaluated with the code published by Li et al. [44].

TABLE 2: Quantitative evaluation results of document relevance and response quality.

Models	Doc_BLEU	NW	avg_len	PPL
SEQ [24] <sup>§</sup>	24.88	0.23	7.31	15.62
SEQS [24] <sup>§</sup>	27.96	0.34	7.21	19.53
D3G [26] <sup>§</sup>	26.76	0.39	6.83	18.40
Transformer [25] <sup>¶</sup>	27.55	0.33	7.91	13.70
Incremental transformer [44] <sup>†</sup>	—	—	—	15.11
Incremental transformer (our impl) <sup>‡</sup>	26.96	0.42	8.52	11.01
BCTCE	28.23	0.36	9.16	17.80

The marks in this table are the same as those in Table 1.

TABLE 3: Human evaluation of baselines and our proposed model.

Models	Fluency	Dialogue coherence	Diversity
SEQ [24] <sup>§</sup>	1.27	0.81	0.42
SEQS [24] <sup>§</sup>	1.13	0.96	0.71
D3G [26] <sup>§</sup>	1.29	1.12	0.84
Transformer [25] <sup>¶</sup>	1.34	1.17	0.90
Incremental transformer (our impl) <sup>‡</sup>	1.35	1.27	0.93
BCTCE	1.34	1.29	0.95

The marks in this table are the same as those in Table 1.

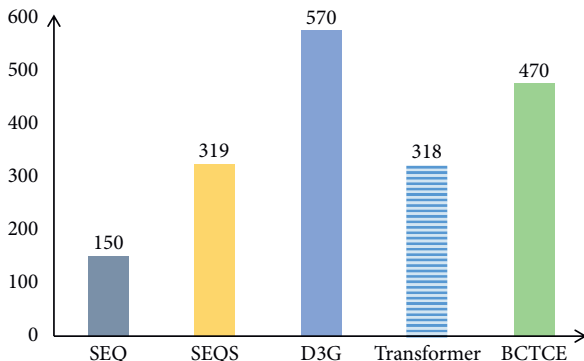


FIGURE 2: Training time of various models.

- (2) -document: we replace the context encoder with the original transformer and take only dialogue utterances as its input
- (3) -history: we remove the dialogue history from the inputs, remaining only the current dialogue utterance and document
- (4) -context encoder: we discard the context encoder and the context-attention layer in the decoder

- (5) -utterance encoder: we discard the utterance encoder and the utterance-attention layer in the decoder
- (6) -bi\_channel: we replace the context encoder with the original transformer and take the concatenation of the dialogue utterances and document as its input

The automatic evaluation results are shown in Tables 4 and 5, respectively. The human evaluation results are shown in Table 6. As shown in Table 4, the ablation models, which remove some modules of BCTCE, perform worse than the basic BCTCE model on the similarity between generated responses and ground truth (BLEU- $n$  scores and METEOR score). The results of “-bi\_channel” indicate that the interaction between the document and the dialogue in bichannel encoding is effective for generating responses. The results of “-context encoder” and “-utterance encoder” show that the context encoder and the utterance encoder are beneficial for response generation. The results of “-document” and “-history” represent that the multiturn dialogue and the document knowledge are important as they contain some vital information useful for generating reasonable response.

Table 5 shows that removing the document or introducing copy mechanism reduces the Doc\_BLEU and NW scores. The results indicate that the BCTCE may pay more

TABLE 4: Ablation study in terms of BLEU scores and METEOR score.

Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR
BCTCE	9.98	3.56	2.05	1.42	5.22
+copy	9.53	3.21	1.70	1.08	4.95
-document	8.55	2.33	0.95	0.46	4.62
-history	8.50	2.84	1.51	1.01	4.54
-context encoder	7.68	1.96	0.74	0.34	4.06
-utterance encoder	9.87	3.34	1.89	1.30	5.13
-bi_channel	9.40	3.05	1.58	1.02	4.80

TABLE 5: Ablation study on document relevance and response quality.

Models	Doc_BLEU	NW	avg_len	PPL
BCTCE	28.23	0.36	9.16	17.80
+copy	27.38	0.34	8.63	16.33
-document	26.25	0.23	8.42	12.97
-history	28.47	<b>0.65</b>	8.05	15.97
-context encoder	23.66	0.39	6.72	<b>11.23</b>
-utterance encoder	28.25	0.41	<b>9.40</b>	16.88
-bi_channel	<b>29.15</b>	0.33	7.94	14.26

TABLE 6: Human evaluation for ablation study. All metrics are scored 0/1/2.

Models	Fluency	Dialogue coherence	Diversity
BCTCE	1.34	<b>1.29</b>	0.95
+copy	1.33	1.26	<b>1.01</b>
-document	1.30	1.27	0.80
-history	1.31	1.06	0.88
-context encoder	<b>1.36</b>	0.81	0.53
-utterance encoder	1.34	1.27	0.87
-bi_channel	1.29	1.26	0.85

attention to the dialogue utterances after removing the document information, and the copy mechanism has less influence on the generated response than expected since the BCTCE has sufficient capability to learn the document knowledge for response generation. The Doc\_BLEU and NW scores increase when removing the dialogue history or utterance encoder from the basic BCTCE model as the lack of sufficient dialogue information makes the model to be more focused on the document. The ablation model “-bi\_channel” increases the Doc\_BLEU score and reduces the NW score, which indicate its generated responses pay a little more attention to the shared information between the document and the dialogue. It is worth noting that the ablation model “-context encoder” significantly outperforms the BCTCE model on PPL and the fluency shown in Table 6. A possible reason is that it tends to generate safe and unremarkable responses (e.g., I don’t know). The decrease of its avg\_len and the diversity shown in Table 6 also supports our argument. Moreover, the copy mechanism effectively reduces the PPL from 17.80 to 16.33 and increases the diversity from 0.95 to 1.01, which indicate that it can improve the response quality of the basic BCTCE model.

*4.7. Case Study.* In this section, we demonstrate three conversation cases and show the responses generated by our BCTCE model and several ablation models (as shown in Table 7). The first case shows that our BCTCE model produces a response that contains an “OOV,” while the BCTCE with a copy mechanism extracts the token “post” and “times” from the document. It indicates that the copy mechanism eases the “OOV” problem, although it reduces the performance of the basic BCTCE model on many evaluation metrics.

In the second case, the BCTCE without the document or context encoder produces a completely incorrect answer, while the basic BCTCE and other ablation models can produce at least one correct answer “jime carrey” contained in the ground truth. These results show that the document is necessary for generating answers, and the BCTCE can take full advantage of the valid information in the document.

The third case indicates the importance of the dialogue history. For the conversation whose current utterance is an uninformative interjection “wow,” the BCTCE model produces a response “all around, it has a lot of great stories,” referring to the “story” in the previous utterances, while the

TABLE 7: Sample responses of multiple variants.

Document	Dialogue	Responses
<p>...the white house retaliates, and in short order the post and times appear together before the supreme court to plead their first amendment argument for the right to publish the material. meanwhile, newspapers across the country pick up the story in solidarity with the post and times. the court rules 6 3 in the newspapers' favor, vindicating graham's decision. nixon demands that the post should be barred from the white house. one year later, a security guard discovers a break-in in progress at the watergate complex ...</p>	<p>... the supreme court ruled 6-3 in favor of the news papers, which vindicated graham's decision. Now i really want to see it. it sounds really juicy, but it will be a little difficult because i feel like the situation will anger me.</p>	<p>Ground truth: but they won. nixon god mad and barred the post from the white house BTMPE: OOV -document: i think the post won the oscar for the OOV -history: i like the OOV of the white house, but they OOV the post to OOV the house. -context encoder: i think that is a good way to end the movie. -utterance encoder: it sounds like a good movie to Watch. -bi_channel: and then the supreme court 6-3 in the supreme court, the court rules. +copy: i wonder if the post joined the times or not.</p>
<p>...bruce nolan (jim carrey) is a television field reporter for eyewitness news at wkbw-tv in buffalo, new york, but desires to be the news anchorman . he is in a relationship with his girlfriend grace connelly (jennifer aniston), but also has a mild crush on his co-worker, susan ortega (catherine bell) , who barely seems to notice him. bruce, however, suffers from constant bad luck and reaches breaking point when he is passed over for promotion by his rival, evan baxter (steve carell) ...</p>	<p>... it did not get a good rating on rotten tomatoes. i liked it though that happens a lot with rotten tomato who is in the movie?</p>	<p>Ground truth: jim carrey and jennifer aniston BTMPE: it is a comedy film starring jim carrey and jennifer aniston as the main character -document: it's got 92% on rotten tomatoes -history: jim carrey plays bruce nolan. -context encoder: the movie is about a man named michael oher who is a stockbroker. -utterance encoder: jim carrey plays the lead -bi_channel: jim carrey is a funny actor. +copy: jim carrey</p>
<p>...the duo tries to make it to the moving truck, but sid's dog, scud, sees them, and gives chase. buzz gets left behind while saving woody from scud, and woody tries rescuing him with andy's rc car, but the other toys, thinking woody eliminated rc as well, attack and toss him off the truck. having evaded scud, buzz and rc retrieve woody, and continue after the truck. upon seeing woody and buzz together on rc, the other toys realize their mistake, and ...</p>	<p>... so tell me the good things you heard about it heard that it had a cute story line/plot and that there are three movies total, toy story 1, 2, and 3 wow</p>	<p>Ground truth: that's about it BTMPE: all around, it has a lot of great stories -document: i'm -history: yeah he did a great job. -context encoder: OOV -utterance encoder: do you know what kind of movie it is? -bi_channel: have a great day +copy: i like the first pixar movies</p>

BCTCE without the dialogue history outputs a dialogue-independent response “yeah he did a great job” that contains an uncertain person pronoun. And the BCTCE without the context encoder produces an uninformative “OOV.” This case shows that the BCTCE without the dialogue history or context encoder generates dialogue-inconsistent responses, while the responses generated by others which use historical

dialogue information are more reasonable and consistent with the current utterance.

## 5. Conclusions

In this paper, we propose the BCTCE model based on the transformer framework for document-driven conversational

response generation. This model consists of a context encoder, an utterance encoder, and a decoder, aiming to generate context-aware responses. The context encoder is used to learn the distributed representations of the input document and dialogue utterances in parallel and then fuse them by the multihead self-attention function. The utterance encoder aims to represent the current utterance as distributed embedding. The decoder focuses on the document knowledge and dialogue coherence for predicting the next response. The competitive models are evaluated by comparing the generated responses with the ground truth. Empirical results show that the BCTCE outperforms state-of-the-art baselines in terms of various BLEU scores, METEOR, and NW. The effectiveness of the modules in the BCTCE is indicated by the ablation study. And the manual evaluation and case study show that our model can capture the useful information contained in the document and dialogue, which helps to generate diverse and reasonable responses with much more relevance with the context. In the future work, we will try to build various encoders and concatenate the output from the encoders to integrate the input sequences and generate reasonable context representation.

### Data Availability

The textual 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.

### Authors' Contributions

All authors have contributed to the creation of this manuscript for important intellectual content and approved the final manuscript.

### Acknowledgments

This work was supported in part by the R&D Program of Beijing Municipal Commission of Education (no. KM202010011011), the Beijing Municipal Natural Science Foundation (nos. 4202014 and 4184084), National Key R&D Program of China (nos. 2019YFC1606401, 2018YFC0831605, and 2016YFD0401205), Support Project of High-Level Teachers in Beijing Municipal Universities in the Period of 13th Five-Year Plan (no. CIT&TCD201804031), the Social Science Foundation of Beijing (no. 19GLB036), and the Humanity and Social Science Youth Foundation of Ministry of Education of China (no. 17YJCZH007).

### References

- [1] L. Shang, Z. Lu, and H. Li, "Neural responding machine for short-text conversation," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 1577–1586, Beijing, China, July 2015.
- [2] I. V. Serban, A. Sordoni, Y. Bengio, A. C. Courville, and J. Pineau, "Building end-to-end dialogue systems using generative hierarchical neural network models," in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pp. 3776–3784, Phoenix, AZ, USA, February 2016.
- [3] A. Sordoni, M. Galley, M. Auli et al., "A neural network approach to context-sensitive generation of conversational responses," *Transactions of the Royal Society of Tropical Medicine & Hygiene*, vol. 51, no. 6, pp. 502–504, 2015.
- [4] O. Vinyals and Q. Le, "A neural conversational model," in *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015*, Lille, France, July 2015.
- [5] H. Zhang, Y. Lan, L. Pang, J. Guo, and X. Cheng, "RECOSA: detecting the relevant contexts with self-attention for multi-turn dialogue generation," in *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Volume 1: Long Papers*, pp. 3721–3730, Florence, Italy, July 2019.
- [6] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological Influence-Aware Recommendation on Social Networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [7] Y. Wu, W. Wu, C. Xing, C. Xu, Z. Li, and M. Zhou, "A sequential matching framework for multi-turn response selection in retrieval-based chatbots," *Computational Linguistics*, vol. 45, no. 1, pp. 163–197, 2019.
- [8] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proceedings of the Annual Conference on Neural Information Processing Systems 2014*, pp. 3104–3112, Montreal, Canada, December 2014.
- [9] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, San Diego, CA, USA, May 2015.
- [10] T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015*, pp. 1412–1421, Lisbon, Portugal, September 2015.
- [11] Z. Lin, M. Feng, C. N. dos Santos et al., "A structured self-attentive sentence embedding," in *5th International Conference on Learning Representations, ICLR 2017*, Toulon, France, April 2017.
- [12] C. Peng, Z. Sun, L. Bing, and Y. Wei, "Recurrent attention network on memory for aspect sentiment analysis," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, September 2017.
- [13] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018.
- [14] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, 2019.
- [15] J. Li, M. Galley, C. Brockett, J. Gao, and B. Dolan, "A diversitypromoting objective function for neural conversation models," *Computer Science*, 2015.
- [16] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, "Recurrent neural network based language model," in *Proceedings of the INTERSPEECH 2010, 11th*



- Annual Conference of the International Speech Communication Association*, pp. 1045–1048, Makuhari, Japan, September 2010.
- [17] H. Wen, M. Gasic, N. Mrksic, P. H. Su, D. Vandyke, and S. Young, “Semantically conditioned LSTM-based natural language generation for spoken dialogue systems,” in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pp. 1711–1721, Lisbon, Portugal, September 2015.
- [18] H. Zhou, T. Young, M. Huang, H. Zhao, J. Xu, and X. Zhu, “Commonsense knowledge aware conversation generation with graph attention,” in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018*, pp. 4623–4629, Stockholm, Sweden, July 2018.
- [19] E. Dinan, S. Roller, K. Shuster, A. Fan, M. Auli, and J. Weston, “Wizard of wikipedia: knowledge-powered conversational agents,” in *Proceedings of the 7th International Conference on Learning Representations, ICLR 2019*, New Orleans, LA, USA, May 2019.
- [20] W. Cui, Y. Xiao, H. Wang, Y. Song, S. Hwang, and W. Wang, “KBQA: learning question answering over QA corpora and knowledge bases,” in *Proceedings of the 43rd International Conference on Very Large Data Bases, VLDB 2017*, pp. 565–576, Munich, Germany, August 2017.
- [21] R. Lian, M. Xie, F. Wang, J. Peng, and H. Wu, “Learning to select knowledge for response generation in dialog systems,” in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019*, pp. 5081–5087, Macao, China, August 2019.
- [22] M. Ghazvininejad, C. Brockett, M. Chang et al., “A knowledge-grounded neural conversation model,” in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, pp. 5110–5117, New Orleans, LA, USA, February 2018.
- [23] H. Le, D. Sahoo, N. Chen, and S. Hoi, “Multimodal transformer networks for end-to-end video-grounded dialogue systems,” in *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019 Volume 1: Long Papers*, pp. 5612–5623, Florence, Italy, July 2019.
- [24] K. Zhou, S. Prabhunoye, and A. W. Black, “A dataset for document grounded conversations,” in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 708–713, Brussels, Belgium, October 2018.
- [25] A. Vaswani, N. Shazeer, N. Parmar et al., “Attention is all you need,” in *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, pp. 5998–6008, Long Beach, CA, USA, December 2017.
- [26] K. Li, Z. Bai, X. Wang, and C. Yuan, “A document driven dialogue generation model,” in *Chinese Computational Linguistics—18th China National Conference, CCL 2019*, pp. 508–520, Kunming, China, October 2019.
- [27] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “BLEU: a method for automatic evaluation of machine translation,” in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics—ACL ’02*, pp. 311–318, Stroudsburg, PA, USA, July 2002.
- [28] M. Denkowski and A. Lavie, “Meteor universal: language specific translation evaluation for any target language,” in *Proceedings of the EACL 2014 Workshop on Statistical Machine Translation*, Baltimore, MD, USA, 2014.
- [29] X. Zhou, L. Li, D. Dong et al., “Multi-turn response selection for chatbots with deep attention matching network,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018 Volume 1: Long Papers*, pp. 1118–1127, Melbourne, Australia, July 2018.
- [30] T. Young, E. Cambria, I. Chaturvedi, H. Zhou, S. Biswas, and M. Huang, “Augmenting end-to-end dialog systems with commonsense knowledge,” in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*, New Orleans, LA, USA, February 2018.
- [31] L. Nio, S. Sakti, G. Neubig, K. Yoshino, and S. Nakamura, “Neural network approaches to dialog response retrieval and generation,” *IEICE Transactions on Information and Systems*, vol. E99.D, no. 10, pp. 2508–2517, 2016.
- [32] K. Cho, B. van Merriënboer, C. Gulcehre et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*, pp. 1724–1734, Doha, Qatar, October 2014.
- [33] J. Chung, Ç. Gülçehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” 2014, <https://arxiv.org/abs/1412.3555>.
- [34] H. Chen, Z. Ren, J. Tang, Y. E. Zhao, and D. Yin, “Hierarchical variational memory network for dialogue generation,” in *Proceedings of the 2018 World Wide Web Conference on World Wide Web—WWW’18*, pp. 1653–1662, Lyon, France, April 2018.
- [35] Y. Kim, C. Denton, L. Hoang, and A. M. Rush, “Structured attention networks,” in *Proceedings of the International Conference on Learning Representations*, Toulon, France, April 2017.
- [36] A. P. Parikh, O. Täckström, D. Das, and J. Uszkoreit, “A decomposable attention model for natural language inference,” in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016*, pp. 2249–2255, Austin, TX, USA, November 2016.
- [37] X. Wu, A. Martinez, and M. Klyen, “Dialog generation using multi-turn reasoning neural networks,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, Volume 1 (Long Papers)*, pp. 2049–2059, New Orleans, LA, USA, June 2018.
- [38] W. Wu, C. Xing, Y. Wu, M. Zhou, Y. Huang, and W. Y. Ma, “Hierarchical recurrent attention network for response generation,” in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*, New Orleans, LA, USA, February 2018.
- [39] T. H. Wen, D. Vandyke, N. Mrksić et al., “A network-based end-to-end trainable task-oriented dialogue system,” in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017 Volume 1: Long Papers*, pp. 438–449, Valencia, Spain, April 2017.
- [40] C. Xing, W. Wu, Y. Wu et al., “Topic aware neural response generation,” in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 3351–3357, San Francisco, CA, USA, February 2017.
- [41] S. Liu, H. Chen, Z. Ren, Y. Feng, Q. Liu, and D. Yin, “Knowledge diffusion for neural dialogue generation,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018 (Volume 1: Long Papers)*, pp. 1489–1498, Melbourne, Australia, July 2018.
- [42] A. Madotto, C.-S. Wu, and P. Fung, “Mem2Seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL*

- 2018 (*Volume 1: Long Papers*), pp. 1468–1478, Melbourne, Australia, July 2018.
- [43] S. Moon, P. Shah, A. Kumar, and R. Subba, “OpenDialKG: explainable conversational reasoning with attention-based walks over knowledge graphs,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 845–854, Florence, Italy, July 2019.
- [44] Z. Li, C. Niu, F. Meng, Y. Feng, Q. Li, and J. Zhou, “Incremental transformer with deliberation decoder for document grounded conversations,” in *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019 Volume 1: Long Papers*, pp. 12–21, Florence, Italy, July 2019.
- [45] V.-K. Tran and L.-M. Nguyen, “Natural language generation for spoken dialogue system using RNN encoder-decoder network,” in *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, Vancouver, Canada, August 2017.
- [46] A. See, P. J. Liu, and C. D. Manning, “Get to the point: summarization with pointer-generator networks,” in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017 (Volume 1: Long Papers)*, pp. 1073–1083, Vancouver, Canada, July 2017.
- [47] J. Gu, Z. Lu, H. Li, and V. O. K. Li, “Incorporating copying mechanism in sequence-to-sequence learning,” in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1631–1640, Berlin, Germany, August 2016.
- [48] D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, San Diego, CA, USA, May 2015.

## Research Article

# Adolescent Sports Behavior and Social Networks: The Role of Social Efficacy and Self-Presentation in Sports Behavior

Lei Lei,<sup>1,2</sup> Huifang Zhang ,<sup>2</sup> and Xin Wang<sup>3</sup>

<sup>1</sup>Sports Department, Northwest A&F University, Yangling 712100, China

<sup>2</sup>School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710064, China

<sup>3</sup>Development Planning Office, Beijing Institute of Computer Technology and Application, Beijing 100854, China

Correspondence should be addressed to Huifang Zhang; zhanghuifang@xjtu.edu.cn

Received 18 June 2020; Revised 16 July 2020; Accepted 21 July 2020; Published 29 August 2020

Guest Editor: Liang Wang

Copyright © 2020 Lei Lei 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.

Social networks are a complex system that members communicate, create new connections or destroy existing connections, and further deliver major impacts on each member's life. Given the spread of the Internet and increased academic pressure, sedentary and prescreen behaviors are very common among adolescents; meanwhile, sports behaviors are gradually decreasing. This situation has had an adverse effect on health. This paper used a questionnaire survey to investigate the influence of social networks on adolescent sports behavior, including the intermediary role of social efficacy and moderating effect of self-presentation. The questionnaire survey was conducted on 568 students from 6 high schools in Shaanxi, Henan, and Shandong Provinces. After this, factor analysis and weighted least squares method were used for the empirical test. Based on theoretical and empirical analysis, this paper found the following: (1) Social networks of adolescents have obvious positive predictive effects on their sports behavior. A single online social network and an offline social network, instrumental network, emotional network, and mixed network have obvious positive predictive effects on adolescent sports behaviors. However, under the influence of multiple types of social networks, an offline social network has a negative predictive effect, while a mixed network has effects that are not as obvious. (2) Social efficacy plays an intermediary role in the relationship between social network and adolescent sports behavior. (3) The moderating effect of self-presentation is not significant.

## 1. Introduction

Today, nearly four-fifths of the world's adolescents lack enough physical activity, and the situation for Chinese adolescents is even worse. Sedentary behavior and prescreen behavior are very common in adolescents' study and daily life activities. In 2019, the overweight and obesity rates of Chinese young students were 28.5% and 11.7%, respectively. The myopia or short sight rate for primary school students, middle school students, and college students is also increasing, currently ranging from 45.7% to 87.7%. Adolescents are the future of the world, and they are a group that needs special attention. Active participation in physical activities is conducive to the development of both the physical and mental health of adolescents, and it is thus an urgent task to promote greater adolescent physical activity levels.

In view of the basic trend of social development, which is changing from relatively closed small group to increasingly open networked community life, human society has reached an unprecedented level in terms of network structure, network connection density, and network member communication. Social networking not only is a static relationship structure but also offers a variety of dynamics for the now microprocess of cognition and interpersonal interaction. To pay close attention to the complex, unique, and continuous interaction between social network and actors, it is necessary to discuss the sociability of actors and their embedded network relationships. Although the influence and intervention of social networking on health risk behaviors, such as smoking, alcohol abuse, and sexual behavior of adolescents, have been extensively studied, the changes and ongoing development of adolescents' physical activity

behavior, a complex social problem, need to be further studied at the level of social networking for the most relevant factors, theories, and mechanisms.

Compared to early childhood and adulthood, adolescents are more influenced by their peers during their process of socialization [1, 2]. Studies have found that social networking can not only promote individuals' participation in collective actions and "catalyze" large-scale collective actions but also establish cross-organizational connections through its networks [3]. Making friendship with others can increase adolescents' motivation for sports. For example, distance cycling in front of friends or in a group will be greater in length than when they are cycling alone, indicating that one's peers do enhance the enthusiasm for sports activities [4]. Indeed, if adolescents have more friends and a larger proportion of same-sex friends, they will participate in more physical activities after school. The intensity of physical activity during recess also positively correlates with the number of friends who participate in sports. This situation also varies according to gender; for example, boys are more easily affected by their network of friends than girls are [5]. In addition, researchers believe that adolescents' physical activity behavior is influenced by their siblings and friends. Friends have a greater influence on adolescent participation in organized physical activities, while siblings mainly affect informal and spontaneous physical activity participation. Besides, the supports of teachers also influence adolescents for teachers and adolescents spending most time together in studying. Sometimes, the teacher-students relationship is important in adolescents' behavior because of regulations and commands. At the same time, parental support is crucial because it can reduce the internal perception barriers of adolescents and affect their daily participation in physical activities. If parents were and are also active participants in physical activity, their adolescent children will be more likely to show active physical activity behavior as well [6].

Although the social network environment should be considered for any intervention of adolescents' physical activities, it is worth noting that the influence of friends and partners on adolescents' physical activities is not regulated by their knowledge of those physical activities, but instead by the interdependent relationship between the network selected by the peers and their physical activity [7]. Therefore, in any intervention of physical activity behaviors of adolescents, one's choice of friends may be critical for promoting and maintaining health and positive behaviors. In other words, social networking may have a positive influence on adolescent sports behavior.

Given the popularization of computers, network platforms have gradually become the main channel for adolescents to use to reach out and make friends. Adolescents not only form offline social network but also form online social networks. Adolescents who are interested in physical activities online prefer to be closer to relevant topics and are more likely to provide information about their physical activities [8]. That is to say, given a digital background, not only may the offline social network actively promote adolescent sports behavior, but it also may promote individual sports behavior as well.

In reality, the theoretical model of "two paths and three layers" of individual behavior suggests that stable personality traits and external situations affect individual behaviors via internal cognitive processes. Moreover, social cognitive theory points out that individual behavior, subject cognition, and environment interact dynamically, while individual cognition is the core of such interaction [9]. Both theories demonstrate that internal cognition plays an important intermediary role between external situations and personal behavior. Bandura pointed out that a sense of efficacy is an important cognitive factor that not only affects an individual's choice of behavior but also determines the effort level and the ability to overcome obstacles [10].

Since self-efficacy is domain-specific, Fan and Mak put forward social efficacy [11] to show that there is an obvious positive correlation between social relations and social efficacy [12]. Individuals with low social efficacy tend to interpret uncertain social situations as dangerous and have negative reactions, including excessive self-concern, anxiety, and operational behavior, all obstacles used to avoid and withdraw from real social situations [13] and further reduce individual behaviors. To sum up then, in a real-life communication situation, social networks, social efficacy, and individual behavior are closely related. Indeed, social efficacy may play an intermediary role in the social network role in influencing adolescent sports behavior.

The popularity of computers and the diversification of network platforms now enable adolescents to present themselves at any time and share and learn on network platforms, such as Facebook or WeChat. With the rise of sports apps and quantitative equipment, many adolescents now use sports software to exercise and show and share themselves on social platforms. For example, the development of Keep and other fitness apps allows adolescents to not only learn the existing sports videos but also share their sports videos and sports records on the same platform and promote their continuous sports activity. On the one hand, this development may raise adolescents' attention toward their own health and strengthen their sports behavior; on the other hand, it can also make other adolescents undertake sports through the communication and sharing of them on social networks. Thus, the adolescents' self-presentation on network platforms is likely to strengthen the influence of social networking on adolescent sports behavior. These above thoughts mean that diverse networks may promote adolescents' self-presentation and that self-presentation may encourage adolescents to do sports more frequently, which shows that self-presentation indeed has a moderating effect.

Through undertaking the above analysis, this paper examines the influence of social networking on adolescent sports behavior, the mediating effect of adolescents' social efficacy, and the moderating effect of self-presentation. Based on social cognition theory, this paper proposes a model that includes a mediating effect and a moderating effect and puts forward three hypotheses: **H1**: social networks positively promote adolescent sports behavior; **H2**: social efficacy plays a mediating role in the relationship between social network and adolescent sports behavior; and **H3**: self-presentation plays a moderating role for social networking and adolescent sports behavior.

At the same time, to deeply analyze the influence of different social networks on adolescents sports, this study divides social networks into online and offline social platforms and emotional, instrumental, and mixed social networks. Of these, teachers and counselors belong to the instrumental network, family members and lovers belong to the emotional network, and roommates and classmates belong to a mixed network [11]. Through an analysis of social networks, social efficacy, and self-presentation, this paper offers both a theoretical basis and practical guidance for promoting the establishment of social networks and facilitating adolescent sports behavior.

## 2. Materials and Methodology

**2.1. Subjects.** The subjects were 630 adolescents from three junior middle schools and three senior high schools in Shaanxi, Henan, and Shandong Provinces. A total of 630 questionnaires were distributed, and of these, 568 valid questionnaires were collected, providing an effective recovery rate of 90.1%. There were 296 boys (52.11%) and 272 girls (47.89%), 376 senior students, and 192 junior students. The average age of the subjects was 14.52 years ( $SD = 1.68$ , with the age range being 12–18).

**2.2. Measurements.** An adolescent sports behavior scale is based on the questionnaire by Mao et al. [14], which included 6 items (e.g., I usually take exercise with friends; I have the habit of exercising), and the questionnaire was scored using 5 answers (1 “totally inconsistent” to 5 “completely consistent”). In this study, the confirmatory factor analysis of the questionnaire fits well:  $\chi^2/df = 2.84$ ,  $RMSEA = 0.05$ ,  $NFI = 0.995$ ,  $GFI = 0.998$ , and  $CFI = 0.997$ . The internal concordance coefficient  $\alpha$  was 0.687.

The social network scale uses a questionnaire designed by Park et al. [15], which contains 14 items and is divided into either two dimensions as (1) an online social network (6 items, e.g., Many of my friends in real life love sports) and (2) an offline social network (4 items, e.g., The number of communities I’ve join online), or three dimensions as (1) an instrumental social network (4 items, e.g., I have a close relationship with my class teacher), (2) an emotional network (4 items, e.g., I have a close relationship with my relatives of the same age), and (3) a mixed network (4 items, e.g., I have a close relationship with my friends in real life).

Some of the items in the questionnaire on different dimensions were the same. For example, “I have a close relationship with my friends in real life” is a measurement item of both an offline social network and a mixed network. The questionnaire is scored using 5 points (1 “totally inconsistent” to 5 “completely consistent”). In the current study, the confirmatory factor analysis of the questionnaire fits well:  $\chi^2/df = 2.84$ ,  $RMSEA = 0.06$ ,  $NFI = 0.998$ ,  $GFI = 0.999$ , and  $CFI = 0.999$ . The internal concordance coefficient  $\alpha$  was 0.654. The coefficients  $\alpha$  of each dimension were 0.940 (online social network), 0.936 (offline social network), 0.898 (instrumental social network), 0.766 (emotional social network), and 0.847 (mixed social network).

The social efficacy scale used a questionnaire compiled by Jeong and Kim [16]. There were 6 items on this questionnaire in the study (e.g., online communication, I can easily become friends with other people; I can easily talk to unfamiliar people). The questionnaire was scored using 5 points (1 “totally inconsistent” to 5 “completely consistent”). In this study, the confirmatory factor analysis of the questionnaire fits well:  $\chi^2/df = 1.24$ ,  $RMSEA = 0.02$ ,  $NFI = 0.998$ ,  $GFI = 0.999$ , and  $CFI = 0.999$ . The internal concordance coefficient  $\alpha$  was 0.634.

The self-presentation scale used a questionnaire compiled by Kim and Lee [17]. It included 4 items (e.g., I will post photos that show the real me; I don’t mind sharing some bad things that happened to me on the Internet). The questionnaire was scored based on 5 points (1 “totally inconsistent” to 5 “completely consistent”). In this study, the confirmatory factor analysis of the questionnaire fits well:  $\chi^2/df = 2.24$ ,  $RMSEA = 0.02$ ,  $NFI = 0.998$ ,  $GFI = 0.999$ , and  $CFI = 0.999$ . The internal concordance coefficient  $\alpha$  was 0.820.

**2.3. Programs and Data Processing.** In this current study, after obtaining the consent of the school leaders, teachers, and students, the questionnaire practice was explained by highly trained surveyors in accordance with standardized instruction, and all questionnaires were collected immediately. SPSS 22.0 software was used to analyze the data, and PROCESS V3.0 was used to test the mediating effect.

First, we used the factor analysis method to reduce the dimensionality of variables. The mathematical model was

$$\begin{cases} x_1 = a_{11}f_1 + a_{12}f_2 + a_{13}f_3 + \cdots + a_{1k}f_k + \varepsilon_1, \\ x_2 = a_{21}f_1 + a_{22}f_2 + a_{23}f_3 + \cdots + a_{2k}f_k + \varepsilon_2, \\ x_3 = a_{31}f_1 + a_{32}f_2 + a_{33}f_3 + \cdots + a_{3k}f_k + \varepsilon_3, \\ \dots \\ x_p = a_{p1}f_1 + a_{p2}f_2 + a_{p3}f_3 + \cdots + a_{pk}f_k + \varepsilon_p. \end{cases} \quad (1)$$

For equation (1),  $x_i$  is the standardized variable;  $f_j$  is the factor variable;  $\varepsilon$  is the special factor;  $a_{ij}$  is the factor load;  $k < p$ . A matrix was used to simplify equation (1) as follows:

$$\mathbf{X} = \mathbf{A}\mathbf{F} + \boldsymbol{\varepsilon}, \quad (2)$$

where  $\mathbf{F}$  is the factor variable matrix and  $\mathbf{A}$  is a factor load matrix and satisfies  $\text{cov}(\mathbf{F}, \boldsymbol{\varepsilon}) = 0$ ,  $D(\mathbf{F}) = \mathbf{I}_m$ ,  $D(\boldsymbol{\varepsilon}) = \boldsymbol{\delta}_i$ .

The factor load  $a_{ij}$  represents the degree of correlation between  $X_i$  and  $F_j$  and the square sum of the elements in the  $i$ -th row of the factor load matrix where  $\mathbf{A}$  is  $h_i^2 = \sum_{j=1}^k a_{ij}^2$ , which represents the explanatory power of all the factor variables for the total variance of  $X_i$ , and  $S_j = \sum_{i=1}^p a_{ij}^2$  is the variance contribution of the variable  $F_j$ .

After determining the factor variables, the factor load matrix needed to be estimated. If we set the eigenvalues of the sample covariance matrix and the corresponding standard orthogonalized eigenvectors, as  $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ ,  $e_1, e_2, \dots, e_p$ , then the covariance matrix can be decomposed into

$$\Sigma = \mathbf{U} \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \dots \\ 0 & & & \lambda_p \end{pmatrix} \mathbf{U}' = \sum_{i=1}^p \lambda_i e_i e_i' \quad (3)$$

When the last several eigenvalues are small, the covariance matrix can be approximately decomposed into

$$\begin{aligned} \Sigma &\approx \left( \sqrt{\lambda_1} e_1, \dots, \sqrt{\lambda_m} e_m \right) \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \dots \\ \sqrt{\lambda_m} e_m' \end{bmatrix} + \begin{pmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \dots \\ & & & \sigma_p^2 \end{pmatrix} \\ &= \mathbf{A} \mathbf{A}' + \sum_{\varepsilon} \end{aligned} \quad (4)$$

where  $\mathbf{A}$  is the factor covariance matrix. Since the factor loading matrix is not unique, factor rotation was executed to make the meaning of the common factor clearer. Considering the orthogonal rotation of two factors, the orthogonal matrix then became

$$\mathbf{Q} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}, \quad (5)$$

and set

$$\mathbf{B} = \mathbf{A} \mathbf{Q} = (b_{ij}), \quad i = 1, 2, \dots, p, \quad j = 1, 2, \quad (6)$$

where  $\mathbf{B}$  is the rotation factor load matrix. At this time, we required the variance of the two columns' data in  $\mathbf{B}$  to be as large as possible, which also meant that the relative variance  $V_j$  should also be as large as possible:

$$V_j = \frac{1}{p} \sum_{i=1}^p \left( \frac{b_{ij}^2}{h_i^2} \right)^2 - \sum_{i=1}^p \left( \frac{b_{ij}^2}{h_i^2} \right), \quad j = 1, 2. \quad (7)$$

Making  $dV/d\phi = 0$ , then  $\phi$  should satisfy the following equation:

$$\tan 4\phi = \frac{D_0 - ((2A_0 B_0)/p)}{C_0 - ((A_0^2 - B_0^2)/p)}, \quad (8)$$

where

$$\begin{cases} A_0 = \sum_{i=1}^p u_i, & B_0 = \sum_{i=1}^p v_i, \\ C_0 = \sum_{i=1}^p (u_i^2 - v_i^2), & D_0 = 2 \sum_{i=1}^p u_i v_i, \\ u_i = \left( \frac{a_{i1}^2}{h_i} \right)^2 - \left( \frac{a_{i2}^2}{h_i} \right)^2, & v_i = \frac{2a_{i1} a_{i2}}{h_i^2}. \end{cases} \quad (9)$$

Using SPSS for the factor analysis, taking adolescent sports and social network as examples, the scree plots were obtained. It can be seen from Figure 1 that the adolescent

sports scale can extract one factor, while the social network scale can extract two factors.

### 3. Results and Discussion

**3.1. Descriptive Analysis.** The mean, standard deviation, and the correlation coefficient of each variable are shown in Table 1. These results show that there are obvious correlations between adolescent sports behavior and different social networks and also obvious correlations between social efficacy, self-presentation, social networking, and sports. In addition, social networks, especially offline social networks, have significant age differences. Different types of social networks and self-presentations also have gender differences. In order to explore the independent effects of social network on adolescent sports, age and gender were used here as the control variables.

Furthermore, we use histogram to basically analyze the sample. From Figure 2(a), it can be seen that the sample is balanced and distributed between male and female, indicating the sample can reflect reality well. In different gender group (Figure 2(b)), the number of adolescents who are 13-14 years old is the most, while the number of 12 years old adolescents is the least.

From Figure 2(c), during the group of 12 years old adolescents, the scores of five social networks are all high, indicating that, for 12-year-olds, their social network type, density, and intensity are stronger. For 13-14 years old adolescents, all types of social networks are not high, especially online social network. Its average value is 1.2 which means the online social network is not strong. This may be because the academic pressure for adolescents in the stage is higher than other groups. For 15-16 years old adolescents, mixed social network and emotional social network are stronger than other social networks, and for 17-18 years old adolescents, online social network, mixed social network, and emotional social network are stronger than other two social networks. From Figure 2(d), no matter how old adolescents are, their sports activities are frequent. For example, for 13-14 years old adolescents, most adolescents do sports more than once a week; only 5% of adolescents do sports once a week or even less than once a week. From Figure 2, we can clearly know the distribution of sample, the social networks of different group, and the sports behavior of different groups; it helps to lay the foundation for the following empirical research.

#### 3.2. Results

**3.2.1. Influence of Social Networking on Adolescent Sports Behavior.** Econometric models are constructed to test the relationships between different social networks and adolescent sports behavior:

$$\text{Sport}_i = \beta X_i + \mathbf{P} \mathbf{Q} + e_1. \quad (10)$$

Among them, Sport is adolescent sports behavior,  $X$  is an adolescents' social network that includes an online/offline social network and an instrumental/emotional/mixed social

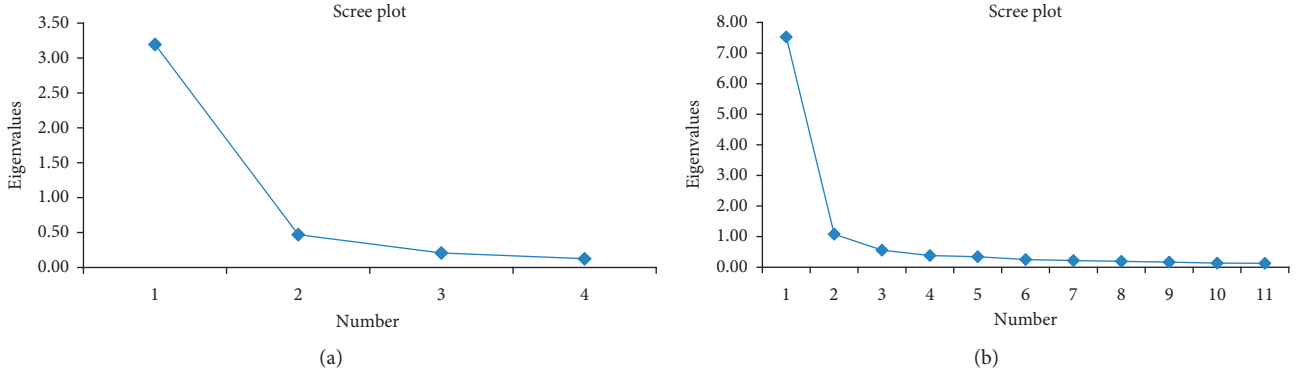


FIGURE 1: Scree plot of adolescent sports (a) and social network (b).

TABLE 1: Mean, standard deviation, and correlation coefficient of variables.

	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Age	14.53	1.69	1										
2. Gender	0.53	0.50	-0.018	1									
3. Sports	2.88	0.93	-0.042	-0.076	1								
4. Social network	2.08	0.87	-0.004*	-0.105*	0.858*	1							
5. Online social network	2.09	0.89	0.015*	-0.084	0.744**	0.967**	1						
6. Offline social network	2.36	0.64	-0.015	-0.113*	0.903**	0.988**	0.916**	1					
7. Mixes social network	2.30	0.71	-0.018	-0.126**	0.876**	0.945*	0.863**	0.966**	1				
8. Instrumental social network	2.34	0.66	0.000	-0.107*	0.734**	0.944**	0.923**	0.924**	0.823**	1			
9. Emotional social network	14.53	1.69	-0.033	-0.066	0.965**	0.844**	0.724**	0.891**	0.850**	0.711**	1		
10. Social efficacy	2.41	0.71	-0.058	-0.078	0.980**	0.838**	0.725**	0.884**	0.880**	0.712**	0.916**	1	
11. Self-presentation	2.43	0.96	0.006	-0.116**	0.808**	0.810**	0.710**	0.848**	0.895**	0.662**	0.816**	0.780**	1

Note:  $N = 568$ , gender is a dummy variable, female = 0, and male = 1. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

network,  $\mathbf{Q}$  is the control variable matrix, and  $\mathbf{e}_1$  is the regression residual. The empirical results are shown in Table 2.

The results in Table 2 indicate that adolescents' social networks have a significant positive effect on their sports behavior ( $\beta = 0.860, p < 0.001$ ). The wider the homogeneity of these adolescents' social relationship is, the more frequent their sports behavior becomes. Online social networks and offline social networks have positive effects on adolescent sports behavior, respectively. However, when adolescents have both an offline social network and an online social network, the offline social network has a positive effect on adolescent sports ( $\beta = 1.381, p < 0.001$ ), while the online social network has a negative effect on adolescent sports ( $\beta = 0.517, p < 0.001$ ). An instrumental network, emotional network, and mixed network have positive effects on adolescent sports, respectively, with coefficients of 0.735, 0.964, 0.880, and are significant at a 1% level. However, when adolescents have the three kinds of above social relations at the same time, the emotional network and the mixed network still have positive effects on adolescent sports, but the influence of the instrumental network on adolescent sports then becomes insignificant.

**3.2.2. Mediating Effect of Social Efficacy.** According to Wen and Ye [18], it is necessary to test the parameters of three regression equations to verify the mediating effect:

$$\text{Sport}_i = c\text{SN}_i + \mathbf{PQ} + \mathbf{e}_2, \quad (11)$$

$$\text{SE}_i = \alpha\text{SN}_i + \mathbf{PQ} + \mathbf{e}_3, \quad (12)$$

$$\text{Sport}_i = c'\text{SP}_i + b\text{SE}_i + \mathbf{PQ} + \mathbf{e}_4. \quad (13)$$

Among these, SN is social network, SE is social efficacy,  $\mathbf{Q}$  is the control variable matrix, and  $\mathbf{e}_2 \sim \mathbf{e}_4$  are the regression residuals. Further,  $c$  is the total effect of the independent variable (social network) on the dependent variable (adolescent sports behavior);  $a$  is the effect of social network on the intervening variable (social efficacy);  $b$  is the effect of social efficacy on adolescent sports behavior after controlling for the influence of social network;  $c'$  is the direct effect of social network on adolescent sports behavior after controlling for the influence of social efficacy. The mediating effect is tested in five steps: first, test the coefficient  $c$  of equation (11); if  $c$  is significant, the intermediary effect is significant; otherwise, there is a masking effect. But whether it is significant or not, follow-up tests are carried out.

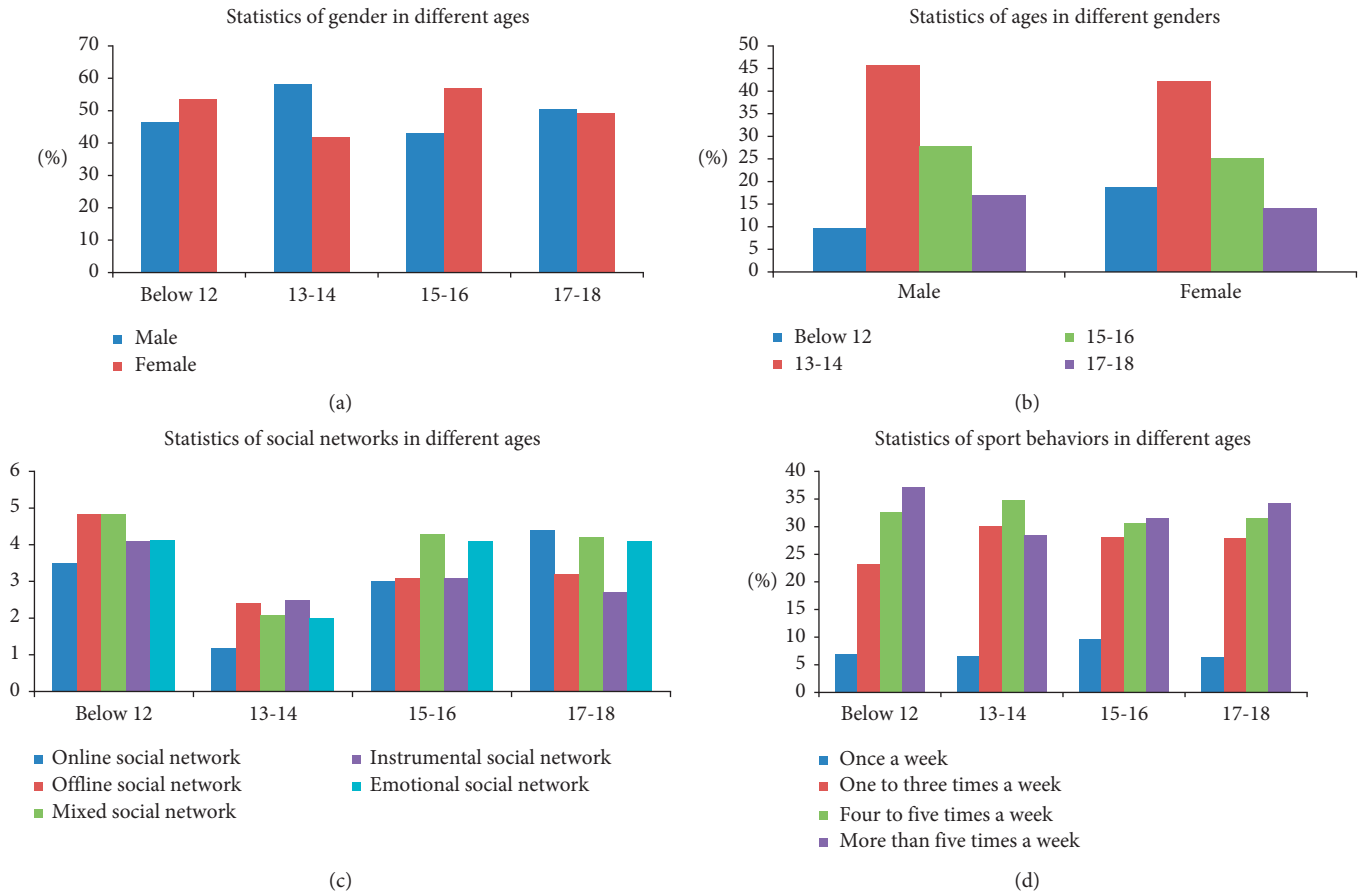


FIGURE 2: Statistic of adolescent social networks and sports behavior.

TABLE 2: Estimated results of the influence of social network on adolescent sports behavior.

Variables	Adolescent sports behavior							
Constant	0.316 (1.568)	0.471* (1.794)	0.210 (1.240)	0.069 (0.465)	0.357 (1.336)	0.100 (0.966)	0.179 (0.945)	0.102 (1.071)
Age	-0.023* (-1.667)	-0.031* (-1.767)	-0.016 (-1.427)	-0.007 (-0.737)	-0.025 (-1.366)	-0.006 (-0.859)	-0.015 (-1.157)	-0.007 (-1.113)
Gender	0.028 (0.599)	-0.028 (-0.458)	0.053 (1.356)	0.074** (2.154)	0.004 (0.068)	-0.024 (-1.004)	0.070 (1.596)	0.004 (0.203)
Social network	0.860*** (37.052)	—	—	—	—	—	—	—
Online social network	—	0.744*** (24.673)	—	-0.517*** (-12.194)	—	—	—	—
Offline social network	—	—	0.905*** (45.545)	1.381*** (32.455)	—	—	—	—
Instrumental network	—	—	—	—	0.735*** (23.892)	—	—	0.016 (0.818)
Emotional network	—	—	—	—	—	0.964*** (81.163)	—	0.793*** (38.244)
Mixed network	—	—	—	—	—	—	0.880*** (40.217)	0.189*** (7.352)
Adj- $R^2$	0.737	0.554	0.816	0.858	0.538	0.931	0.767	0.942

Note:  $t$  value is in parentheses. Symbols \*, \*\*, and \*\*\* indicate that coefficient is significant at 10%, 5%, and 1% levels.

Second, test the coefficient  $a$  of equation (12) and the coefficient  $b$  of equation (13) in turn; if both are significant, then the indirect effect is significant; go to the fourth step; if at least one is not significant, go to the third step. Third, use Bootstrap method to directly check  $H_0: ab = 0$ . If it is significant, then the indirect effect is significant; go to the fourth step; otherwise, the indirect effect is not significant; stop the analysis. Forth, examining the coefficient  $c'$  of equation (13), if it is not significant, the direct effect is not

significant, indicating that there is only an intermediary effect. If it is significant, that is, the direct effect is significant, go to the fifth step. Fifth, compare the signs of  $ab$  and  $c'$ . If they are the same, it is partly intermediary effect. If the sign is different, it is a masking effect. If  $a$ ,  $b$  and  $c$  are all significant, the mediation effect is significant; otherwise, there are other effects. For example, if  $c$  is not significant, there is a masking effect; if  $c'$  is significant, there also exist direct effects. Besides the above five steps, a nonparametric percentile Bootstrap



method was used to test the mediating effect. If the 95% confidence interval does not contain a zero, then the mediating effect becomes significant. The empirical regression results and significance test results for the mediating effect are shown in Table 3.

The results in Table 3 show that  $R^2$  in the three equations was bigger than 0.7 and the goodness of fit for the model was good. In equation (11), social network has a significant positive effect on adolescent sports behavior,  $c = 0.860, p < 0.001$ . In equation (12), social network has a significant effect on social efficacy,  $c = 0.839, p < 0.001$ . In equation (13), social network and social efficacy have a significant effect on adolescent sports behavior,  $c' = 0.122, b = 0.879, p < 0.001$ . According to the mediating effect test procedure of Wen and Ye [18], the mediating effect of social efficacy is significant because  $a, b, c,$  and  $c'$  are all significant. Further, the Bootstrap method was used to test the mediating effect of social efficacy. 5000 Bootstrap samples were randomly selected from the original sample for indirect effect estimation. Table 3 shows the estimated value of the indirect effect of social efficacy to be 0.7374 and 95% confidence intervals were [0.7007, 0.7704], excluding zero. Thus, the mediating effect of social efficacy became significant.

**3.2.3. Moderating Effect of Self-Presentation.** As for the moderating effect of self-presentation, the following econometric model was constructed:

$$\text{Sport}_i = \alpha \text{SN}_i + \beta \text{SP}_i + \gamma \text{SN}_i * \text{SP}_i + \mathbf{PQ} + e_5, \quad (14)$$

where SP is the self-presentation. The stepwise regression method was used to estimate the moderating effect of self-presentation. The results are shown in Table 4.

The results of the study show that although both social networking and self-presentation have significant positive effects on adolescent sports behavior, the interaction terms of social network and self-presentation do not have any significant influence on adolescent sports behavior, thereby indicating that the moderating effect of self-presentation is not significant.

In addition, this study found that age has a negative predictive effect on adolescent sports behavior; that is, the older the adolescents are, the less likely they are to exercise. Gender had no significant effect on the prediction of adolescent sports behavior.

### 3.3. Discussion

**3.3.1. Influence of Social Networks on Adolescent Sports Behavior.** The results of this study indicate that social networks have significant influences on adolescent sports behavior, demonstrating that adolescents with larger and denser social networks have more frequent sports behavior. People are mutually linked to each other, and so are their sports behaviors. Thus, the existence of social networks means that persons' movements are interdependent [19]. Sports behavior can be spread from person to person, and

there is also a phenomenon called "peer effects" [20], which is just as the proverb says: "If you live with a lame person you will learn to limp." However, it is interesting that, with the strengthening of the social network establishment, the effects of online social networks and offline social networks on adolescent sports behavior are not the same. When adolescents have both an offline social network and an online social network, the offline social network has a positive effect on adolescent sports, while the online social network has a negative effect on adolescent sports. This outcome may be because, at the present stage, Chinese adolescents make friends online mainly through games, chats, and other ways [21]. Moreover, adolescents in junior and senior high schools have heavy daily learning tasks and usually surf the Internet for leisure and entertainment. Therefore, most adolescents' online social network relationships are entertainment groups. Besides, during the formation of online network relationships, most adolescents will use computers or mobile phones for a long time, a factor that further reduces the possibility of adolescent sports [22]. Under the influence of the characteristics of network groups and the time used for online communication, offline social networks have a negative influence on adolescent sports behavior.

For social networks with different characteristics, the instrumental network, emotional network, and mixed network have positive effects on adolescent sports, respectively. In other words, without the influence of others, social relations, such as those with counselors, teachers, family members, and friends, have positive effects on adolescent sports behaviors. However, when adolescents have the above three kinds of social relations at the same time, the influence of the instrumental network, that is, counselors or teachers, on adolescent sports behavior is not significant. That is to say, the influence of counselors or teachers on adolescent sports behavior is easily affected by other people. This outcome may be because the instrumental network connectivity is not strong [23]. Even for teachers and students who live and study very closely, there is no similar influence like those of friends, indicating that although physical activity has the interpersonal infectivity of social network, its infectivity changes based on relative activities and the genders of friends [24]. For example, in a given social network, less active runners tend to affect more active runners, while the opposite is not true. Both men and women affect men, while only women affect other women [25]. The reason why it is easy for physical activity behaviors to spread through adolescent friendship or a family network is that first of all, adolescents are more willing to experience peer relations during adolescence, and their peer groups are based on common behaviors (sports, computer games, video games, diet, etc.). These directly or indirectly affect their physical activities and health [26]. Secondly, family relationship is the network relationship with the strongest connection among adolescents. Under the influence of that strong connection, an emotional network will have an obvious influence on adolescent sports behavior.

Overall, different types of social networks have positive effects on adolescent sports behavior separately, but different social networks are also interacted, so their influences on adolescent sports behavior will change by the interactions.

TABLE 3: Estimated results and significance test results of mediating effect of social efficacy.

Variables	Equation (11) Sports	Equation (12) Social efficacy	Equation (13) Sports	
Constant	0.316 (1.568)	0.460** (2.148)	-0.088 (-1.191)	
Age	-0.023* (-1.667)	-0.032** (-2.228)	0.006 (1.123)	
Gender	0.028 (0.599)	0.018 (0.364)	0.012 (0.716)	
Social network	0.860*** (37.052)	0.839*** (34.099)	0.122*** (7.916)	
Social efficacy	—	—	0.879*** (56.91)	
Adj-R <sup>2</sup>	0.737	0.706	0.956	
<i>Indirect effect(s) of SN on sport</i>				
Intervening variable: social contact	Effect 0.7374	Boot standard error 0.0178	Boot LLCI 0.7007	Boot ULCI 0.7704

Note: *t* value is in parentheses. Symbols \*, \*\*, and \*\*\* indicate that coefficient is significant at 10%, 5%, and 1% levels.

TABLE 4: Estimated results of moderating effect of self-presentation.

Variables	Adolescent sports behavior		
Constant	0.383* (1.654)	0.330* (1.768)	0.317* (1.691)
Age	-0.028* (-1.764)	-0.024* (-1.936)	-0.024* (-1.916)
Gender	0.036 (0.673)	0.049 (1.134)	0.046 (1.069)
Self- presentation	0.811*** (30.426)	0.332*** (9.093)	0.329*** (8.993)
Social network	—	0.592*** (16.231)	0.589*** (16.109)
Social network*self-presentation	—	—	0.014 (0.891)
Adj-R <sup>2</sup>	0.654	0.774	0.774

Note: *t* value is in parentheses. Symbols \*, \*\*, and \*\*\* indicate that coefficient is significant at 10%, 5%, and 1% levels.

That is to say, the effect of social networks on sports behavior not only is influenced by the strength and density of social networks but also influenced the interaction of different social networks.

**3.3.2. Mediating Effect of Social Efficacy and the Moderating Effect of Self-Presentation.** Social efficacy plays a mediating role in the relationship between social networking and adolescent sports behavior, and these results can be explained using the following perspectives.

First, according to the theoretical model of “two paths and three layers” of trust behavior, social relationship is a situational factor, and its influence on adolescent sports behavior is affected by internal cognition [27]. In a positive and extroverted social relationship, adolescent sports behaviors tend to receive more social support, which then can enable adolescents to form a more positive internal cognition of themselves and others, thereby producing a stronger motivation to undertake sports [28].

Secondly, from the perspective of self-efficacy theory, situational condition is one of the important factors that can affect individual self-efficacy [29]. If one is in a familiar and pleasant environment, individual’s self-efficacy will increase. Doing sports is an important way to relieve adolescent psychological stress and improve physical health. Thus, adolescents receive positive demonstrations and responses from their social network and their motivation to exercise will strengthen in order to achieve the recognition of their social relationships.

The moderating effect of self-presentation in the relationship between a social network and adolescent sports behavior is not significant in that the positive effect of social

network on adolescent sports behavior is not affected by adolescents’ self-presentation. This outcome may be because, on the one hand, the level of self-presentation in Chinese adolescents’ sports is low. According to the *Research Report on Internet Users’ Information Security* released by the China Internet Network Information Center (CNNIC, 2018), more than 73% of adolescents’ self-presentation focuses on entertainment and study, not so much on sports [30]. On the other hand, although self-presentation can reduce the loneliness of individuals and show them the enthusiasm of sports behaviors, it also presents certain negative information about sports, such as fatigue and spending time, thereby weakening the positive effect of self-presentation [31].

To sum up then, this study concludes that social networking not only directly predicts adolescent sports behavior but also indirectly affects their actual sports behavior through social efficacy. The implication of this mediator model is that adolescent sports behavior can also be generated through interpersonal interaction and that behavior can be affected by individual cognition, thereby showing the relevance of self-cognition gained by social interaction [32].

## 4. Conclusions

The main conclusions of this study are the following:

- (1) The social networks of adolescents have obvious positive effects on their sports behavior. Single online social networks and offline social networks, instrumental networks, emotional networks, and mixed networks have obvious positive effects on adolescent sports behavior, respectively. However, under the influence of multiple types of social networks, offline

social networking has a negative predictive effect on adolescent sports behavior, while the effect of mixed networks is not so obvious.

- (2) Social efficacy plays an intermediary role in the relationship between social networking and adolescent sports behavior. The moderating effect of self-presentation is not significant.

Although this paper found the importance of social networks in adolescents sports behavior and discovered the role of social efficacy, there are some limitations to this study, which need to be addressed in future studies. For example, all variables in this study were investigated using a questionnaire; it can measure variables in static level but may ignore diversity and complexity of variables; furthermore, this study found that the moderating effect of self-presentation is not significant. This finding may have a theoretical rationale, but it may also be due to the low degree of self-presentation of the sample. In the future, that sample needs to be expanded for further research and more precise evaluation.

## Data Availability

The survey 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 study was supported by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Project no. 18YJC890012).

## References

- [1] L. P. Spear, "The adolescent brain and age-related behavioral manifestations," *Neuroscience & Biobehavioral Reviews*, vol. 24, no. 4, pp. 417–463, 2000.
- [2] L. Steinberg and K. C. Monahan, "Age differences in resistance to peer influence," *Developmental Psychology*, vol. 43, no. 6, pp. 1531–1543, 2009.
- [3] S. Sarah-Jeanne, N. R. James, C. B. Julie et al., "Effect of peers and friends on youth physical activity and motivation to be physically active," *Journal of Pediatric Psychology*, vol. 34, no. 2, pp. 217–225, 2009.
- [4] M. Jennifer, D. L. H. Kayla, L. M. Barnett et al., "Friendship network characteristics are associated with physical activity and sedentary behavior in early adolescence," *PLoS ONE*, vol. 10, no. 12, pp. 1–15, 2015.
- [5] M. J. Edwards, R. Jago, J. S. Simon et al., "The influence of friends and siblings on the physical activity and screen viewing behaviours of children aged 5–6 years: a qualitative analysis of parent interviews," *BMJ Open*, vol. 5, no. 5, Article ID e006593, 2015.
- [6] K. D. L. Haye, G. Robins, P. Mohr et al., "How physical activity shapes, and is shaped by, adolescent friendships," *Social Science & Medicine*, vol. 73, no. 5, pp. 719–728, 2011.
- [7] Z. Li, F. Xiong, X. Wang et al., "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [8] B. Yu and X. Hu, "Toward training and assessing reproducible data analysis in data science education," *Data Intelligence*, vol. 1, no. 4, pp. 333–344, 2019.
- [9] A. Bandura, "Self-efficacy mechanism in human agency," *American Psychologist*, vol. 37, no. 2, pp. 122–147, 1982.
- [10] A. Bandura, *Self-efficacy: The Exercise of Control*, Worth Publishers, New York, NY, USA, 1997.
- [11] C. Fan and A. S. Mak, "Measuring social self-efficacy in a culturally diverse student population," *Social Behavior and Personality: An International Journal*, vol. 26, no. 2, pp. 131–144, 1998.
- [12] M. Wei, D. W. Russell, and R. A. Zakalik, "Adult attachment, social self-efficacy, self-disclosure, loneliness, and subsequent depression for freshman college students: a longitudinal study," *Journal of Counseling Psychology*, vol. 52, no. 4, pp. 602–614, 2005.
- [13] F. Xiong, W. Shen, H. Chen et al., "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, 2019.
- [14] H. Y. Mao, H. C. Hsu, and S. D. Lee, "Gender differences in related influential factors of regular exercise behavior among people in Taiwan in 2007: a cross-sectional study," *PLoS one*, vol. 15, no. 1, pp. 1–22, 2020.
- [15] S. Park, J. Y. Kang, and L. A. Chadiha, "Social network types, health, and health-care use among South Korean older adults," *Research on Aging*, vol. 40, no. 2, pp. 131–154, 2018.
- [16] E. J. Jeong and D. H. Kim, "Social activities, self-efficacy, game attitudes, and game addiction," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 4, pp. 213–221, 2011.
- [17] J. Kim and J.-E. R. Lee, "The Facebook paths to happiness: effects of the number of Facebook friends and self-presentation on subjective well-being," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 6, pp. 359–364, 2011.
- [18] Z. Wen and B. Ye, "Analyses of mediating effects: the development of methods and models," *Advances in Psychological Science*, vol. 22, no. 5, pp. 731–745, 2014.
- [19] L. G. Smith, L. Banting, R. Eime et al., "The association between social support and physical activity in older adults : a systematic review," *International Journal of Behavioral Nutrition & Physical Activity*, vol. 14, no. 1, p. 56, 2017.
- [20] K. P. Smith and N. A. Christakis, "Social networks and health," *Annual Review of Sociology*, vol. 34, no. 1, pp. 405–429, 2008.
- [21] S. Aral and C. Nicolaides, "Exercise contagion in a global social network," *Nature Communications*, vol. 8, Article ID 14753, 2017.
- [22] S. Ji, S. Pan, E. Cambria et al., "A survey on knowledge graphs: representation, acquisition and applications," 2020, <http://arxiv.org/abs/2002.00388>.
- [23] T. W. Valente, K. Fujimoto, C.-P. Chou, and D. Spruijt-Metz, "Adolescent affiliations and adiposity: a social network analysis of friendships and obesity," *Journal of Adolescent Health*, vol. 45, no. 2, pp. 202–204, 2009.
- [24] P. Manasatchakun, P. Chotiga, J. Hochwalder, . Roxberg, M. Sandborgh, and M. Asp, "Factors associated with healthy aging among older persons in northeastern Thailand," *Journal of Cross-Cultural Gerontology*, vol. 31, no. 4, pp. 369–384, 2016.
- [25] L. Adam and P. S. Damon, "The intersection of sport management and sociology of sport research: a social network perspective," *Sport Management Review*, vol. 15, no. 2, pp. 244–256, 2012.

- [26] Z. Wu, S. Pan, F. Chen et al., "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 99, 2020.
- [27] L. N. Howard, "Sport sociology, NASSS, and undergraduate education in the United States: a social network perspective for developing the field," *Sociology of Sport Journal*, vol. 27, no. 1, pp. 76–88, 2010.
- [28] Q. Zhu, "Self-disclosure in online support groups for people living with depression," Master's Thesis, National University of Singapore, Singapore, 2011.
- [29] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.
- [30] L. S. Eller, E. L. Lev, C. Yuan, and A. V. Watkins, "Describing self-care self-efficacy: definition, measurement, outcomes, and Implications : definition, measurement, outcomes, and implications," *International Journal of Nursing Knowledge*, vol. 29, no. 1, pp. 38–48, 2018.
- [31] B. Cornwell, L. P. Schumm, E. O. Laumann et al., "Assessment of social network change in a national longitudinal survey," *Journals of Gerontology*, vol. 69, no. 2, pp. 75–82, 2014.
- [32] L. Guo, Q. Zhang, W. Hu, Z. Sun, and Y. Qu, "Learning to complete knowledge graphs with deep sequential models," *Data Intelligence*, vol. 1, no. 3, pp. 224–243, 2019.

## Research Article

# The Collective Behaviors of Self-Excitation Information Diffusion Processes for a Large Number of Individuals

Lifu Wang<sup>1</sup> and Bo Shen<sup>2</sup> 

<sup>1</sup>Department of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China

<sup>2</sup>China Key Laboratory of Communication and Information Systems,  
Beijing Municipal Commission of Education Beijing Jiaotong University, Beijing, China

Correspondence should be addressed to Bo Shen; [bshen@bjtu.edu.cn](mailto:bshen@bjtu.edu.cn)

Received 25 June 2020; Accepted 24 July 2020; Published 24 August 2020

Guest Editor: Hongshu Chen

Copyright © 2020 Lifu Wang and Bo Shen. 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 opinion dynamics is a complex and interesting process, especially for the systems with a large number of individuals. It is usually hard to describe the evolutionary features of these systems. In some previous works, it has been shown that the self-excitation type model has superior performance in learning and predicting opinions. Following this line, we consider the self-excitation opinion model and study the collective behaviors of the self-excitation model. We propose a McKean–Vlasov-type integrodifferential equation to describe the asymptotic behaviors of the model and show that the introduced equation, by coupling with the initial distribution, has the ability of capturing the influence of the self-excitation process, which describes the mutually exciting and recurrent nature of individuals. We also find that the steady-state distribution is a “contraction” of the initial distribution in the linear and bounded confidence (DW model) interaction cases, which is different from the results of the model with nonself-excitation interaction.

## 1. Introduction

Expressing opinions and then influencing others are the primary forms of social behavior of people. The dynamics of public opinion has long been a major concern of social science and computational social science. There has been an increasing interest on the perception and prediction of social opinion, such as quantitative investment firms which measure investor sentiment and trade using social media Karppi and Crawford [1] and prediction of election results [2].

Based on sentiment analysis with deep learning, there are many frameworks to perceive opinion on the social network. For the prediction of social opinion, there are some noticeable models proposed at the beginning of the 21<sup>st</sup> century, such as the Sznajd model [3], Hegselmann–Krause model [4, 5], Deffuant–Weisbuch model [6], and so on, which had been established to understand the co-ordinated movements of opinion as a group. However, these works

held some limitations: (i) most of these models on opinion dynamics are theoretical and have not proved their effectiveness quantitatively. (ii) These models are represented by a cellular automaton, which makes it difficult to analyze the global behavior of the dynamics.

More recently, some researchers began to study the models, which can provide more accurate predictions [7–9]. De et al. [9] proposed a framework of opinion dynamics named SLANT, providing accurate predictions of users’ opinions. In the early models, the opinion and the states of the next moment only depend on the states of the last moment, and thus have nothing to do with more previous ones. It means that opinion formation is seen as a Markov process. However, in the real world, this assumption may not be precise. The model given in the study by DE et al. [9] indicates that self-excitation opinion models with exponential function weight have superior performance in learning and predicting opinions.

In many works [9–13], self-excitation models have shown efficient abilities to capture users' behavior pattern. These models focus on the microbehaviors basing on cellular automata, which would not express the global dynamics of opinion evolution analytically.

In this paper, we study the self-excitation opinion dynamics on the large homogeneous network. We show that when the number of the individuals is very large, the opinion distribution of this system evolves according to a nonlinear partial differential equation (PDE) of McKean–Vlasov type. Due to the dependence on the initial value in the local rules, the individuals with different initial values will have different dynamic equations; thus, the mean-field equation deriving process is different from the cases of the previous works [14, 15]. We show that this PDE can be decomposed into two parts where the first part is the same as the Fokker–Planck model in the study by Toscani [16]. And the second part can capture the mutually exciting and recurrent nature of individual behaviors by an integral term related to the initial value. This term leads to different steady state behaviors from the Markov model, such as the model given in the study by Toscani [16]. Generally, the stationary solutions of the Toscani type model will trend to a single-point support Dirac measure when noise parameters tend to 0, but the model with additional term makes the steady state to be a “contraction” of the initial distribution.

The rest of this paper is organized as follows. Section 2 is for related work. In Section 3, we introduce the mathematical model of opinion dynamics we study and the main mean-field theorem in this paper. In Section 4, using this theorem, we consider the linear interaction model and the bounded confidence (DW) model and study the steady-state distribution of these models. Section 5 is the proof outline of this theorem. Section 6 is a summary. The detailed proof is in the appendix.

## 2. Related Work

**2.1. Information Diffusion Model.** The early theory of information diffusion in mathematical sociology was proposed by Sznajd in the study by Sznajd–Weron and Sznajd [3]. Sznajd's model is designed to explain the features of opinion dynamics, in which every individual is on a lattice and have a state 0 or 1 to express the opinion. Individuals update their opinion by the opinions of their neighbors. The numerical simulations of this model are also investigated by many researchers [17–19]. On the complete graph, Slanina and Lavicka [20] show that when the number of individuals is very large, the probability densities evolve according to the partial differential equations

$$\frac{\partial}{\partial \tau} P(m, \tau) = \frac{\partial^2}{\partial m^2} \left[ (1 - m^2) P(m, \tau) \right], \quad (1)$$

$$\frac{\partial}{\partial \tau} P(m, \tau) = -\frac{\partial^2}{\partial m^2} \left[ (1 - m^2) m P(m, \tau) \right], \quad (2)$$

where  $P(m)$  is the probability distribution of opinion. Equation (1) describes the Ochrombel simplification of the

Sznajd model, and (2) does the original case. The authors showed the existence of phase transition in the original formulation and smooth behavior without transition in the model modified by Ochrombel.

For the theory of continuous opinion dynamics, the Deffuant–Weisbuch model [6] and Hegselmann–Krause model [4, 5] deal with two different cases. Both these models are so-called bounded confidence models, i.e., the individuals can only be influenced by some others whose opinions are close enough to theirs. The characteristics of bounded confidence lead to the formation of communities, which coincides with the phenomenon in the real world. A very beautiful analysis of the convergence of the Deffuant–Weisbuch model is given in the study by Zhang et al. [21], which reveals the community structure of the model. Bhattacharyya [22] proves the convergence of the Hegselmann–Krause model. On the other hand, numerical simulations on complex networks [23] show that “the more heterogeneous the complex network is, the weaker the ability of polarization and consensus of the complex network will be.” Models with both continuous opinion and discrete states are also considered, such as the SHIR model [24], where the individuals have a four-state variable to represent susceptible, hesitated, infected, and removed. Opinion dynamics of the social network and the impact on the recommendation are studied in the studies by Xiong et al. [25] and Li et al. [26].

Como and Fagnani [27] gave a general framework to prove that the scaling limits of the many-body continuous opinion system can be described by a measured value ordinary differential equation (ODE). In such models, every individual will change its opinion to  $\omega X + (1 - \omega)X'$ , a convex combination with the neighbor with probability  $k(X, X')$ . When  $k(X, X') = 1_{[0, \epsilon]}(|X - X'|)$ , it becomes the Deffuant–Weisbuch model. The key ideas of their work are the large deviation principle and some estimations based on the Lipschitz property, which are the same as the general method used to prove the chaotic propagation properties [28].

Based on the Boltzmann equation with granular gas like interactions, Toscani [16] introduced a collision model, in which the communication between individuals was considered as a collision with the following form:

$$w' = (1 - \eta P(w, w_*))w + \eta P(w, w_*)w_* + \zeta D(w), \quad (3)$$

where function  $D$  and  $P$  describe the local relevance of the compromise and diffusion for a given opinion. Toscani also considered the quasi-invariant opinion limit for the Boltzmann-type equation. The main idea of that is to scale the interaction frequency, strength, and the diffusion in the integral equation, and then, the equation reduces to a Fokker–Planck type equation. Then, on the base of this model, considering the influence of the structure of the social network, Albi et al. [29] consider the kinetic opinion dynamics on the large scale networks evolving over time. The evolutions of both the network and the opinion were involved.

**2.2. Self-Excitation Point Process in Social Systems.** Point processes type models are generally used to analyze the impact of events on the system such as CSMA/CD [30], financial data [31], and the dynamics of book sales [32]. A very succinct and effective analysis for the time series of daily views for videos on YouTube is introduced [33]. The authors showed that most of the video viewing record data could be described statistically by a Poisson process, but still, about 10% data show a point process with the instantaneous rate  $\lambda(t)$ ,

$$\lambda(t) = V(t) + \sum_{i,t_j \leq t} \mu_i \phi(t - t_i), \quad (4)$$

where  $\mu_i$  is the number of potential viewers influenced directly by person  $i$  who views the video through the social network, and  $V(t)$  captures all spontaneous views that are not triggered by epidemic effects on the network.

For the self-excitation social opinion models, De et al. [9] proposed an opinion model on a social network, in which every user has a latent opinion about the given topic and can post messages about the topic. The experiments on real data show that the self-excitation model, with the intensity depending on the messages sent from a neighbor in the past and the weight function having the form like  $e^{-\omega t}$ , performs much more better than the Poisson process ( $\omega = 0$ ) model.

Wang et al. [13] used point process to predicting user activities. They proposed a generic framework for point process prediction problem, and they used a mass transport equation to update the transport rate and compute the conditional mass function.

**2.3. Mean Field Theory.** Mean-field theory (MFT) is introduced to study many-body problem by using a single averaged effect to approximate the effect of all individuals. The core of MFT is to estimate the error of the mean-field approximation. If there is no interaction, all individuals are independent of each other, and then, the law of large numbers works and gives the mean-field approximation. In the general case, we still want to be able to use the law of large numbers. That is to say, if one picks a chaotic (i.i.d) initial distribution of particles, we hope that this distribution is still chaotic as time-evolving, which is the so-called propagation of chaos.

The theory of propagation of chaos in the cases of Wiener noise is summed in the study by Sznitman [28]. If  $N$  particles with initial ‘‘chaotic’’ distribution  $u_0^{\otimes N}$  satisfy the stochastic differential equation (SDE),

$$dX_i = dW_i + \frac{1}{N} \sum_j b(X_i, X_j) dt, \quad (5)$$

the McKean–Vlasov mean-field equation is

$$dX = dW + \int b(X, y) u(dy) dt, \quad (6)$$

where  $u$  is the law of  $X$ . The method to show this is to use the Lipschitz character of  $b$  and Gronwall’s lemma. McKean–Vlasov equation for SDE with Poisson jumps is in the study by

Andreis et al. [15]. The authors used Burkholder–Davis–Gundy inequality for martingales. The main step in the study by Andreis et al. [15] is to use Doob–Meyer decomposition and estimate the compensated Poisson process.

Using the large deviation principle [34], Arous and Guionnet [35] studied the mean-field simplification dynamics for Langevin spin glass. The framework of their proof makes use of Sanov’s theorem and Varadhan’s lemma in the path space  $C[0, T]$  to get the rate function for the interaction system.

The main obstacle to use these methods to prove MFT is the unboundedness of coefficients in SDE. Because of Varadhan’s lemma, if we use a stop time to get the localization of SDE, we can get the so-called local large deviation principle. Dawson and Gartner [36] gave some compactness criteria to convert the ‘‘local’’ result into ‘‘global’’ one. They also show that a Lyapunov function for the system of weakly interacting diffusions will let the compactness condition be satisfied. Puhalskii [37] set up the whole framework of local to global LDP and introduce the  $C$ -exponential tightness conditions.

### 3. Mathematical Model

**3.1. Self-Excitation Opinion Model.** In this paper, we consider a simplified version of the model [9] and study the large number limit of the self-excitation opinion model with the following form:

There are  $N$  individuals with opinion  $x_i$ , which is a function of the time  $t$ , such that

$$x_i(t) = x_i(0) + \sum_j h(x_j, x_i) k(t) * dN_j(t), \quad (7)$$

where  $h(x_j, h_i)$  is a interaction function,  $*$  is the convolution operator,  $k(t) = e^{-\omega t}$  is the triggering kernel, and  $N(t)$  is a Poisson process. This model has been previously studied [11, 12].

The original model proposed in the study by De et al. [9] has the following form:

$$\begin{aligned} x_i(t) &= b_i + \sum_j \alpha_{i,j} \sum_{t_j \in H_j(t)} m_j k(t - t_j) \\ &= b_i + \sum_j \alpha_{i,j} m_j k(t) * dN_j(t), \end{aligned} \quad (8)$$

where  $m_j$  is the message sent by user  $j$ ,  $\alpha_{i,j}$  is the interaction intensity between  $i$  and  $j$ ,  $b_i = x_i(0)$  is the initial opinion of user  $i$ ,  $H_j(t)$  is the history of events up to time  $t$  for user  $j$ , and  $k(t)$  is a triggering kernel. The values of message  $m$  come from a sentiment distribution related to the opinion  $x$  and  $E[m_i | x_i] = x_i$ .

As the models have been previously studied [11, 12], we only consider the opinion part, the simplified model (7). In the case that  $h(x_i, x_j) = h(x_j)$ , where  $x_j$  is the opinion of the individual  $j$ , De et al. and Wang et al. [9, 11] showed the validity of the model in the study by DE et al. [9] on the real world data. And when  $h(x_j, x_i) = x_j - x_i$ , the form of this model will degenerate into the form of the DW model [6] and Toscani model [16] when  $k(t) = 1$ .

In studies by De et al. and Wang et al. [9, 11], the self-excitation opinion model with exponential function type weight was shown to be very effective in the opinion dynamics; therefore, we set  $k(t) = e^{-\omega t}$ , where  $\omega$  is a constant. Due to the fact that differential of the convolution of two functions  $d(f * g) = f(0)g + g * df$ , where there is a jump part in  $g$ , and  $dk(t) = -\omega k(t)dt$ . We can use a stochastic differential equation (SDE) to describe  $x(t)$  [11].

$$dX_i = \omega(b_i - X_i(t))dt + \sum_j \alpha h(X_j, X_i) dN_j. \quad (9)$$

It is also necessary to add  $\sigma dW$  into the model to represent noise, where  $W$  is the Wiener noise with  $W[0] = 0$ .

$$dX_i = \omega(X_i(0) - X_i(t))dt + \sum_j \alpha h(X_j, X_i) dN_j + \sigma dW_i. \quad (10)$$

The above two equations are given in the study by Wang et al. [11]. Because of the dependence on the path of the history, it is apparently not a Markov process. We consider the limit that the population size tends to infinity. Let  $X^N = (X_1, X_2, \dots, X_N)$  be the opinion variable of the  $N$ -particles system.  $X$  satisfies

$$dX_i = \omega(b_i - X_i(t))dt + \frac{1}{N} \sum_{j \neq i} \alpha h(X_j, X_i) dN_j + \sigma dW_i, \quad (11)$$

where  $b_i = X_i[0]$ . In the case that  $\omega = 0$  and  $h(x, y) = (x - y)1_{[0, \epsilon]}(|x - y|)$ , this is a Deffuant–Weisbuch-like model.

If we only consider the case  $\omega = 0$  and replace the Poisson Noise  $dN$  with  $dt$  and consider the form

$$dX_i = dW_i + \frac{1}{N} \sum_j h(X_i, X_j) dt. \quad (12)$$

The mean-field theory of this equation is well studied [28]. The McKean–Vlasov mean-field equation is

$$dX = dW + \int b(X, y) u(dy) dt. \quad (13)$$

In this paper, we first use Doob–Meyer decomposition for  $dN$ . Then, we estimate the compensated Poisson process by Burkholder–Davis–Gundy inequality for martingales. This method helps us to reduce our equation to the form in (12). As for the dependence on the initial values, we regard the initial as a random media. Using the methods in the study by Pra and Hollander [38]. The McKean–Vlasov equation can be obtained using Sanov’s theorem and Varadhan’s lemma in the path space  $C(0, T)$ . In this paper, we use a direct estimation to prove the convergence to the McKean–Vlasov equation.

**3.2. Main Results.** The main result in this paper is the following theorem:

**Theorem 1.** *When  $N$  tends to infinity, there is a limit process  $X$  for  $X^N$  in equation (11), with SDE:*

$$dX = \int dP(y) h(y, X) \lambda dt + \omega(b - X) dt + \sigma dW, \quad (14)$$

where  $P$  is the law of  $X$ , and  $b = X[0]$  is a random variable. And given the initial distribution  $\mu$ , the law of  $X$  is  $P = \int P^b d\mu(b)$ , where  $P^b$  is the solution of the Fokker–Planck equation associated with the McKean–Vlasov process:

$$\begin{aligned} \frac{\partial}{\partial t} P^b &= -\frac{\partial}{\partial x} [\beta^{b,P} P^b] + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} P^b, \\ P^b(0) &= \delta_b, \end{aligned} \quad (15)$$

$$\beta^{b,P} = \int \mu(db') \int P^{b'}(dy) \alpha \lambda h(y, x) + \omega(b - x),$$

where  $\delta_b$  is a Dirac measure with support on  $b$ .

From this theorem, we can study the asymptotic behavior of the self-excitation opinion model. In Section 4, we will study the steady state of this system by using Fokker–Planck equation and show the effects of self-excitation, which makes it impossible to give a perfect consensus, such that it can avoid Abelson’s diversity puzzle [39].

In order to prove this theorem, we need some assumptions. Stemming from the physical meaning, we can assume that the support of initial distribution  $\mu(b)$  is bounded, such as the uniform distribution on  $[0, 1]$ . Also, for the proof of McKean–Vlasov limit, we need the function  $h(x)$  to have good enough properties to make

$$E \left[ h \left( X_j^b, X_i^b \right) \right] < \infty, \quad (16)$$

$$\int x^2 P(dx) < \infty, \quad (17)$$

where  $P$  is the solution of the McKean–Vlasov equation.

If  $h$  is bounded, the first condition is satisfied trivially. In the appendix, we will give other examples of  $h$  that can make these conditions be satisfied. Also, we assume that  $h(x, y)$  is Lipschitz for both of the variables. However, this is not the necessary condition [14].

$$|h(x_1, y_1) - h(x_2, y_2)| < L|x_1 - x_2| + K|y_1 - y_2|. \quad (18)$$

Most of these conditions can be relaxed, but we hope to prove our results without excessive technical details on tightness; therefore, these restrictions are imposed.

## 4. The Properties of McKean–Vlasov Dynamics

**4.1. McKean–Vlasov Equation.** For the equation

$$\begin{aligned} \frac{\partial}{\partial t} P^b &= -\frac{\partial}{\partial x} [\beta^{b,P} P^b] + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} P^b, \\ P^b(0) &= \delta_b, \end{aligned} \quad (19)$$

$$\beta^{b,P} = \int \mu(db') \int P^{b'}(dy) \lambda h(y, x) + \omega(b - x),$$

we consider a simple case



$$\begin{aligned}\lambda &= 1, \\ \sigma &= 1, \\ h(y, x) &= \alpha(y - x).\end{aligned}\quad (20)$$

Then, we can see that

$$\beta^{b,P} = \alpha(m - x) + \omega(b - x), \quad (21)$$

where  $m = \int x dP(x)$ , the mean value of  $x$ .

Since  $h$  is an odd function, it is easy to see that

$$\begin{aligned}\frac{dE(X(t))}{dt} &= E\left[\alpha \int P(dY)(Y - X) + \omega(b - X(t))\right] \\ &= 0 + \omega E[X(0) - X(t)].\end{aligned}\quad (22)$$

So,  $E(X(t) - X(0)) = Ce^{-\omega t}$  since  $E(X(0) - X(0)) = 0$ ,  $C = 0$ , and  $m = E(X(t)) = E(X(0))$ .

Then, we have the Fokker-Planck equation associated with the Mckean-Vlasov process:

$$\frac{\partial}{\partial t} P^b = -\frac{\partial}{\partial x} [\alpha(m - x) + \omega(b - x)P^b] + \frac{1}{2} \frac{\partial^2}{\partial x^2} P^b. \quad (23)$$

Integral over  $\mu(b)$

$$\frac{\partial}{\partial t} P = -\frac{\partial}{\partial x} [(\alpha(m - x))P + \omega \int d\mu(b)(b - x)P^b] + \frac{1}{2} \frac{\partial^2}{\partial x^2} P. \quad (24)$$

Another example is the bounded confidence type model [4, 6] where every agent interacts only within a certain level of confidence:

$$h(y, x) = (y - x)k(y - x), \quad (25)$$

where  $k(y - x)$  is a continuous function, and if  $y - x < \Delta_1$ ,  $k(y - x) = 1$ , and if  $y - x > \Delta_2$ ,  $k(y - x) = 0$ . Since, in this case,  $h(x)$  is bounded and Lipschitz, our assumptions are satisfied.

Note that for the case that  $h(y, x) = \alpha y$ , which is used in the SLANT system [9],

$$\frac{\partial}{\partial t} P^b = -\frac{\partial}{\partial x} [\alpha m(t) + \omega(b - x)P^b] + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} P^b, \quad (26)$$

where  $m(t) = \int x dP(x, t)$ . Since

$$\begin{aligned}\frac{dE(X(t))}{dt} &= E\left[\alpha \int P(dY)Y + \omega(b - X(t))\right] \\ &= (\alpha - \omega)E(X) + \omega E(X(0)).\end{aligned}\quad (27)$$

As a sanity check, this formula matches (4) in the study by De et al. [9]. Solving this equation, we get

$$\begin{aligned}m(t) + \frac{\omega}{\alpha - \omega} m(0) &= Ce^{(\alpha - \omega)t}, \\ C &= \frac{\alpha}{\alpha - \omega}.\end{aligned}\quad (28)$$

*4.2. Discussion.* Como and Fagnani [27] showed that when the population size tends to infinity, the limit behavior of the Deffuant-Weisbuch model can be described by a measure-valued ODE. And also, Toscani introduced a kinetic model of opinion formation [16]. In this model, the opinion is changed by the binary interaction of collision, and the dynamics of the opinion distribution is modeled by Boltzmann-type integrodifferential equation.

In these models, the dynamics of opinion distribution is depicted as follows:

$$\frac{d}{dt} \langle \phi, \mu_t \rangle = \langle \phi, H(\mu_t) \rangle, \quad (29)$$

where  $\mu_t$  is the measure of opinion distribution, and  $\phi$  is an arbitrary test function. The right side of (29) has the following form [27]:

$$\langle \phi, H(\mu_t) \rangle = \iint (\phi(1 - \omega)x + \omega y) - \phi(x)k(x, y) d\mu(x) d\mu(y), \quad (30)$$

and in Boltzmann case [16],

$$\begin{aligned}\langle \phi, H(\mu_t) \rangle &= \iint \phi(w') + \phi(w'_*) - \phi(w) - \phi(w_*) d\mu(w) d\mu(w_*), \\ w' &= 1(-\eta P(w, w_*))w + \eta P(w, w_*)w_*, \\ w'_* &= (1 - \eta P(w_*, w))w + \eta P(w_*, w)w.\end{aligned}\quad (31)$$

Intuitively, both models can be derived from the stochastic differential equation with Poisson jump interactions by the method mentioned in the study by Méléard [14]. Events that interact with other particles are regarded as a Poisson process.

For the system with  $h(y, x) = y - x$ ,

$$\frac{\partial}{\partial t} P = -\frac{\partial}{\partial x} [(\alpha(m - x))P + \omega \int d\mu(b)(b - x)P^b] + \frac{1}{2} \frac{\partial^2}{\partial x^2} P, \quad (32)$$

in the case that  $\omega = 0$ , which is the nonself-excitation case, it degenerates into the general opinion dynamics model, or equally, the quasi-invariant opinion limit for Toscani's Boltzmann equation [16]:

$$\frac{\partial}{\partial t} P = \frac{\lambda}{2} \frac{\partial^2}{\partial x^2} (D(|x|^2)P) - \frac{\partial}{\partial x} (P(|x|)(m - x)P). \quad (33)$$

The  $\omega$  term is a modification to consider the influence of the self-excitation process.

In this paper, we only consider the homogeneous network. However, using the method in the study by Pra and Hollander and Albi et al. [38, 40], we can study the large system of agents interacting through a network with a given distribution of the number of connections by the following way:

- (1) Consider a random media  $c \in \{0, 1, 2, \dots, c_{\max}\}$ , which is a discrete variable describing the number of connections.  $c$  should be a random variable to

represent the complex network. For example,  $c$  obeys power law distribution.

- (2) Construct a new interaction equation as in the study by Albi et al. [40].

$$dX_i = \omega(b_i - X_i(t))dt + \sum_j \alpha h(X_j, X_i, c) dN_j. \quad (34)$$

Then, it is the same as Theorem 1 in the study by Pra and Hollander [38] such that in this random media, we can also prove a mean-field theorem. And there is a distribution  $P(c, x, t)$  about the variable  $c$ , opinion  $x$ , and time  $t$ . Using a similar argument as Albi et al. [40], we have obtained the stationary solutions.

#### 4.3. Steady-State Distribution of the Linear Case.

Considering a simple case of initial distribution  $P_0 = 1/2\delta_{-10} + 1/2\delta_{10}$ . We will use this simple form to explain how the initial distribution is coupled to the equation. Then, the equation will become

$$\begin{aligned} \frac{\partial}{\partial t} P^{[-10]} &= -\frac{\partial}{\partial x} \left[ (\alpha(\bar{x} - x))P^{[-10]} + \omega(-10 - x)P^{[-10]} \right] \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial x^2} P^{[-10]}, \\ \frac{\partial}{\partial t} P^{[10]} &= -\frac{\partial}{\partial x} \left[ (\alpha(\bar{x} - x))P^{[10]} + \omega(10 - x)P^{[10]} \right] \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial x^2} P^{[10]}. \end{aligned} \quad (35)$$

The overall distribution  $P$  is equal to  $1/2P^{[-10]} + 1/2P^{[10]}$ .

The evolution process of this distribution is shown in in Figure 1(a). And Figures 1(b) and 1(c) show the differences between two kinds of steady states.

As shown in Figure 1(b), for the initial distribution  $P_0$ , the steady state distribution is normal. In fact, this is true for any initial distribution, and we have the following theorem.

**Theorem 2.** *In the case that  $\omega = 0$ , for all type of initial distribution  $P_0$ , we have the steady-state distribution  $P(\infty) = 1/Z e^{\alpha(x-m)/\sigma^2}$ , where  $Z$  is the normalization constant.*

This is because the form of PDE for  $P$  has the form

$$\frac{\partial}{\partial t} P = -\frac{\partial}{\partial x} (\alpha(m - x))P + \sigma^2 \frac{1}{2} \frac{\partial^2}{\partial x^2} P. \quad (36)$$

The steady-state equation

$$-\frac{\partial}{\partial x} (\alpha(m - x))P + \sigma^2 \frac{1}{2} \frac{\partial^2}{\partial x^2} P = 0, \quad (37)$$

and this steady state is asymptotically stable, such that when  $t \rightarrow \infty$ , the solution of the equation will trend to the steady state. It is easy to see that the steady state is a normal distribution as shown in Figure 1(b).

From Theorem 2, this  $\omega = 0$  model cannot describe the actual situation since, in this model, the final state of the opinion distribution has no community structures but a perfect consensus (in fact, this is true for any monotonic function  $h$ , since the steady state is the Gibbs distribution of a potential energy  $V(x)$ , such that  $\nabla V = h$ ), which is so-called Abelson's diversity puzzle [39], a persistent research puzzle in the social sciences. Generally, this puzzle is solved by the bounded-confidence mode, such as HK [4] and DW [6] models. It is proved in the study by Como and Fagnani [27] for a large number of individuals case and by Zhang et al. [21] for the general case that the steady state of the DW model has the form  $\sum_i C_i \delta_{x_i}$ , where  $x_i$  is the opinion of the community  $i$ , and  $|x_i - x_j| > R$  if  $i \neq j$ , where  $R$  is the bounded confidence distance.

However, in the case  $\omega > 0$ , the equation has the steady state as in Figure 1(c) when the initial distribution has the form  $1/2(\delta_{-10} + \delta_{10})$ . In the general case, we have

**Theorem 3.** *In the case that  $\omega > 0$ , given an initial distribution  $\mu(x)$ , we have the steady-state distribution  $P(x, \infty) = 1/Z \int \mu(b) e^{-((\alpha+\omega)x - (am+\omega b)/\sigma^2)}$ , where  $Z$  is the normalization constant.*

Let  $P = \int d\mu(b) P^b$ , and  $P^b$  is the solution of Mckean-Vlasov equation,

$$\frac{\partial}{\partial t} P^b = -\frac{\partial}{\partial x} \left[ \alpha(m - x) + \omega(b - x)P^b \right] + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} P^b. \quad (38)$$

The steady state of  $P^b$  is easy to calculate:

$$P^b(x, \infty) = \frac{1}{Z} e^{-((\alpha+\omega)x - (am+\omega b)/\sigma^2)}, \quad (39)$$

where  $Z$  is the normalization constant, so we have  $P(\infty) = \int d\mu(b) P^b(\infty)$ .

Considering the limit behavior  $\sigma \rightarrow 0$ , the steady state of Mckean-Vlasov equation when  $h(y, x) = y - x$  has the form

$$P(x, \infty) = \int d\mu(b) \delta\left(x - \frac{\alpha m + \omega b}{\alpha + \omega}\right). \quad (40)$$

So that even in the case that the variance of the noise trends to 0, this equation will not converge to a single Dirac measure, which is different from the  $\omega = 0$  case. Intuitively, the term  $\omega(b - x)$  makes the opinion try not to deviate from the initial position too far, which can be considered as soft-bounded confidence.

In the case that  $h(y, x) = \alpha y$ , which is used in the SLANT system in the study by De et al. [9], supposing  $\omega > \alpha$ , the steady state will have the similar form since the equation of which is

$$-\frac{\partial}{\partial x} \left[ \alpha m + \omega(b - x)P^b \right] + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} P^b = 0, \quad (41)$$

where  $m = m(\infty) = (\alpha/\omega - \alpha)m(0)$  (this is from (25)), the final mean value of  $x$ . We have  $P(\infty) = \int d\mu(b) \delta(x - (\alpha m + \omega b)/\omega)$ .

#### 4.4. Steady-State Distribution of Bounded Confidence Models.

In the case that  $\omega = 0$  and  $h(y, x) = (y - x)k(y - x)$ , the steady-state distribution is studied [27]. We can follow their methods to study the weak solution of our equation.

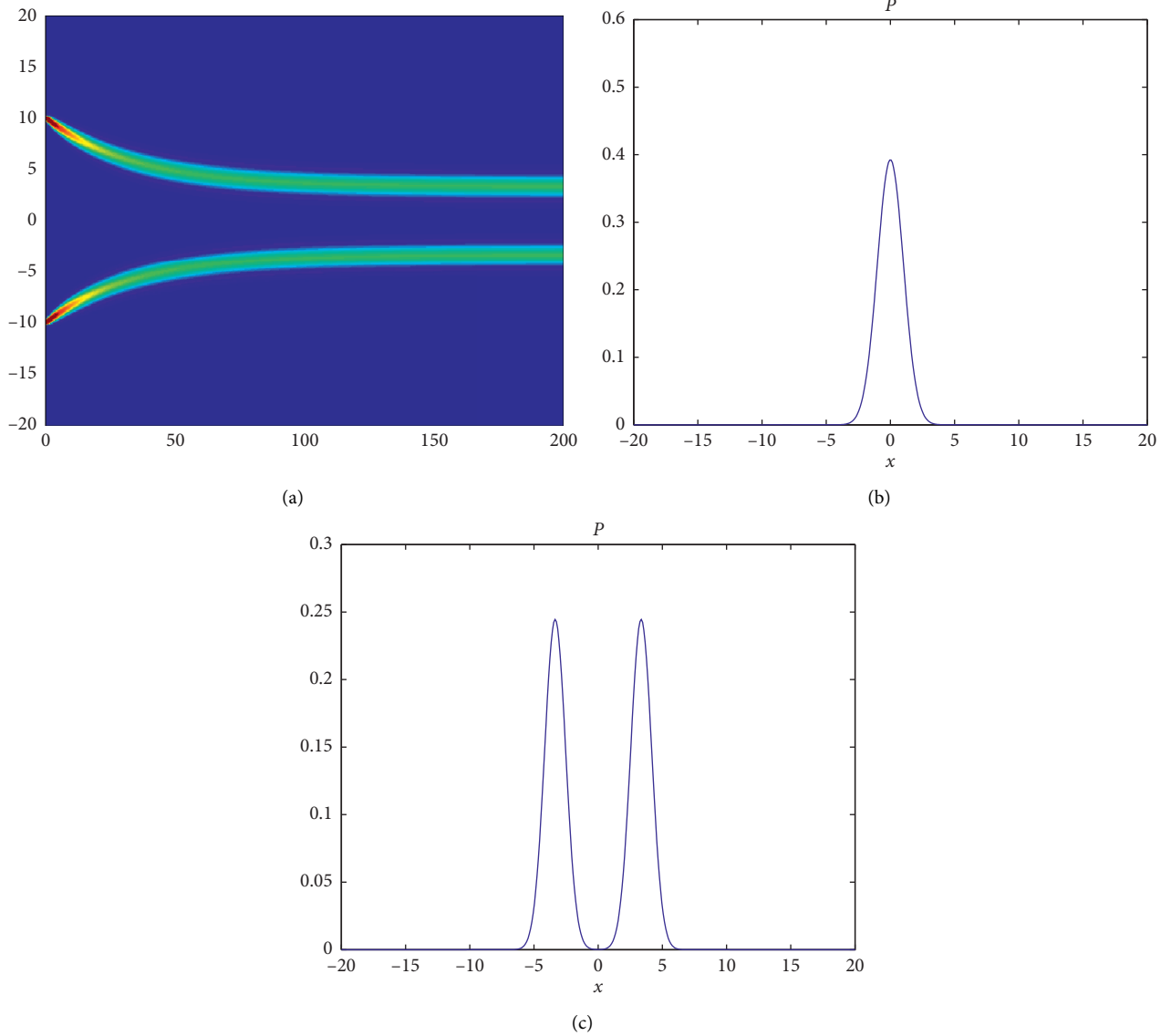


FIGURE 1: Simulation results. (a) Evolution of the opinion density, with  $s=0:02$ ;  $a=0:02$ ;  $w=0:01$ . (b) Steady state for the dynamics with parameters  $s=0:02$ ;  $a=0:02$ ;  $w=0$ . (c) Solution profiles at time  $t=200$ , with parameters  $s=0:02$ ;  $a=0:02$ ;  $w=0:01$ .

**Theorem 4.** Let  $\omega = 0$ ,  $h(y, x) = (y - x)k(y - x)$ , and  $k(z) = 0$  if  $\|z\| > R$ , otherwise  $k(z) > 0$ . We have  $P(x, \infty) \rightarrow \sum_i C \delta(x - x_i)$  when  $\sigma \rightarrow 0$ , where  $\|x_i - x_j\| > R$  if  $i \neq j$ .

*Proof.* Considering equation (19) with  $\omega = 0$ , let  $Q_t = \int \mu(b)x^2 P^b(x) = \int x^2 P(dx)$ . We have

$$\begin{aligned}
 \frac{d}{dt} Q_t &= - \int \mu(db) x^2 \frac{\partial}{\partial x} \int P(dy) [(y - x)k(y - x) + \omega(b - x)P^b(dx)] + \frac{1}{2} \int \sigma^2 x^2 \frac{\partial^2}{\partial x^2} P(dx) \\
 &= \int 2x(y - x)k(y - x)P(dy)P(dx) \frac{1}{2} \int \sigma^2 x^2 \frac{\partial^2}{\partial x^2} P(dx) \\
 &\stackrel{(a)}{=} - \int (y - x)^2 k(y - x)P(dy)P(dx) + \sigma^2 \int P(dx),
 \end{aligned} \tag{42}$$

where  $k(y-x)$  is a bounded confidence kernel. In (a) of (42), we use the symmetry of  $P(x)$  and  $P(y)$  and the integration by parts.

Considering the steady state and  $\sigma \rightarrow 0$  limit, we have  $d/dt(Q_t) = 0$ , yet if in the support of  $P(\infty)$ , there are two points  $x$  and  $y$  such that  $\|x-y\| < R$  and  $x \neq y$ , and we have

$$\begin{aligned} \frac{d}{dt}Q_t &= - \int \mu(b)x^2 \frac{\partial}{\partial x} \int P(dy) [(y-x)k(y-x)P^b(dx)] + \frac{1}{2} \int \sigma^2 x^2 \frac{\partial^2}{\partial x^2} P(dx) = \int 2x(y-x)k(y-x)P(dy)P(dx) \\ &+ \int 2\omega x(b-x)\mu(db)P^b(dx) + \frac{1}{2} \int \sigma^2 x^2 \frac{\partial^2}{\partial x^2} P(dx) = - \int (y-x)^2 k(y-x)P(dy)P(dx) \\ &+ \int 2\omega x(b-x)\mu(db)P^b(dx) + \sigma^2 \int P(dx). \end{aligned} \quad (43)$$

Since  $-\int (y-x)^2 k(y-x)P(dy)P(dx) \leq 0$  and  $\sigma \rightarrow 0$ , we have

$$\int x(b-x)\mu(db)P^b(dx) \geq 0, \quad (44)$$

such that in the sense of mean value,  $|x| < |b| = |x(0)|$ . And the degree of offset is depending on the value of  $\int (y-x)^2 k(y-x)P(dy)P(dx)$ . We can see that  $\int (y-x)^2 k(y-x)P(dy)P(dx)$  will increase as the radius  $R$  of the kernel  $k(y-x)$  grows. When  $R \rightarrow \infty$  and  $k(y-x) = 1$ , the model will degenerate into the linear case in Theorem 3.  $\square$

## 5. The Outline of the Proof

*5.1. SDE Related to Initial Values.* Since the process we study is not a Markov process, we need to consider SDE with the form

$$dX = F(X[0], X, t)dt + dW. \quad (45)$$

It has no Markov generator, but the non-Markovian of it is “not so bad”. The law of  $X$  can be solved by the following method. For a given  $X[0]$ , we can solve the law by  $e^{tA}P_0$ , where  $A$  is the generator of the Fokker–Planck equation, and  $P_0 = \delta_{x_0}$  is the initial Dirac distribution. Then, the solution of the SDE has the form

$$\int dP_0(X_0)e^{tA(X_0)}\delta_{X_0}, \quad (46)$$

where  $A(X_0)$  is the generator with given  $X[0]$ ,  $\delta_{X_0}$  is the Dirac measure with the support on  $X_0$ , and  $P_0$  is the initial distribution of  $X_0$ .

The essence of this method is to regard  $X[0]$  as an additional random variable. When SDE has the above form, we need the double layer empirical measure, which will regain the symmetry.

We will consider the SDEs

$-\int (y-x)^2 k(y-x)P(y)P(x) < 0$ , which is a contradiction, so that our claim follows.  $\square$

Supposing  $\omega > 0$ , it is hard to obtain a steady state directly. However, we can use  $Q_t$  function to give an estimation. We can show that in the bounded confidence case, the steady-state solution is still a “contraction” of initial distribution.

$$dX_i = \omega(b_i - X_i(t))dt + \sum_{j \neq i} \alpha h(X_j, X_i) dN_j + \sigma dW_i, \quad (47)$$

where  $b_i$  is a stochastic variable such that  $X_i[0] = b_i$ ; then, we will prove the large number law for the double layer empirical measure  $1/N \sum_i \delta_{X_i, b_i}$ , which will give the McKean–Vlasov limit. This is equal to consider SDEs

$$dX_i^b = \omega(b - X_i^b(t))dt + \sum_{j \neq i} \int d\mu(b') \alpha h(X_j^b, X_i^b) dN_j + \sigma dW_i, \quad (48)$$

where  $X^b$  is the stochastic process with given  $b$ , and  $E(X) = E(\int d\mu(b)X^b)$ .

*5.2. Intermediate Process.* Andreis et al. [15] proved that the McKean–Vlasov limit of the equation

$$dX_i^N = F(X_i^N)dt + \sigma(X_i^N)dW_i + \frac{1}{N} \sum_{j \neq i}^N h(X_j^N, X_i^N)dN_j, \quad (49)$$

has the form

$$dX = F(X) + \int dP(Y)h(Y, X)\lambda dt + \sigma(X)dW, \quad (50)$$

where  $P$  is the law of  $X$ .

We use the same way as Andreis et al.[15] to reduce the jump-SDE to the “averaging dynamics.” Their method is to decompose the jump terms into a martingale and a continuous part. Then, the martingale can be easily estimated by Burkholder–Davis–Gundy inequality.

We will show that there is an intermediate process  $Y_i$  with the following SDE:

$$dY_i^b = (b - Y_i)dt + \sigma dW_i + \int d\mu(b') \frac{1}{N} \sum_{j=1}^N \alpha h(Y_j^b, Y_i^b) \lambda dt, \quad (51)$$

such that

$$\int d\mu(b) \frac{1}{N} \sum_{j=1}^N E \left[ \sup \|X_i^b - Y_i^b\| \right] \leq \frac{C}{\sqrt{N}} \quad (52)$$

By comparing these two process directly, we can notice that

$$\int \mu(b) E \left[ \sup_r \|X_i^b - Y_i^b\| \right] \leq \int \mu(b) (F + \Theta), \quad (53)$$

where  $F = \omega \int_0^t \|X_i^b(s) - Y_i^b(s)\| ds$ , and  $\Theta = E[\sup_r \|1/N \sum_j \int_0^r ds (\alpha h(X_j, X_i) dN_j(s) - \alpha h(Y_j, Y_i) \lambda ds)\|]$ .

Using Doob–Meyer decomposition,

$$dN = d\tilde{N} + \lambda dt, \quad (54)$$

where  $\tilde{N}$  is a martingale.

Intuitively, the supremum of a martingale would not be very large. This can be proved by Burkholder–Davis–Gundy inequality and our assumption on  $h$ . Then, we can obtain the approximation (52). See the second part of the appendix for a detailed proof of (52).

**5.3. Mckean–Vlasov Process.** The intermediate process has the form like

$$dX_i = \frac{1}{N} \sum_j f(X_i, X_j) dt + g(X_i, \omega_i) dt + dW, \quad (55)$$

where  $\omega_i$  is a random variable.

The asymptotic behavior of the double layer empirical measure, with the case that  $f$  and  $g$  are bounded, has been studied [38]. They use Varadhan’s lemma to reduce the system to the case without interaction. Since  $f$  and  $g$  are not bounded, we use the method in the study by Sznitman [28] to compare the two SDEs directly. Using the Lipschitz condition, we can control  $E[\sup_r \|X(r) - \bar{X}(r)\|]$  by the first Wasserstein distance  $\rho$ .

The first Wasserstein distance is defined by

$$W_1(\mu_1, \mu_2) = \inf \int_{R^d \times R^d} |x - y| \eta(dx, dy): \eta \in H(\mu_1, \mu_2), \quad (56)$$

where  $H(\mu_1, \mu_2)$  is the set of all probability measures on  $R^d \times R^d$  with marginals  $\mu_1$  and  $\mu_2$ .

By Kantorovich–Rubinstein duality theorem, it has a dual representation:

$$\sup_{\|f\|_L \leq 1} \int f(x) d(\mu_1 - \mu_2)(x), \quad (57)$$

where  $\|f\|_L$  is the minimal Lipschitz constant for  $f$ .

Using this representation, since we have assumed that the coefficients are Lipschitz, we can see that

$$E \left[ \sup_r \|X(r) - \bar{X}(r)\| \leq L \int_0^T E[\rho(\mu_X^N(t), \mu_t)] dt \right], \quad (58)$$

where  $\mu_X^N$  is the empirical measure, and  $\mu_t$  is the solution of Mckean–Vlasov equation.

For  $E[\rho(\mu_X^N(t), \mu_t)]$ , we have

$$E[\rho(\mu_X^N(t), \mu_t)] \leq E[\rho(\mu_X^N(t), \mu_X^N(t))] + E[\rho(\mu_X^N(t), \mu_t^N)], \quad (59)$$

where  $\mu_X^N$  is the empirical measure for  $N$  independent Mckean–Vlasov processes.

Then, we can derive our result by considering the Wasserstein distance between the solution of Mckean–Vlasov equation and empirical measure. Using Theorem 1 in the study by Zanella and Guillin [41], if the two order moment is finite,

$$E[\bar{X}^2] < \infty, \quad (60)$$

when the size tends to infinity and the Wasserstein distance trends to zero, so our claim follows. The detailed proof is given in the appendix.

## 6. Conclusion

In this paper, we studied the asymptotic behaviors of the exponential function weight of the self-excitation interaction system  $X^N$  on the large homogeneous network when  $N$ , the number of the individuals, trends to infinity. We proved that there is a Mckean–Vlasov process  $\bar{X}$ , such that  $\lim_{N \rightarrow +\infty} E|X^N - \bar{X}|$  is zero, whose opinion distribution evolve according to a Mckean–Vlasov type integrodifferential equation, which couples with the initial distribution. The steady state of this equation is also studied. We investigated the steady state distribution of linear and bounded confidence models and showed that the steady distribution of these models is a “contraction” of initial distribution. By virtue of the self-excitation interaction, even if we do not consider bounded confidence, it will give not a perfect consensus, but a distribution depending on the initial distribution, such as this model, can avoid Abelson’s diversity puzzle.

## Appendix

### A. The Condition under which the Assumptions Hold

In this section, we will give some examples that the assumptions can be satisfied. We will show that when  $h$  has linear growth, the assumptions (16) and (17) are satisfied.

For SDE,

$$dX_i^b = \frac{1}{N} \sum_{j \neq i} \int d\mu(b') \alpha h(X_j^b, X_i^b) dN_j + \omega(b - X_i^b) dt + \sigma dW_i, \quad (A.1)$$

and it is shown in the study by Situ [42] that since both  $b - X$  and  $h$  has linear growth, this SDE has a strong solution, and also, we can estimate the moment.

In order to prove (16), since  $h$  is linear growth, we can turn to consider the second moment by Jensen's inequality  $E[h(y, x)] \leq L(E[|y|] + E[|x|]) \leq 2L(E[x^2])^{1/2}$ . Using Ito formula,

$$\begin{aligned} \int d\mu(b) \frac{1}{N} \sum_i (X_i^b(T))^2 &= \int d\mu(b) \left[ \frac{1}{N} \sum_i (X_i^b(0))^2 + \int_0^T X_i^b(s) b - (X_i^b(s))^2 ds \right] \\ &+ \int d\mu(b) d\mu(b') \frac{1}{N} \left[ \sum_{i,j} \int_0^T 2X_i^b(s) \frac{1}{N} \alpha h(X_j^b, X_i^b) + \frac{1}{N^2} \alpha^2 \left| h(X_j^b, X_i^b) \right|^2 dN_j \right] + C + \text{martingale}. \end{aligned} \quad (\text{A.2})$$

Since  $|h(x, y)| \leq L(|X| + |Y|)$  and  $dN = \lambda dt + \text{martingale}$ , we can show that

$$\begin{aligned} \int d\mu(b) \frac{1}{N} \sum_i E(X_i^b(T))^2 &\leq \int d\mu(b) \frac{1}{N} \sum_i [X_i^b(0)^2 \\ &+ K \int_0^T E(X_i^b(s))^2 + C ds \\ &+ \int_0^T E((X_i^b(s)b - E(X_i^b(s))^2) ds)]. \end{aligned} \quad (\text{A.3})$$

We use  $E[X] \leq E[|X|] \leq (E[X^2])^{1/2} \leq K(1 + E[X^2])$  by Jensen's inequality. Then, using Gronwall lemma, our claim follows.

As for the second condition (17), the above method can be also applicable.

## B. Reduce to the Averaging Dynamics

The SDE for stochastic process  $X_i$  is

$$dX_i^b = g(b, X_i^b) dt + \sigma dW_i + \int d\mu(b') \frac{1}{N} \sum_{j \neq i} \alpha h(X_j^b, X_i^b) dN_j. \quad (\text{B.1})$$

Following the study by Andreis et al. [15], it is useful to introduce an intermediate process  $Y_i$  with SDE

$$dY_i^b = g(b, Y_i^b) dt + \sigma dW_i + \int d\mu(b') \frac{1}{N} \sum_{j=1}^N \alpha h(Y_j^b, Y_i^b) \lambda dt. \quad (\text{B.2})$$

Since the Poisson process is a semimartingale,  $N(t)$  can be decomposed into  $N(t) = \tilde{N}(t) + \lambda t$  by Doob-Meyer decomposition, where  $\tilde{N}(t)$  is a martingale. The  $\lambda dt$  in  $Y$  is the second part of Doob-Meyer decomposition of  $N$ . To simplify the symbol, we set  $\lambda = 1$ .

We will show that for large  $N$ ,

$$\int d\mu(b) \frac{1}{N} \sum_{j=1}^N E \left[ \sup_{r \in [0, t]} \|X_j^b(r) - Y_j^b(r)\| \right] \leq \frac{C}{\sqrt{N}} \quad (\text{B.3})$$

so,  $X$  can be approximated by  $Y$ .

Let

$$\begin{aligned} G_i^b &= E \left[ \int_0^t \|g(X_i^b, b) - g(Y_i^b, b)\| ds \right], \\ \Theta_i &= E \left[ \sup_r \left\| d\mu(b') \left[ \frac{1}{N} \int \sum_{j \neq i} \int_0^r \alpha h(X_j^b, X_i^b) dN_j - \frac{1}{N} \sum_j \int_0^r \alpha h(Y_j^b, Y_i^b) ds \right] \right\| \right], \end{aligned} \quad (\text{B.4})$$

so,

$$\begin{aligned} & \int d\mu(b) \frac{1}{N} \sum_i E \left[ \sup_r \|X_i^b(r) - Y_i^b(r)\| \right] \\ & \leq \int d\mu(b) \left[ \frac{1}{N} \sum_i (G_i + \Theta_i) \right]. \end{aligned} \quad (\text{B.5})$$

In our model (9),  $g(x_i, b) = \omega(b - x_i)$ , so that we have  
 $G_i = E \left[ \int_0^t \|\omega(Y_i^b - X_i^b)\| \right]$ .  
 For  $\Theta$ ,

$$\begin{aligned} \Theta_i \leq & E \left[ \sup_r \left\| \int d\mu(b') \frac{1}{N} \sum_{j \neq i} \int_0^r \alpha h(X_j^b, X_i^b) d\bar{N}(t) \right\| \right] + E \left[ \sup_r \left\| \int d\mu(b') \frac{1}{N} \sum_j \int_0^r \alpha \left( h(X_j^b, X_i^b) - h(Y_j^b, Y_i^b) \right) ds \right\| \right] \\ & + \frac{1}{N} E \left[ \sup_r \left\| \int_0^r \int d\mu(b') \alpha h(X_i^b, X_i^b) ds \right\| \right]. \end{aligned} \quad (\text{B.6})$$

As in the study by Andreis et al. [15], we can use the Burkholder–Davis–Gundy inequality for martingales.

$$\begin{aligned} \Theta_i & \leq \frac{k}{N} E \left[ \left( \sum_{j \neq i} \int_0^t \left\| \int d\mu(b') h(X_j^b, X_i^b) \right\|^2 ds \right)^{1/2} \right] + \frac{1}{N} \int_0^t E \left[ \sup_r \sum_j \alpha \left\| \int d\mu(b') h(X_j^b, X_i^b) - h(Y_j^b, Y_i^b) \right\| \right] ds + \frac{H}{N} \\ & \leq \frac{k}{N} \int_0^t E \left[ N \left( \left\| \int d\mu(b') h(X_j^b, X_i^b) \right\|^2 ds \right)^{1/2} \right] + \frac{1}{N} \int_0^t E \left[ \sum_j \alpha \left\| \int d\mu(b') h(X_j^b, X_i^b) - h(Y_j^b, Y_i^b) \right\| \right] ds + \frac{H}{N} \\ & \leq \frac{k}{N} E \left[ \left( Nt \left( \sup_r \left\| \int d\mu(b') h(X_j^b(r), X_i^b(r)) \right\| \right)^2 \right)^{1/2} \right] + \int_0^t \frac{1}{N} E \left[ \sum_j \alpha \left\| \int d\mu(b') h(X_j^b, X_i^b) - h(Y_j^b, Y_i^b) \right\| \right] ds + \frac{H}{N} \\ & \leq \frac{C}{\sqrt{N}} + L \int_0^t E \left[ \|X_i^b - Y_i^b\| \right] ds + \frac{L}{N} \sum_j \int d\mu(b') \int_0^t E \left[ \|X_j^b - Y_j^b\| \right] ds + \frac{H}{N}, \end{aligned} \quad (\text{B.7})$$

where we use the symmetry, assumption (16), and the Lipschitz condition for  $h$ . Combine the two results; then, we get

$$\begin{aligned} \int d\mu(b) \frac{1}{N} \sum_j E \left[ \sup \|X_i^b - Y_i^b\| \right] & \leq \frac{C}{\sqrt{N}} + K \int d\mu(b) \frac{1}{N} \sum_j \int_0^t E \left[ \|X_i^b - Y_i^b\| \right] ds \\ & + \frac{H}{N} \leq \frac{C}{\sqrt{N}} + K \int d\mu(b) \frac{1}{N} \sum_j \int_0^t E \left[ \sup \|X_i^b - Y_i^b\| \right] ds + \frac{H}{N}. \end{aligned} \quad (\text{B.8})$$

Then, using the standard technology of Gronwall's lemma, we get the approximation.

### C. McKean–Vlasov process

As shown in the previous section, we can turn to analyze the stochastic process  $Y$ . Let  $\bar{X}$  be a stochastic process with SDE:

$$d\bar{X} = \int dP(y)\alpha h(y, \bar{X})dt + \omega(b - \bar{X})dt + \sigma dW, \quad (\text{C.1})$$

where  $P$  is the law of  $\bar{X}$ , and  $b$  is a random variable, such that  $b = X[0]$ .

Using the trick in (47), we can show that

$$P(y) = \int d\mu(b)P^b(y), \quad (\text{C.2})$$

where  $P^b$  is the law of SDE with given  $b$ , the initial distribution is  $\delta_b$ , and  $\mu(b)$  is the given initial distribution of  $\bar{X}$ . We will show that

$$\lim_{N \rightarrow \infty} \int d\mu(b) \frac{1}{N} \sum_i E \left[ \sup_{t \leq T} |Y_t^{b,i} - \bar{X}_t^{b,i}| \right] = 0. \quad (\text{C.3})$$

Let  $Q^N$  be the empirical measures  $\sum_j (1/N)\delta_{\bar{X}_j}$ , we have

$$\begin{aligned} Y_t^{b,i} - \bar{X}_t^{b,i} &= \int_0^t \frac{1}{N} \int d\mu(b') \sum_j h(Y_s^{b,i}, Y_s^{b',j}) - \int h(\bar{X}_s^{b,i}, y) Q(dy) ds + \omega \int_0^t \bar{X}_s^{b,i} - Y_s^{b,i} ds \\ &= \int_0^t ds \frac{1}{N} \int d\mu(b') \sum_j h(Y_s^{b,i}, Y_s^{b',j}) - h(\bar{X}_s^{b,i}, Y_s^{b,j}) + (h(\bar{X}_s^{b,i}, Y_s^{b,j}) - h(\bar{X}_s^{b,i}, \bar{X}_s^{b,j})) \\ &\quad + \frac{1}{N} \sum_j \int d\mu(b') (h(\bar{X}_s^{b,i}, \bar{X}_s^{b',j}) - \int h(\bar{X}_s^{b,i}, y) Q(dy)) + \omega \int_0^t \bar{X}_s^{b,i} - Y_s^{b,i} ds. \end{aligned} \quad (\text{C.4})$$

Since  $h$  is Lipschitz, we can see that

$$\begin{aligned} \int d\mu(b) E \left[ \left| \sup_r Y^{b,i}(r) - \bar{X}^{b,i}(r) \right| \right] &\leq L \int d\mu(b) \sup_r \int_0^r ds \left( E \left[ |Y^{b,i} - \bar{X}^{b,i}| \right] \right) + \frac{1}{N} \sum_j \int d\mu(b') E \left[ |Y^{b',j} - \bar{X}^{b',j}| \right] \\ &\quad + \int d\mu(b) E \left[ \frac{1}{N} \sum_j \int d\mu(b') h(\bar{X}^{b,i}, \bar{X}^{b',j}) - \int h(\bar{X}^{b,i}, y) Q(dy) + \omega \int_0^t \|\bar{X}_s^{b,i} - Y_s^{b,i}\| \right]. \end{aligned} \quad (\text{C.5})$$

Summing over  $i$ ,

$$\begin{aligned} &\frac{1}{N} \sum_i \int d\mu(b) E \left[ \sup_r |Y^{b,i}(r) - \bar{X}^{b,i}(r)| \right] \\ &\leq \frac{1}{N} \left\{ L \sup_r \int_0^r ds \sum_i \int ds(b) \left[ E \left[ |Y^{b,i} - \bar{X}^{b,i}| \right] \right] + E \left[ \frac{1}{N} \sum_j \int d\mu(b') h(\bar{X}^{b,i}, \bar{X}^{b',j}) - \int h(\bar{X}^{b,i}, y) Q(dy) + \omega \int_0^t \sum_i E \|\bar{X}_s^{b,i} - Y_s^{b,i}\| \right] \right\}, \end{aligned} \quad (\text{C.6})$$

then, using Gronwall's lemma

$$\begin{aligned} &\frac{1}{N} \sum_i \int d\mu(b) E \left[ \sup_r |Y^{b,i}(r) - \bar{X}^{b,i}(r)| \right] \\ &\leq L(T) \sup_r \int_0^r \int d\mu(b) E \left[ \frac{1}{N} \sum_j \int d\mu(b') h(\bar{X}^{b,i}, \bar{X}^{b',j}) - \int h(\bar{X}^{b,i}, y) Q(dy) \right] ds. \end{aligned} \quad (\text{C.7})$$



So, we only need to estimate the right side.

Using the Lipschitz condition for  $h$  again, this can be controlled by  $E[\rho[(Q_X^N(t), Q_{\bar{X}}(t))]]$  where  $\rho$  is the Wasserstein distance between the two measures, and  $Q_{\bar{X}}$  is the solution of McKean–Vlasov equation (72).  $Q^N$  is the empirical measures  $\sum_j (1/N)\delta_{x_j}$ . It follows from [41] Theorem 1 that if the moments is bounded (this is satisfied by (17))

$$E\left[\overline{X^2}\right] < \infty. \quad (\text{C.8})$$

Thus, our claim follows.

## Data Availability

This paper is not data-driving, and the model we study is from other data-driving ones. The data used in this article are only the data of the simulation of PDE, which are contained in Figure 1.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors thank Wenyu Zhang for his great help in writing and Prof. Haibo Wang who introduced the recent developments in the mean-field game theory to us. This research was funded by the National Key Research and Development Program of China, grant number 2018YFC0831300.

## References

- [1] T. Karppi and K. Crawford, “Social media, financial algorithms and the hack crash,” *Theory, Culture & Society*, vol. 33, no. 1, pp. 73–92, 2016.
- [2] A. Ceron, L. Curini, and S. M. Iacus, “Using sentiment analysis to monitor electoral campaigns,” *Social Science Computer Review*, vol. 33, no. 1, pp. 3–20, 2015.
- [3] K. Sznajd-Weron and J. Sznajd, “Opinion evolution in closed community,” *International Journal of Modern Physics C*, vol. 11, no. 6, pp. 1157–1165, 2000.
- [4] R. Hegselmann and U. Krause, “Opinion dynamics driven by various ways of averaging,” *Computational Economics*, vol. 25, no. 4, pp. 381–405, 2005.
- [5] R. Hegselmann and U. Krause, “Opinion dynamics and bounded confidence models, analysis and simulation,” *Journal of Artificial Societies & Social Simulation*, vol. 5, no. 3, p. 2, 2002.
- [6] G. r. Weisbuch, G. Deffuant, F. d. r. Amblard, and J.-P. Nadal, “Meet, discuss, and segregate!” *Complexity*, vol. 7, no. 3, pp. 55–63, 2002.
- [7] A. Das, S. Gollapudi, and K. Munagala, “Modeling opinion dynamics in social networks,” in *Proceedings of the ACM International Conference on Web Search and Data Mining*, pp. 403–412, New York, NY, USA, 2014.
- [8] A. De, S. Bhattacharya, P. Bhattacharya, N. Ganguly, and S. Chakrabarti, “Learning a linear influence model from transient opinion dynamics,” in *Proceedings of the ACM International Conference on Conference on Information and Knowledge Management*, pp. 401–410, Shanghai, China, 2014.
- [9] A. De, I. Valera, N. Ganguly, S. Bhattacharya, and M. G. Rodriguez, *Learning and Forecasting Opinion Dynamics in Social Networks*, 2015, <http://arxiv.org/abs/1506.05474v3>.
- [10] K. Zhou, H. Zha, and L. Song, “Learning social infectivity in sparse low-rank networks using multi-dimensional Hawkes processes,” in *Proceedings of Machine Learning Research, PMLR*, C. M. Carvalho and P. Ravikumar, Eds., in *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, vol. 31, pp. 641–649pp. 641–, Scottsdale, AZ, USA, 2013.
- [11] Y. Wang, E. Theodorou, A. Verma, and L. Song, “A stochastic differential equation framework for guiding online user activities in closed loop,” in *Proceedings of Machine Learning Research, PMLR*, A. Storkey and F. Perez-Cruz, Eds., in *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, vol. 84, pp. 1077–1086pp. 1077–, Playa Blanca, Spain, 2018.
- [12] Y. Wang, G. Williams, E. Theodorou, and L. Song, “Variational policy for guiding point processes,” in *Proceedings of the International Conference on Machine Learning*, pp. 3684–3693, Sydney, Australia, 2017.
- [13] Y. Wang, X. Ye, H. Zha, and L. Song, “Predicting user activity level in point processes with mass transport equation,” in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio et al., Eds., vol. 30, pp. 1645–1655, Curran Associates, Inc., Red Hook, NY, USA, 2017.
- [14] S. Méléard, “Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models,” in *Probabilistic Models for Nonlinear Partial Differential Equations*, pp. 42–95, Springer, Berlin, Germany, 1996.
- [15] L. Andreis, P. D. Pra, and M. Fischer, “McKean-Vlasov limit for interacting systems with simultaneous jumps,” 2017, <http://arxiv.org/abs/1704.01052v1>.
- [16] G. Toscani, “Kinetic models of opinion formation,” *Communications in Mathematical Sciences*, vol. 4, no. 3, pp. 481–496, 2006.
- [17] A. T. Bernardes, D. Stauffer, and J. Kertész, “Election results and the Sznajd model on Barabasi network,” *The European Physical Journal B*, vol. 25, no. 1, pp. 123–127, 2002.
- [18] D. Stauffer, “Better being third than second in a search for a majority opinion,” *Advances in Complex Systems*, vol. 5, no. 1, pp. 97–100, 2002.
- [19] D. Stauffer and P. M. C. de Oliveira, “Persistence of opinion in the Sznajd consensus model: computer simulation,” *The European Physical Journal B-Condensed Matter*, vol. 30, no. 4, pp. 587–592, 2002.
- [20] F. Slanina and H. Lavicka, “Analytical results for the Sznajd model of opinion formation,” *The European Physical Journal B-Condensed Matter*, vol. 35, no. 2, pp. 279–288, 2003.
- [21] J. Zhang, G. Chen, Y. Hong, J. Zhang, G. Chen, and Y. Hong, “Convergence analysis of asymmetric homogeneous deffuant-weisbuch model,” in *Proceedings of the China Conference on Control and Decision-Making*, pp. 2394–2399, Taiyuan, China, 2012.
- [22] A. Bhattacharyya, M. Braverman, B. Chazelle, and H. L. Nguyen, “On the convergence of the hegselmann-krause system,” in *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science, ITCS’13*, pp. 61–66, ACM, New York, NY, USA, 2013.
- [23] L. Guo and X. Cai, “Bifurcation phenomena of opinion dynamics in complex networks,” in *Proceedings of the International Conference on Complex Sciences*, pp. 1146–1153, Shanghai, China, 2009.

- [24] Y. Liu, S.-M. Diao, Y.-X. Zhu, and Q. Liu, "SHIR competitive information diffusion model for online social media," *Physica A: Statistical Mechanics and its Applications*, vol. 461, pp. 543–553, 2016.
- [25] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, 2019.
- [26] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [27] G. Como and F. Fagnani, "Scaling limits for continuous opinion dynamics systems," *The Annals of Applied Probability*, vol. 21, no. 4, pp. 1537–1567, 2011.
- [28] A. S. Sznitman, "Topics in propagation of chaos," *Biophysical Journal*, vol. 108, no. 2, p. 501, 1991.
- [29] G. Albi, L. Pareschi, and M. Zanella, "Opinion dynamics over complex networks: kinetic modeling and numerical methods," *Kinetic & Related Models*, vol. 10, no. 1, pp. 1–32, 2016.
- [30] B. J. Cho, *A Simulation Study on Interference in CSMA/CA Ad-Hoc Networks Using Point Process*, Aalto University, Helsinki, Finland, 2010.
- [31] V. Chavez-Demoulin and J. A. McGill, "High-frequency financial data modeling using Hawkes processes," *Journal of Banking & Finance*, vol. 36, no. 12, pp. 3415–3426, 2012.
- [32] F. Deschates and D. Sornette, "Dynamics of book sales: endogenous versus exogenous shocks in complex networks," *Physical Review E*, vol. 72, no. 2, Article ID 016112, 2005.
- [33] R. Crane and D. Sornette, "Robust dynamic classes revealed by measuring the response function of a social system," *Proceedings of the National Academy of Sciences*, vol. 105, no. 41, pp. 15649–15653, 2008.
- [34] J.-D. Deuschel and D. W. Stroock, "Large deviations," vol. 342 American Mathematical Society, Providence, RI, USA, 2001.
- [35] G. B. Arous and A. Guionnet, "Large deviations for Langevin spin glass dynamics," *Probability Theory and Related Fields*, vol. 102, no. 4, pp. 455–509, 1995.
- [36] D. A. Dawson and J. Gärtner, "Large deviations from the McKean-Vlasov limit for weakly interacting diffusions," *Stochastics*, vol. 20, no. 4, pp. 247–308, 1987.
- [37] A. Puhalskii, *Large Deviations and Idempotent Probability*, Chapman & Hall/CRC, London, UK, 2001.
- [38] P. D. Pra and F. D. Hollander, "McKean-Vlasov limit for interacting random processes in random media," *Journal of Statistical Physics*, vol. 84, no. 3-4, pp. 735–772, 1996.
- [39] R. P. Abelson, "Mathematical models of the distribution of attitudes under controversy," in *Contributions to Mathematical Psychology*, N. Frederiksen and H. Gulliken, Eds., Holt McDougal, New York, NY, USA, 1964.
- [40] G. Albi, L. Pareschi, and M. Zanella, "Opinion dynamics over complex networks: kinetic modelling and numerical methods," *Kinetic & Related Models*, vol. 10, no. 1, pp. 1–32, 2017.
- [41] N. Zanella and A. Guillin, "On the rate of convergence in Wasserstein distance of the empirical measure," *Probability Theory & Related Fields*, vol. 36, no. 3, pp. 1–32, 2013.
- [42] R. Situ, *Theory of Stochastic Differential Equations with Jumps and Applications: Mathematical and Analytical Techniques with Applications to Engineering*, Springer Science & Business Media, Berlin, Germany, 2006.

## Research Article

# Statistical Analysis of Dispelling Rumors on Sina Weibo

Yue Wu,<sup>1</sup> Min Deng,<sup>1</sup> Xin Wen,<sup>1</sup> Min Wang,<sup>1</sup> and Xi Xiong<sup>1,2</sup> 

<sup>1</sup>School of Computer and Software Engineering, Xihua University, Chengdu 610039, China

<sup>2</sup>School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China

Correspondence should be addressed to Xi Xiong; [flyxiongxi@gmail.com](mailto:flyxiongxi@gmail.com)

Received 17 June 2020; Accepted 15 July 2020; Published 19 August 2020

Guest Editor: Liang Wang

Copyright © 2020 Yue 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.

Analyzing the process and results of dispelling rumors is a prerequisite for designing an effective anti-rumor strategy. Current research on this subject focuses on the simulation experiments, short of empirical study. By using the False Information Publicity Results of Sina Weibo as the data source of empirical research, this article compares the typical features of rumor and anti-rumor accounts. Furthermore, taking COVID-19 as the target topic, distributions of the reported time, frequency, platform penalty levels, and diffusion parameters of rumors related to COVID-19 are given, and some interesting results are obtained.

## 1. Introduction

With the rapid development of online social media, such as Twitter, Facebook, YouTube, Sina, etc., a large amount of unreliable information is generated and propagated at an unprecedented rate. Compared with truth, rumor is proved to be more novel and inspired fear, disgust, and surprise in replies. Statistical results on big data showed that false rumors diffused significantly farther, faster, deeper, and more broadly than the truth [1]. Rumors not only have a strong ability to spread but also can pose a threat to the network security and social stability [2]. For instance, on April 23, 2013, fake news claiming two explosions happened in the White House and Barack Obama got injured was widely diffused on Twitter, caused severe social panic and a loss of \$136.5 billion in the stock market [3]. The negative impact is more dangerous and destructive when faced with sudden public events (public health incident, natural disaster, accident disaster, social security incident, and economic crisis). For example, when the coronavirus disease 2019 (COVID-19) spreads all over the globe, numerous rumors appeared on social media. Some of them are masking healthy behaviors and providing incorrect suggestions that result in poor physical and mental health outcomes [4]. In Iran, more than 700 people have died after ingesting toxic methanol because false rumors it helps cure the coronavirus [5]. In

Nigeria, there are several cases of overdose of chloroquine after the rumors about the effectiveness of chloroquine for treating COVID-19 spread [6]. Many other rumors caused psychological panic and panic buying of groceries.

Recognizing the dangers of rumors, Tedros Adhanom Ghebreyesus, Director-General of the WHO, said, “We’re not just fighting an epidemic; we’re fighting an infodemic,” at a gathering of foreign policy and security experts in Munich, Germany, in mid-February, referring to fake news that “spread faster and more easily than this virus.” In response, a team of WHO “Mythbusters,” including Facebook, Google, Pinterest, Tencent, Twitter, TikTok, and YouTube, unite to fight against the spread of rumors [7]. In early February, Facebook announced that it would be removing false claims and conspiracy theories about the disease posted on its social media platforms [8]. On March 18, Twitter updated its safety policy to prohibit tweets that “could place people at a higher risk of transmitting COVID-19.” [9]. Sina Weibo’s official account created a special home page for fighting with the rumors; meanwhile, it encouraged users to report fake news [10].

Social media platforms set different strategies for dispelling rumors, including deleting posts and accounts, deducting scores, lowering credit ratings, tagging rumor labels, training professional whistleblowers, and so on. In the field of extreme information confrontation, it has been

verified that, for achieving the best results, four strategies should be balanced. The strategies are deleting accounts, disintegrating clusters, supporting anti-hate clusters, and setting competing hate narratives [11, 12]. However, during the actual rumor confrontation, the effect of the strategies has not been analyzed, nor the features of rumor and anti-rumor accounts, and the dispelling process. To discover more patterns from real data, we will conduct an in-depth empirical analysis based on the False Information Publicity Results of Sina Weibo.

In this paper, Section 2 introduces related work. Section 3 reports the results of empirical analysis, including the different features between rumor spreaders and whistleblowers and the confrontation process of rumors related to COVID-19.

## 2. Related Work

It is always difficult to distinguish true information from false rumor, especially if the data are well formatted and structured [13]. Thus, source credibility is vital as external information in the task of rumor detection. As the COVID-19 pandemic continues to spread around the world, people are suggested to get the accurate and timely news from trusted healthcare professionals and government officials, such as WHO, CDC, CIDRAP, and so on. [14]. Furthermore, some institutions, such as the new Coronavirus Misinformation Tracking Center of NewsGuard [15], provide over 157 reports about low-quality and questionable new outlets that spread the rumor about COVID-19, which reminds users to stay away from the unreliable sources. Other organizations, such as Media Bias/Fact Check, provide users with a reputation search for news sites [16]. These professional reputation scores are valuable when facing with a large number of news. However, not all websites have reputation scores, and most messages are provided by anonymous users in social media. How to rank the credibility of information sources automatically and effectively is a big challenge. In unsupervised and semisupervised rumor detection methods, different features of information sources are analyzed. Yang et al. [17] analyzed the characteristics of accounts, including the user's identity, personal descriptions, gender, avatar type, name type, registration time, place, number of followers, friends, and posts. The classifying accuracy of using account-based features alone is 72.6%. These features have been used widely in subsequent methods, such as hierarchical neural networks [18]. Besides the classical features, Liu et al. [19] considered the features of the credibility identification, diversity, and the relationship between profile location and event location. Han and Guo [20] classified sources into three levels from low to high and analyzed the corresponding source descriptions, characteristics, and examples. It is proved that these features are valid on some data sets. However, whether they are always effective in various topics on different platforms requires further statistical analysis. After distinguishing rumors from the truth, the next task is to dispel rumors. For

solving this problem, many researchers designed several kinds of simulation models. Zhang et al. [21] improved the ICSAR model (I: ignorance; C: information carrier; S: information spreader; A: information advocate; R: removal) to accurately analyze rumor propagation and refutation in metropolises. Zhang et al. [22] designed a novel rumor and authoritative information propagation model in complex social networks considering the superspreading phenomenon and analyzed the dynamic interaction between rumor and authoritative information. Bodaghi et al. [23] proposed a novel model based on two hypotheses: one is that the number of followers influences the impact level of the received rumor/anti-rumor messages by those users, and the second deals with the effect of time on the impact of rumor/anti-rumor posts. The evaluation of the model supported their hypotheses. Zhang et al. [24] designed a novel two-stage rumor propagation and refutation model with time effect for online social networks and obtained the corresponding mean-field equations in both homogeneous and heterogeneous networks. Askarizadeh et al. [25] proposed an evolutionary game model to analyze the rumor process in the social network considering the impacts of people's decisions on rumor propagation and control. The analysis shows that propagation of convincing anti-rumor messages and locating rumor control centers impact debunking the rumor. In addition to modeling, Takahashi and Igata [26] investigated the difference of word distribution between rumor and anti-rumor tweets in the two instances. Zhiyuan et al. [27] found that a large number of anti-rumor accounts reported messages about themselves, and many rumor spreaders are spammers.

Some of our former works have been applied in the area of extracting reliable information in the social networks. Xiong et al. [28, 29] made effective suggestion of location despite a large amount of rumor and noise. A study by Xiong et al. [30] helps to recognize the real information from web data.

Different from previous works, this article focused on empirical analysis and tried to discover useful patterns of dispelling rumors.

## 3. Empirical Analysis

*3.1. Data Preprocessing.* Sina Weibo is a popular social networking site in China with more than 500 million registered users [31]. In this paper, the False Information Publicity Results of Sina Weibo is used as the data source. There are 39297 records of anti-rumors dated from 2012-5-29 to 2020-5-8. Monthly data goes up and down frequently, and yearly data arrived at the peak in 2014 and then remained stable in the following years, as shown in Figure 1.

For analyzing the text, three kinds of features, including users, Weibo content, and dispelling results, are collected, as illustrated in Table 1. Then ICTCLAS [32] is used for word segmentation, and each word is given a weighted score computed by TF-IDF [33]. In the case study, rumors are filtered according to the topic keywords, such as COVID-19, coronavirus, pandemic, and so on.

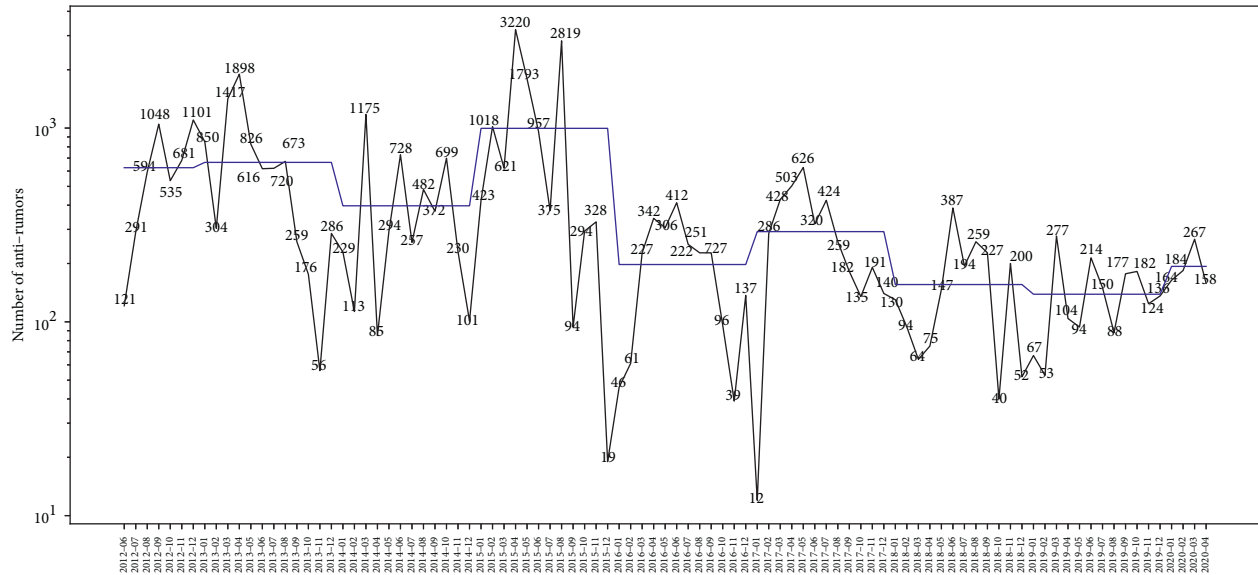


FIGURE 1: Number of anti-rumors over time. Black line represents the monthly data, and the blue line is the average data of each year.

TABLE 1: Instructions of features in the false information publicity results.

Category	Description	Features
Users	Rumor spreaders	User name, location, label, description, number of followers, friends, posts, education background, credit rank
	Anti-rumor users	
Weibo content	Rumor information	Content, release time, number of comments, reposts, and likes
	Anti-rumor record	
Confrontation result	Penalties	Level, score deduction, penalty days, account suspended

### 3.2. Characteristics of Rumor and Anti-Rumor Spreaders.

In this dataset, there are 34754 rumor spreaders and 15866 anti-rumor accounts, most of the accounts appeared only once. After data filtering, 2553 rumor spreaders and 2653 anti-rumor accounts who appeared more than once are conserved. The characteristics of the two types of accounts are compared, including basic features, high-frequency words of descriptions, geographic location, number of followers, fans, and posts in Sina platform, and number of posts in False Information Publicity Results, as shown in Tables 2 and 3 and Figures 2–4.

Table 2 shows that, compared with anti-rumors, more rumor spreaders are deleted. This is because in Sina Weibo if an account spreads rumor with a serious consequence, it would be deleted immediately. In Table 2, the proportion of verified anti-rumor accounts is a little more than that of verified rumor spreaders, which is reasonable. However, compared with anti-rumor accounts, rumor spreaders have more comprehensive information, including brief introductions, tags, and educational information.

Table 3 illustrates the high-frequency words of descriptions between the two kinds of accounts. The names of rumor spreaders are mainly cities, while those of anti-rumor accounts are related to police. Besides that, the labels of the two types of users are mainly about entertainment, and this is depended on the nature of the platform.

From the perspective of geographical distribution in Figure 2, rumor spreaders are more dispersed than anti-

rumor accounts. The areas where both types of accounts are concentrated are Beijing, Shanghai, and Guangdong, which are the most developed cities and provinces in China. Besides, the anti-rumor reporters are also distributed in the three coastal provinces; they are Fujian, Zhejiang, and Jiangsu.

Figure 3 gives the distributions of followers, fans, and posts between rumor and anti-rumor accounts. The three subfigures have a common pattern, that is when the numbers of followers, fans, and posts are small, anti-rumor accounts are more than rumor spreaders. But, when the numbers raise, results in the three situations are reversed.

Figure 4 shows the distribution of posts by anti-rumor accounts and rumor spreaders. Overall, the number of spreaders who post just once is more than that of anti-rumor accounts. In addition, a small number of anti-rumor accounts reported many times. Some of them are professionals, for example, the account who reported about 104 times is found to be a professional whistleblower.

### 3.3. Rumors Related to COVID-19

**3.3.1. Number of Reported Rumors.** Rumors related to COVID-19 from January 2 to May 8, 2020, in Sina Weibo are obtained as the topic data source. Totally, there are 526 rumors on the topic of COVID-19 posted by 505 accounts. The number of reported rumors fluctuates every day, as shown in Figure 5.

TABLE 2: Basic features between rumor spreaders and anti-rumor accounts.

	Deleted (%)	Verified (%)	Brief introduction (%)	Tagged (%)	Educational information (%)
Rumor spreaders	37.30	57.62	81.92	70.70	31.65
Anti-rumor users	16.77	58.15	74.41	70.56	31.15

TABLE 3: High-frequency words of descriptions between rumor spreaders and anti-rumor accounts.

	User name		Label	
	Rumor spreaders	Anti-rumor accounts	Rumor spreaders	Anti-rumor accounts
Top 1	上海 (Shanghai)	网警 (internet police)	美食 (food)	美食 (food)
Top 2	北京 (Beijing)	执法 (law enforcement)	生活 (life)	旅游 (tourism)
Top 3	搞笑 (humor)	巡查 (patrol)	旅游 (tourism)	电影 (movie)
Top 4	全球 (global)	公安 (policeman)	搞笑 (humor)	音乐 (music)
Top 5	生活 (life)	中国 (China)	新闻 (news)	IT

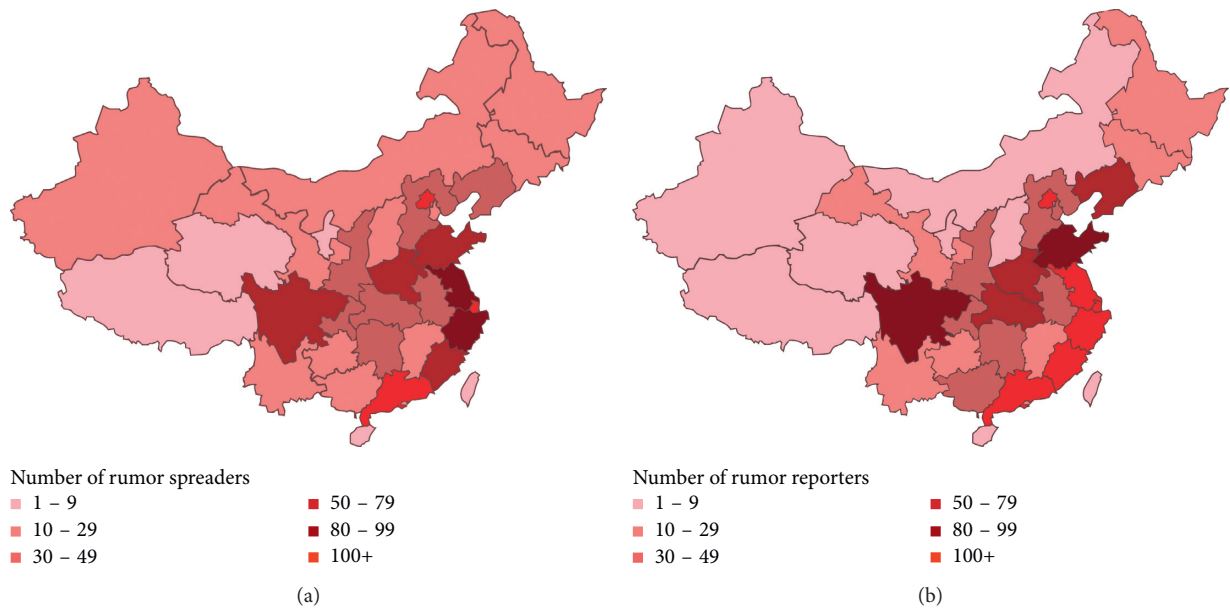


FIGURE 2: Geographical distributions between (a) rumor spreaders and (b) anti-rumor accounts.

**3.3.2. Report and Punishment.** Reported time and frequency distribution of rumors related to COVID-19 are shown in Table 4 and Figure 6, respectively. In Table 4, 81% of the rumors are reported within one day, 10% of them are even marked within 10 minutes, which reflects the rumor supervision of the Sina platform is relatively strict, and the anti-rumor reporters respond quickly to rumors.

Besides, about 60% of the rumors are reported only once by whistleblowers, but there are also some rumors reported more than 100 times by different anti-rumor accounts, as shown in Figure 6. Through further analysis, it is found that the accounts who have been reported many times have comparatively lots of fans. For example, the account whose rumor was reported more than 600 times has 7119 followers.

The levels distribution of punishment in the False Information Publicity Results of Sina Weibo is shown in Figure 7. It illustrates that the number of rumor spreaders with punishment level 1 (the lightest punishment) accounts for nearly 50%. In addition, ranging from level 1 to level 5, as the level increases, the number of people decreases rapidly. Moreover, there are a few rumor spreaders in level 6 and level 7 (heaviest punishment), the two add up to only about 1.5%, and it seems that

most rumors related to COVID-19 in Sina Weibo are not very harmful. As to the harmful rumors, we found that 11.95% of the rumors were deleted, and 3.31% of the accounts could not be displayed on the Sina platform. Furthermore, the correlation coefficients between punishment level and the number of reported times, reposts, comments, and likes are tested. All the correlation coefficients are around 0.2, which illustrates that there is no obvious positive relationship between them. We speculate the punishment level is related to the possible harmfulness of the message.

**3.3.3. Rumor Propagation.** In social networks, there are three common behaviors, reposting, commenting, and giving somebody thumbs up. Distributions of the three behaviors are shown in Figure 8. Besides, these rumors are ranked according to the number of comments; features of the top-10 and the last-10 rumors are compared, as shown in Tables 5 and 6, respectively; in-depth analysis of the comments of the top 4 rumors are given in Figures 9 and 10 and Table 7.

In the three behaviors, giving a “like” is much easier than the other two behaviors, which just needs to press a button

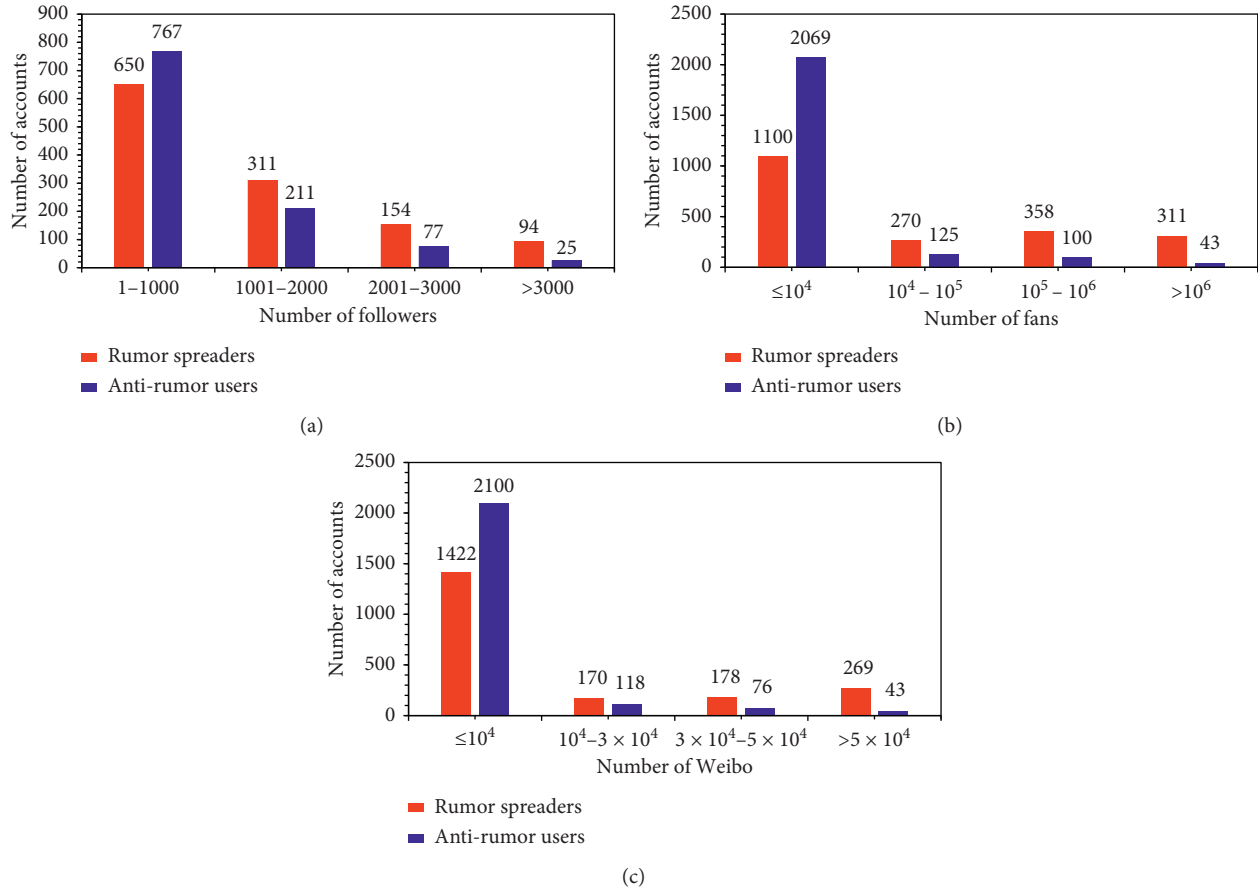


FIGURE 3: Distributions of followers, fans, and posts in Sina Weibo. (a) Distributions of followers. (b) Distributions of fans. (c) Posts in Sina Weibo.

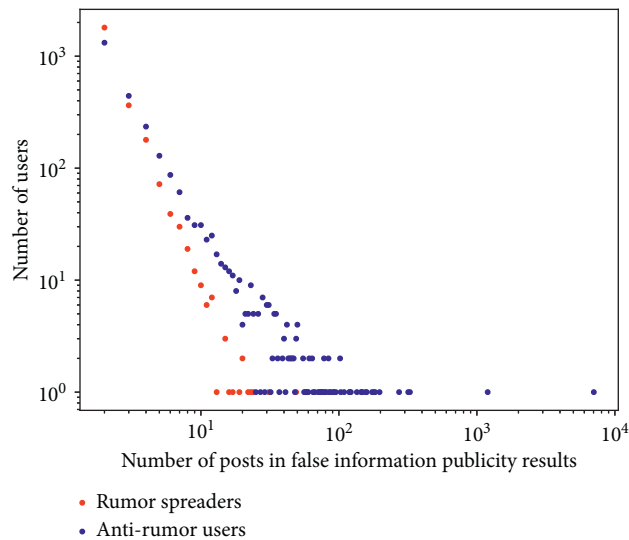


FIGURE 4: Distributions of posts in false information publicity results.

and will not lead to further broadcasting. Reposting is easier than commenting but may need to be responsible for the results of information dissemination. Commenting is the most time-consuming action. So, Figure 8 illustrates that

the most popular action is liking and the next one is reposting. A few of rumors related to COVID-19 even get more than 105 likes, and some rumors are reposted more than 104 times.

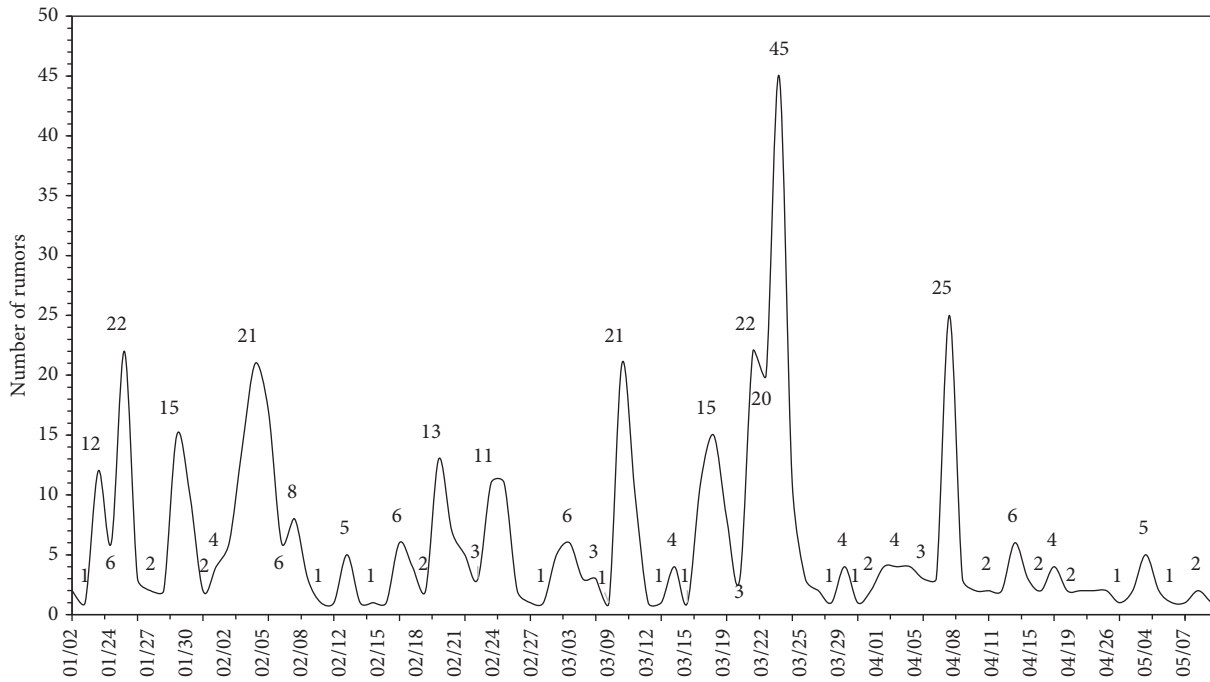


FIGURE 5: Number of reported rumors related to COVID-19.

TABLE 4: Reported time distribution of rumors about COVID-19.

Time	Within 10 min	Within 1 hour	Within 1 day	More than 1 day
Proportion	10.03%	27.02%	81.27%	18.73%

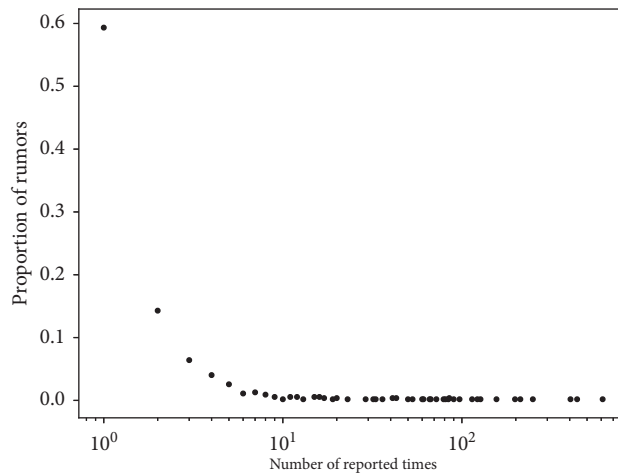


FIGURE 6: Distribution of reported times of different rumors.

There are 10 features in Tables 5 and 6, including the number of followers, fans, posts, and credit level of the rumor spreaders, average reposts, comments, and likes of ten posts written before the rumors, the earliest time and frequency of being reported, and whether the rumor spreader has been reported previously. (1) As a whole, the accounts in the top-10 set have more fans than those who are in the last-10 set, and the number of followers and posts are various on different accounts. (2) The credit

level of accounts in the two sets is at least medium, so it is unreliable to predict the authenticity of an account's information by using the credit level provided by the platform. (3) The accounts in the last-10 sets cannot attract other users' attention, and the average numbers of their reposts, comments, and likes are zero. In contrast, the messages posted by the top-10 accounts often receive much more feedback from others. (4) All the accounts in the top-10 set are reported equal or more than three



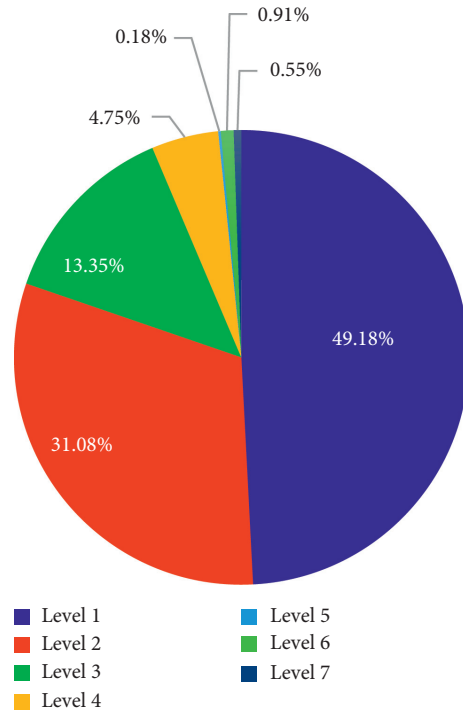


FIGURE 7: The levels distribution of punishment for anti-rumor accounts (level 1: minus 2 points; level 2: minus 5 points, no posting for 7 days; level 3: minus 10 points, no posting for 15 days; level 4: minus 20 points, no posting for 30 days; level 5: minus 20 points, no posting for 90 days; level 6: no posting forever; level 7: delete account).

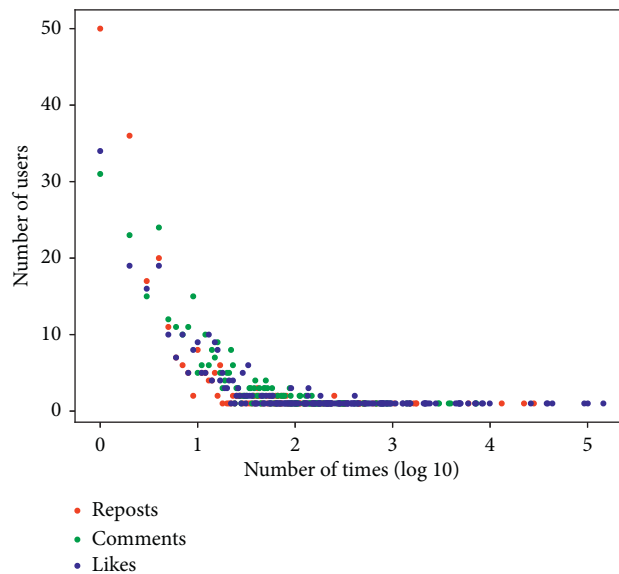


FIGURE 8: Distribution of reposts, comments, and likes of rumors about COVID-19.

times, while all the accounts in the last-10 are reported equal or less than three times. (5) In the cases of the last-10 rumors, some rumors are reported immediately after it appeared on the platform. For example, a piece of rumor whose writer has more than 170 thousand fans was reported one hour after it was written. This illustrates that as long as the supervisors pay close attention to the accounts with big fans, the scope of rumors they posted can

be controlled. (6) Moreover, we found that in the data of the top-10 set, there are three accounts posted rumors beforehand, which reminds us that we need to pay more attention to the accounts who have already posted rumors previously.

Figure 9 illustrates the number of comments on the four most popular rumors related to COVID-19. All the comments of the four rumors arrive at their maximum

TABLE 5: Accounts information in the top-10 set.

Rank	Top 1	Top 2	Top 3	Top 4	Top 5	Top 6	Top 7	Top 8	Top 9
Followers	776	718	680	627	1173	3066	577	190	599
Fans	247675	688008	272292	8422285	26942	11907	43040	56797	10994
Posts	7173	24477	4510	37837	2305	6883	9172	5896	7555
Credit	High	High	Medium	High	Medium	High	High	Medium	Medium
Average reposts	3336	23	1	1451	8	1	2	1	17
Average comments	218	21	4	319	9	3	2	1	10
Average likes	12157	82	16	6427	31	17	10	2	24
Reported times	32	5	611	4	61	440	85	90	16
The earliest report time	23 h 12'	1 d 6 h 34'	23 h 13'	16 h 1'	14 h 14'	8 h 22'	1 h	7 h 18'	2 h 28'
Reported before	N	Y	N	N	N	N	N	Y	N

TABLE 6: Accounts information in the last-10 set.

Rank	Last 1	Last 2	Last 3	Last 4	Last 5	Last 6	Last 7	Last 8	Last 9
Followers	454	4871	179	178	443	508	2445	736	689
Fans	82	1750902	37	82	498	2456	4712	372	1773
Posts	438	42780	92	234	766	1266	4579	1182	17663
Credit	High	Medium	High	High	Medium	Medium	Medium	High	Medium
Average reposts	0	0	0	0	0	0	0	0	0
Average comments	0	1	0	0	1	0	0	0	0
Average likes	1	0	0	1	0	1	1	0	0
Reported times	1	1	1	1	1	1	3	2	1
The earliest report time	3'	1 h 5'	3 h 15'	15 d 37'	13 d 13 h	11 d 23 h	11 h 22'	36'	1 d 30'
Reported before	N	N	N	N	N	N	N	N	N

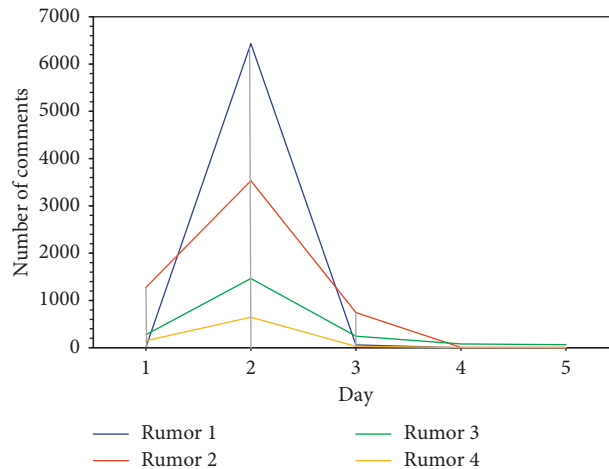


FIGURE 9: Numbers of comments of five popular rumors vary with time.

value the next day after publishing. We checked the date of the rumors and found that except Rumor 2 which was written at 13:44, the other three rumors were written after 21:00. We hypothesized it is the release time that impacts the number of comments on the first day, which reminds us that the evening is a good time to refute rumors. In the meantime, it can be seen that the numbers of comments of the four rumors decrease sharply on the third day, and this is because the Sina platform took measures and marked the rumors, so when a message is confirmed as a fake message, people's attention will be greatly reduced. Figure 10 visualizes the comment networks of the four popular rumors related to COVID-19. There is an obvious core network in Figures 10(a)–10(c), respectively, in which the

center node is the source of the rumor. In Figure 10(d), there are two groups, the center node of one group is the rumor source, and another center node is an account who argued frequently with others. In addition to the core networks, there are some nodes which caused a small number of users to reply. Through text analysis, we found most of these kinds of nodes are accounts who expressed controversial views. Besides, the anti-rumor accounts who appeared in the comment networks are represented as big and red nodes. From Figures 10(a), 10(b), and 10(d), it can be seen that some anti-rumor users do not reply to the rumor source directly, nor do they cause extensive discussion. This may be an important reason for not controlling the spread of rumors in time.

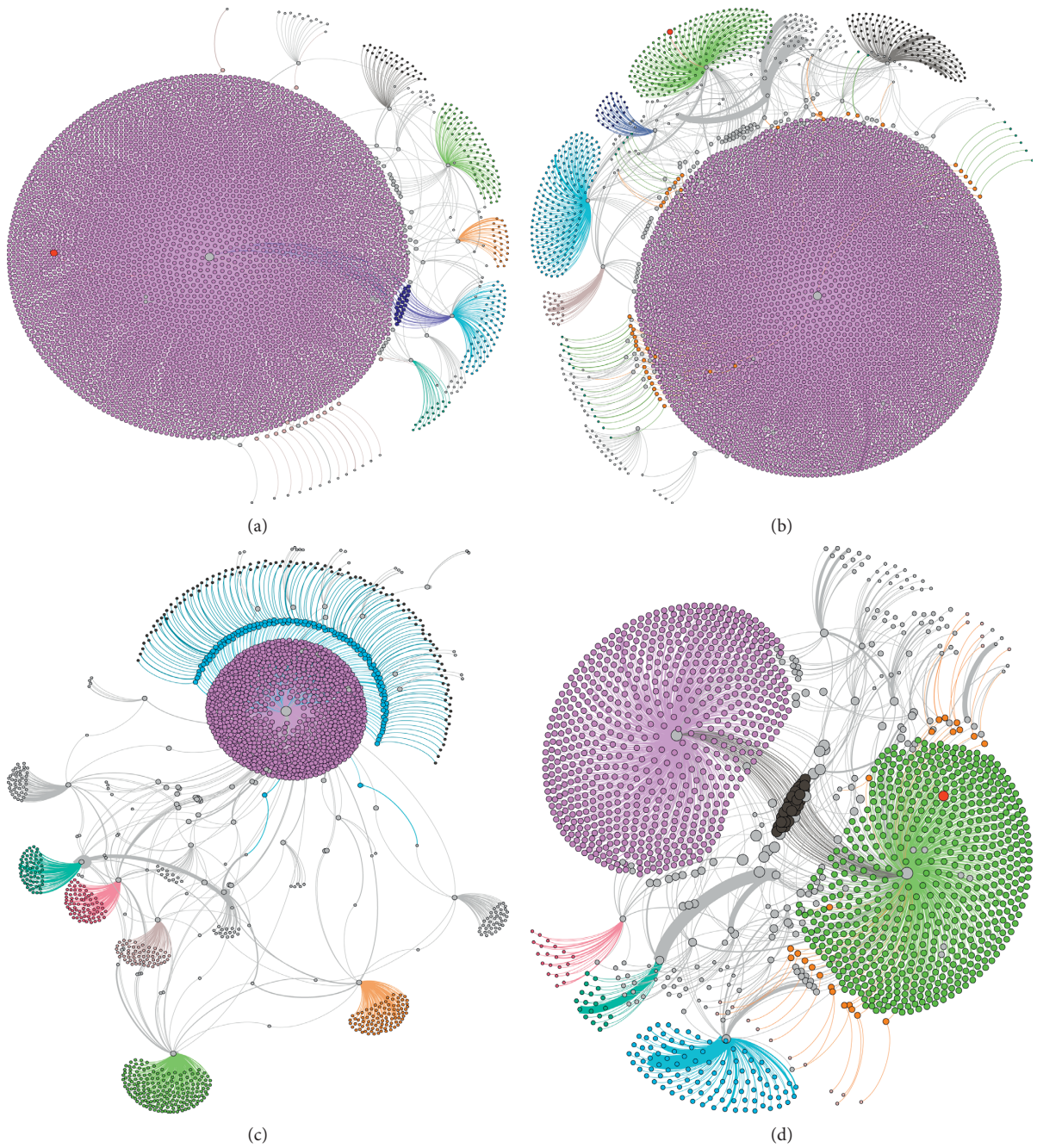


FIGURE 10: Comment networks of the four popular rumors. (a) Rumor 1. (b) Rumor 2. (c) Rumor 3. (d) Rumor 4.

TABLE 7: Network parameters of the comment networks.

Number	Nodes	Edges	Diameter	Clustering coefficient	Average distance	Modularity index
1	5070	5148	4	0.562	2.109	0.137
2	5212	5360	4	0.412	2.284	0.347
3	2029	2145	4	0.137	2.737	0.611
4	1719	1883	5	0.236	3.113	0.659

Several network parameters are shown in Table 7. From Table 7, the four popular rumors about COVID-19 did not spread deeply, because the diameters of the

networks are equal or less than five. Judging from Modularity Index, comment networks of Rumor 3 and Rumor 4 have reached a certain degree of modularity.

## 4. Conclusion

In this work, we performed a case study of dispelling rumors. We compared classical features between the two kinds of accounts and found that the percentage of being deleted, high-frequency words of usernames, geographical distributions of followers, fans, and posts in Sina Weibo and posts in False Information Publicity Results are quite different. The geographical distribution of rumor spreaders is relatively more dispersed, their personal information is more complete, and many of them own a large number of followers, fans, and Weibo. The geographic location of anti-rumor accounts is concentrated, their user names are usually about police, and some of them are very professional that reported rumor over  $10^4$  times. Furthermore, we analyzed the reported rumors of COVID-19. The results suggest that most of the rumors related to COVID-19 do not have serious harm, 60% of the rumors are reported only once, and about 81% of the rumors are reported in one day. As for dispelling rumors, the credit rating provided by Sina is unreliable to detect rumor spreaders. In addition, the time after 9 PM is a prime time to debunk rumors. If the rumor is controlled well in the evening, the next day's outbreak may be avoided. Besides, the accounts who had written rumors previously and whose daily message is popular should be focused on observation. Meanwhile, the anti-rumor accounts should reply to the rumor source directly.

## Data Availability

Readers can access the data of the manuscript titled "Confrontation analysis between rumor spreaders and anti-rumor reporters" from the False Information Publicity Results of Sina Weibo. In the manuscript, the URL of the data source is given in Reference [10], that is, <https://service.account.weibo.com/?type=5&status=4&display=0&retcode=6102>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The work was partially supported by the National Natural Science Foundation of China under Grant no. 81901389 and the China Postdoctoral Science Foundation under Grant no. 2019M653400.

## References

- [1] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.
- [2] T. Chen, X. Li, H. Yin et al., "Call attention to rumors: deep attention based recurrent neural networks for early rumor detection," in *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 40–52, Springer, Melbourne, Australia, June 2018.
- [3] <http://www.dailymail.co.uk/news/article-2313652/AP-Twitter-hackers-break-news-White-House-explosions-injured-Obama.html>.
- [4] S. Tasnim, M. M. Hossain, and H. Mazumder, "Impact of rumors or misinformation on coronavirus disease (COVID-19) in social media," *Journal of Preventive Medicine and Public Health*, vol. 53, pp. 171–174, 2020.
- [5] <https://www.independent.co.uk/news/world/middle-east/coronavirus-iran-deaths-toxic-methanol-alcohol-fake-news-rumours-a9487801.html>.
- [6] <https://www.cnn.com/2020/03/23/africa/chloroquine-trump-nigeria-intl/index.html>.
- [7] <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19>.
- [8] <https://www.mobihealthnews.com/news/doctor-coronavirus-turns-twitter-his-experience.html>.
- [9] <https://techcrunch.com/2020/03/18/twitter-coronavirus-covid-19-misinformation-policy/>.
- [10] <https://service.account.weibo.com/?type=5&status=4>.
- [11] N. F. Johnson, N. Velásquez, N. J. Restrepo et al., "The online competition between pro-and anti-vaccination views," *Nature*, vol. 582, pp. 230–233, 2020.
- [12] N. F. Johnson, R. Leahy, N. J. Restrepo et al., "Hidden resilience and adaptive dynamics of the global online hate ecology," *Nature*, vol. 573, pp. 261–265, 2019.
- [13] S. Hamidian and M. T. Diab, "Rumor detection and classification for Twitter data," 2019, <http://arxiv.org/abs/1912.08926>.
- [14] <https://libguides.stthomas.edu/covid-19>.
- [15] <https://www.newsguardtech.com/coronavirus-misinformation-tracking-center/>.
- [16] <https://mediabiasfactcheck.com/>.
- [17] F. Yang, Y. Liu, X. Yu, and M. Yang, "Automatic detection of rumor on Sina Weibo," in *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, pp. 1–7, Beijing, China, August 2012.
- [18] H. Guo, J. Cao, Y. Zhang, J. Guo, and J. Li, "Rumor detection with hierarchical social attention network," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 943–951, Turin, Italy, October 2018.
- [19] X. Liu, A. Nourbakhsh, Q. Li, R. Fang, and S. Shah, "Real-time rumor debunking on twitter," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pp. 1867–1870, Melbourne, Australia, October 2015.
- [20] H. Han and X. Guo, "Construction on framework of rumor detection and warning system based on web mining technology," in *Proceedings of the 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, pp. 767–771, IEEE, Singapore, June 2018.
- [21] N. Zhang, H. Huang, M. Duarte, and J. Zhang, "Risk analysis for rumor propagation in metropolises based on improved 8-state ICSAR model and dynamic personal activity trajectories," *Physica A: Statistical Mechanics and Its Applications*, vol. 451, pp. 403–419, 2016.
- [22] Y. Zhang, Y. Su, L. Weigang, and H. Liu, "Rumor and authoritative information propagation model considering super spreading in complex social networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 506, pp. 395–411, 2018.
- [23] A. Bodaghi, S. Goliaei, and M. Salehi, "The number of followings as an influential factor in rumor spreading," *Applied Mathematics and Computation*, vol. 357, pp. 167–184, 2019.
- [24] Y. Zhang, Y. Su, W. Li, and H. Liu, "Modeling rumor propagation and refutation with time effect in online social

- networks,” *International Journal of Modern Physics C*, vol. 29, no. 8, Article ID 1850068, 2018.
- [25] M. Askarizadeh, B. Tork Ladani, and M. H. Manshaei, “An evolutionary game model for analysis of rumor propagation and control in social networks,” *Physica A: Statistical Mechanics and Its Applications*, vol. 523, pp. 21–39, 2019.
- [26] T. Takahashi and N. Igata, “Rumor detection on twitter,” in *Proceedings of the 13th International Symposium on Advanced Intelligence Systems*, IEEE, Kobe, Japan, pp. 452–457, November 2012.
- [27] L. Zhiyuan, Z. Yue, T. Cuncao, and S. Maosong, “Statistical semantic analysis of Chinese social media rumors,” *Science of China: Information Science*, vol. 45, no. 12, p. 1536, 2015.
- [28] X. Xiong, F. Xiong, J. Zhao et al., “Assessment of public attention, risk perception, dynamic discovery of favorite locations in spatio-temporal social networks,” *Information Processing and Management*, vol. 57, p. 102337, 2020.
- [29] X. Xiong, S. Qiao, Y. Li, N. Han, F. Xiong, and L. He, “Affective impression: sentiment-awareness POI suggestion via embedding in heterogeneous LBSNs,” *IEEE Transactions on Affective Computing*, p. 1, 2020.
- [30] X. Xiong, Y. Li, R. Zhang, Z. Bu, G. Li, and S. Ju, “DGI: recognition of textual entailment via dynamic gate matching,” *Knowledge-Based Systems*, vol. 194, Article ID 105544, 2020.
- [31] Z. Hou, F. Du, H. Jiang, X. Zhou, and L. Lin, “Assessment of public attention, risk perception, emotional and behavioural responses to the COVID-19 out-break: social media surveillance in China,” *SSRN Electronic Journal*, 2020.
- [32] H. P. Zhang, H. K. Yu, D. Y. Xiong, and Q. Liu, “HHMM-based Chinese lexical analyzer ICTCLAS,” in *Proceedings of the Second SIGHAN Workshop on Chinese Language Processing*, vol. 17, pp. 184–187, Association for Computational Linguistics, Sapporo, Japan, June 2003.
- [33] T. Joachims, *A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization*, Carnegie-Mellon Univ Pittsburgh pa Dept of Computer Science, Pittsburgh, PA, USA, 1996.

## Research Article

# A Hierarchical Attention Recommender System Based on Cross-Domain Social Networks

Rongmei Zhao <sup>1</sup>, Xi Xiong <sup>1,2</sup>, Xia Zu <sup>3</sup>, Shenggen Ju,<sup>4</sup> Zhongzhi Li,<sup>1</sup> and Binyong Li<sup>1</sup>

<sup>1</sup>School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China

<sup>2</sup>School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China

<sup>3</sup>College of Computer Science, Sichuan University, Chengdu 610065, China

<sup>4</sup>School of Management, Chengdu University of Information Technology, Chengdu 610103, China

Correspondence should be addressed to Xia Zu; [sunnyzu@163.com](mailto:sunnyzu@163.com)

Received 14 June 2020; Revised 6 July 2020; Accepted 23 July 2020; Published 12 August 2020

Guest Editor: Hongshu Chen

Copyright © 2020 Rongmei Zhao 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.

Search engines and recommendation systems are an essential means of solving information overload, and recommendation algorithms are the core of recommendation systems. Recently, the recommendation algorithm of graph neural network based on social network has greatly improved the quality of the recommendation system. However, these methods paid far too little attention to the heterogeneity of social networks. Indeed, ignoring the heterogeneity of connections between users and interactions between users and items may seriously affect user representation. In this paper, we propose a hierarchical attention recommendation system (HA-RS) based on mask social network, combining social network information and user behavior information, which improves not only the accuracy of recommendation but also the flexibility of the network. First, learning the node representation in the item domain through the proposed Context-NE model and then the feature information of neighbor nodes in social domain is aggregated through the hierarchical attention network. It can fuse the information in the heterogeneous network (social domain and item domain) through the above two steps. We propose the mask mechanism to solve the cold-start issues for users and items by randomly masking some nodes in the item domain and in the social domain during the training process. Comprehensive experiments on four real-world datasets show the effectiveness of the proposed method.

## 1. Introduction

With the acceleration of people's daily life, much time can be saved to get useful information quickly in a practical way, and recommender systems play a critical role in filtering information. Collaborative filtering recommendation [1] is the mainstream traditional recommendation algorithm. However, due to the data sparsity, it seriously affects the recommendation quality. The collaborative filtering recommendation combined with the neural network [2] (e.g., CNN, RNN, and CDAE) alleviates this problem. Besides, taking advantage of the productive relationship in the social networks [3–5] can effectively solve the cold-start problem [6, 7], but there is a malicious fraud problem by distrusting users.

On one hand, in social networks, there are explicit or implicit relationships between most users, which influence

each other's behaviors. However, not every user and connection in a social network is trustworthy. To reduce the impact of untrusted users, social networks can be divided into many small ego-networks (representing all the users both directly and indirectly trusted by  $u$ ) according to users' trust. Adding a trusted network into the recommender system can effectively solve the problem of a fraud attack. The traditional attribute graph-based semisupervised classification methods, such as label propagation (LP) [8] and label spreading (LS) [9], can effectively classify the unstructured data and have been used in community (consisting of similarity nodes) detection. In [10], they propose a new method to extend the global reputation model with local reputation, through an unsupervised process calculated in the user's ego-network. On the other hand, the widespread existence of social media has dramatically enriched the social

activities of network users and produced productive social relationships; the integration of these social relationships has improved the quality of the recommender system and solved the cold-start problem. In recent years, researchers have proposed a large number of recommender systems [11, 12] based on social networks that homogenize social relationships. However, it may restrain user representation learning in each respective domain, since users behave and interact differently in two domains, which makes their representations heterogeneous [13]. Some research [14, 15] considers the heterogeneity, and learning separated latent representations and then transferred them from a source domain to a target domain for a recommendation. However, learning the representations is challenging due to the inherent data sparsity problem in both domains.

Based on the latest advances in deep learning, especially the development of graph convolutional networks (GCN) [16], it is more accessible to aggregate features in social networks. The recommender system proposed by researchers based on graph convolutional neural network [17] is superior to the previous recommendation algorithms in recommendation quality; Chebyshev [18] is a classical graph convolution based on the spectral method. However, graph convolution assigns the same weight to neighbor nodes, and the feature aggregation depends on the entire graph, which limits the flexibility and generalization of the graph. Therefore, a graph attention network [19] is proposed, which uses the attention mechanism to weigh the sum of the features of neighboring nodes, and the feature weights of neighboring nodes are related to the nodes themselves, whether it is a graph convolutional neural network or a graph attention network, processing the user nodes and item nodes in the same way, without considering the heterogeneity of the two types of nodes. Besides, the training time is longer when the network is more massive.

In this paper, we consider the heterogeneity of information and relationships in the social and item domains, propose the Context-NE model to learn the representation in the item domain. We can learn representations of users interacting with multiple items through the model. The context contains the basic properties of the node and the user's comments on the item so that it can fuse the structure information and text information in the item domain. Besides, based on the graph convolutional networks, we propose a hierarchical attention mechanism and mask mechanism, which can improve the quality of recommendations and the generalization ability and flexibility of the network. Our significant contributions can be summarized as follows:

- (i) We propose a network embedding method Context-NE, which learns the node embedding in the item domain. It mainly represents the information of multiple items visited by the user and obtains the user's node representation in this domain.
- (ii) We use the hierarchical attention mechanism to aggregate the node information in the social domain. At first, we use the K-head self-attention mechanism the same as GAT to aggregate the

information of neighbor nodes; then, we use the attention mechanism to aggregate the output information of the K-head self-attention mechanism to obtain the final representation of the node.

- (iii) We apply the mask mechanism during training and processing the features of a mask node according to the mask mechanism. Newly joined nodes (e.g., new users and new items) can be regarded as a masked node, increasing the flexibility and generalization of the network.
- (iv) We conducted extensive experiments to evaluate the performance of the proposed HA-RS model on citation networks and Yelp datasets, and the results showed that our approach performed better than the most advanced baseline approaches.

The remainder of this paper is structured as follows: Section 2 discusses the related works. Section 3 gives preliminary definitions in the paper and gives a description of the problem. Section 4 introduces the method of hierarchical attention recommender system based on cross-domain social network, which is divided into four parts. The experimental results are demonstrated in Section 5. Lastly, the full study is concluded in Section 6.

## 2. Related Works

In this section, we will introduce the study of the traditional recommender system, trust recommender system, and graph convolution recommender system, respectively. This paper builds on the latest advanced methods based on social networks.

*2.1. Traditional Recommender Systems.* It is mainly divided into two categories. One is content-based recommender systems [20] which use content information of users and items, such as their respective occupation and genre, to predict the next purchase of a user or rating of an item. The other is collaborative filtering models [21] which find the choices of users by mining the historical behavior data of users, divide the groups of users based on different preferences, and recommend commodities with similar interests. The traditional the recommendation algorithm has serious cold-start problems.

*2.2. Trust Recommender System.* Social network-based recommender systems [22] can mitigate the cold-start problem, and most of the information on social networks is unreliable due to different backgrounds or preferences of users. The trust network is introduced into the recommender system to solve the problem of a fraud attack. Yang et al. [23] propose a hybrid method (TrustMF) that combines both a truster model and a trustee model from the perspectives of trusters and trustees; that is, both the users who trust the active user and those who are trusted by the user will influence the user's ratings on unknown items. Guo et al. [24] propose TrustSVD, a trust-based matrix factorization technique, which takes into account both the explicit and implicit

influence of ratings and trust information when predicting ratings of unknown items. Pasquale et al. [10] propose a model capable of integrating local and global reputation in online social networks and calculating the user’s ego-network through an unsupervised method. In this paper, it claims that the local reputations generally perform better than global ones, although they could not always be sufficient for calculating reliable values.

**2.3. Graph Convolution Recommender System.** Topological relations [25] and information transmission [26] between nodes in social networks are very important for social recommendation. Therefore, fully extracting the information of network structure is one of the key parts of recommendation algorithm. Based on the graph embedding method and drawing on the ideas of CNN, a graph neural network (GNN) [27] is proposed to collect information from the graph structure. GNN is a compelling, structured data modeling architecture; however, GNN uses the same parameters in the iteration and cannot aggregate edge information. To solve some of the limitations of graph neural networks, some variants of graph neural networks (e.g., GCN, DCNN, GAT, and GGNN) have gradually appeared. Instead of directly propagating all the attributes of each node in GCN and GAT, Masked GCN [28] only propagates a portion of its attributes to the neighbors. PinSage [29] combines efficient random walks and graph convolutions to generate embeddings of nodes (i.e., items) that incorporate both graph structure and node feature information. STAR-GCN [30] can produce node embeddings for new nodes by reconstructing masked input node embeddings, which primarily tackles the cold-start problem.

Despite the compelling success achieved by these methods in general cases, they ignore the heterogeneity of social networks. It may restrain user representation learning in each respective domain since users behave and interact differently in social networks. GraphRec [31] and DiffNet [32] consider the heterogeneity of social network, but they separately learn a latent representation of users and items. Poor network flexibility and scalability is a common problem in graph-based recommendation algorithms. The main concern of this paper is to consider the heterogeneity of social networks and improve the flexibility of networks.

### 3. Definitions and Problem Statement

The social network used in this paper is a heterogeneous graph, which contains multiple types of nodes (users and item nodes), as well as various types of edges (e.g., the relationship between users and the interaction between users and items).

*Definition 1.* Social domain and item domain: we divide the social network into the social domain and the item domain, which represent the user-user connections and user-item interactions.  $I = \{I_1, I_2, \dots, I_m\}$  represents the item sets and  $m$  indicates that there are  $m$  items in the item domain. The user sets are represented by  $U = \{U_1, U_2, \dots, U_n\}$  and  $n$

represents the total of users in the social domain. The edges between nodes represent friend relationships are represented by  $F$ ; the edges between the user and the item represent their interaction and are represented by  $T$ . The entire network can be represented through a quaternary  $G = \{U, B, F, T\}$  and the overall network is shown in Figure 1.

*Definition 2.* User embedding and item embedding: item embedding contains information such as the user’s reviews of the item, the score, and the attributes of the item (e.g., item name and type). The initial user representation includes the primary attributes of the user (e.g., name and age).

*Definition 3.* Social homogeneity and item homogeneity: take the user  $U_1$  as an example, the user directly connected to the user is a first-order neighbor (e.g.,  $U_2$  and  $U_3$ ), the user directly connected to the first-order neighbor is a second-order neighbor, and so on. User  $U_1$  tends to choose items  $I_1$  and  $I_2$  that first-order neighbors and second-order neighbors prefer (i.e., users and friends have similar preferences), which is called social homogeneity. User  $U_1$  also tends to choose item  $I_6$  which is similar to item  $I_5$  selected by the user’s historical behavior, which is called item homogeneity. We will further introduce item homogeneity and social homogeneity in Sections 4.1 and 4.2, respectively.

*Definition 4.* Network mask: we randomly selected some nodes in network  $G$  and masked some attributes of these nodes. We describe the mask mechanism in detail in Section 4.3; the users ( $U_3$  and  $U_4$ ) and the items ( $I_1$  and  $I_6$ ) are the masked nodes in the network  $G$ .

**3.1. Task Description.** We divided the heterogeneous network into the social domain and item domain. Firstly, we learn the user representation with user preference in the item domain and then aggregate the social influence in the user domain. Finally, we use the user representation to recommend items.

## 4. Model

The HA-RS model framework is shown in Figure 2 and mainly includes two parts, and one is the user’s node representation model Context-NE in the item domain; it can be seen on the left side of Figure 2. The other is the hierarchical attention embedding model, which is used to learn the feature of user friends; it can be seen on the right side of Figure 2. The following content will describe the model in detail.

**4.1. Context-NE.** Considering the item homogeneity, users tend to choose the item similar to the historical item; that is, the user’s past behavior will affect the user’s next behavior. We obtain the user representation in item domain through the Context-NE model, which contains the historical preference of the user. The Context-NE



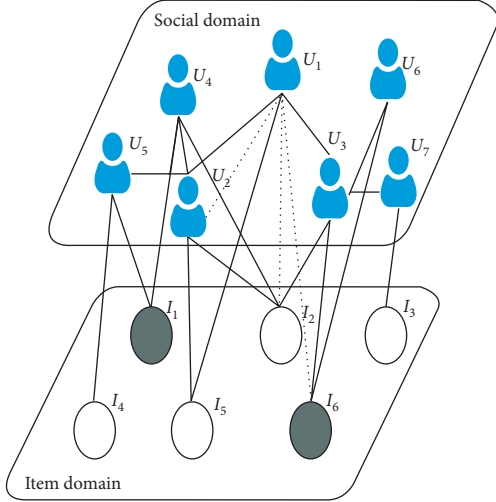


FIGURE 1: Hierarchical representation of heterogeneous social networks. The users ( $U_3$  and  $U_4$ ) and the items ( $I_1$  and  $I_6$ ) are the masked nodes in the network  $G$ .

model learns the user embedding in the item domain, and we call the information of users/items (refers to the point of interest) as a context in this paper. The user's context includes the primary attributes of the user,  $\text{context} = \{\text{user'id, user'name, friend'name, fans'number}\}$ . The context of the item includes the user's review of the item and the primary attributes of the item,  $\text{context} = \{\text{review, bussiness'id, bussiness'name, city, latitude, longitude, categories}\}$ . Considering the interactive sparse of the user in the item domain, we use LSTM to learn context sequence vector other than the score matrix. The Context Description transforms each word in the sequence into the corresponding word vector  $w$  through the LSTM model. It then merges context information (in particular, the context-dependence of the review) into the sequence embedding  $S$ ,  $S = \{w_1, w_2, w_d\}$ ,  $S \in R^d$ . The Convolution Layer extracts local features from the sequence embedding  $S$ , the number of convolution kernels is  $M$ , the size is  $C$ , and the size of the sliding window is  $L_w$ . The user context vector and item context vector after convolution are expressed as  $\vec{U}_c$  and  $\vec{I}_c$ , respectively. Calculation formula is as shown in the following equations:

$$\vec{U}_c = \text{Conv1}(K_m, S^u) + b_m, \quad (1)$$

$$\vec{I}_c = \text{Conv1}(K_m, S^i) + b_m, \quad (2)$$

where  $K_m$  is the  $m$ -th convolution kernel,  $b_m$  is bias,  $S^u$  represents the sequence embedding of the user, and  $S^i$  represents the sequence embedding of the item.

The pooling layer uses mean pooling to reduce the output dimension and prevent overfitting. The sliding window size is also  $L_w$ . The user context vector and item context vector after pooling are expressed as  $\vec{U}_i$  and  $\vec{I}_{ij}$ , respectively, and the calculation formula is as shown in the following equations:

$$\vec{U}_i = \text{meanpool}(\vec{U}_c), \quad (3)$$

$$\vec{I}_{ij} = \text{meanpool}(\vec{I}_c). \quad (4)$$

We get the user representation of the  $i$ -th user in the item domain after the user embedding and the item embedding are nonlinearized by the Tanh function, which is represented by  $\vec{Z}_i$ ; the calculation formula is as shown in the following equation:

$$\vec{Z}_i = r_j \text{Tanh}(\vec{U}_i^T W_{ij} \vec{I}_{ij} + b_{ij}), \quad (5)$$

where  $\vec{I}_{ij}$  represents the embedding of item  $j$  that user  $i$  has visited,  $\vec{U}_i$  represents the representation of user  $i$ , and  $W_{ij}$  and  $b_{ij}$  represent weight parameters and bias parameters, respectively.  $r_j$  represents the user's preference for each item,  $\text{check}_j$  represents the number of interactions between user  $i$  and item  $j$ , and  $\text{check}_i$  represents the total number of interactions between user  $i$  and items. The more times user  $i$  interacts with the same item, the more the user likes the item; that is, the larger the  $r_j$  value, the greater the proportion of the item embedded in the user representation.  $r_j$  can be calculated by the following equation:

$$r_j = \frac{\text{check}_j}{\text{check}_i}. \quad (6)$$

**4.2. Hierarchical Attention.** In the social domain, hierarchical attention mechanism and graph convolution are mainly used to fuse the feature information of user neighbors into user embedding. The user representation finally obtained contain the user's preference information and the social impact in social networks, and utilizing it to recommend will be closer to the user's preferences.

Learning the user representation in the social domain should consider the friend relationships, that is, neighbor nodes in the network graph. However, it is not worth calculating all neighbor nodes in large social networks. Every user has too many neighbors, but not every neighbor has a significant impact on the user's preferences. For example, user  $A$  has a neighbor  $C$  that infrequently interacts; the impact of  $C$  on user  $A$  will be minimal. Besides, we will consider  $L$ -layer neighbor in the aggregating process, if we aggregate the features of all the neighbor nodes that will increase the noise in the user's representation. Meanwhile, the time cost of calculating all  $L$ -layer neighbor nodes in a large-scale network is very high.

We prune the neighbor node in this paper. First, we calculate the node similarity between the user and the neighbor and, then, sort and select the node of the top  $T$  as the user's neighbor when the number of neighbors of the user is higher than  $T$ ; otherwise, there is no need to delete the neighbor, and the pruning strategy can be expressed as the following equation:

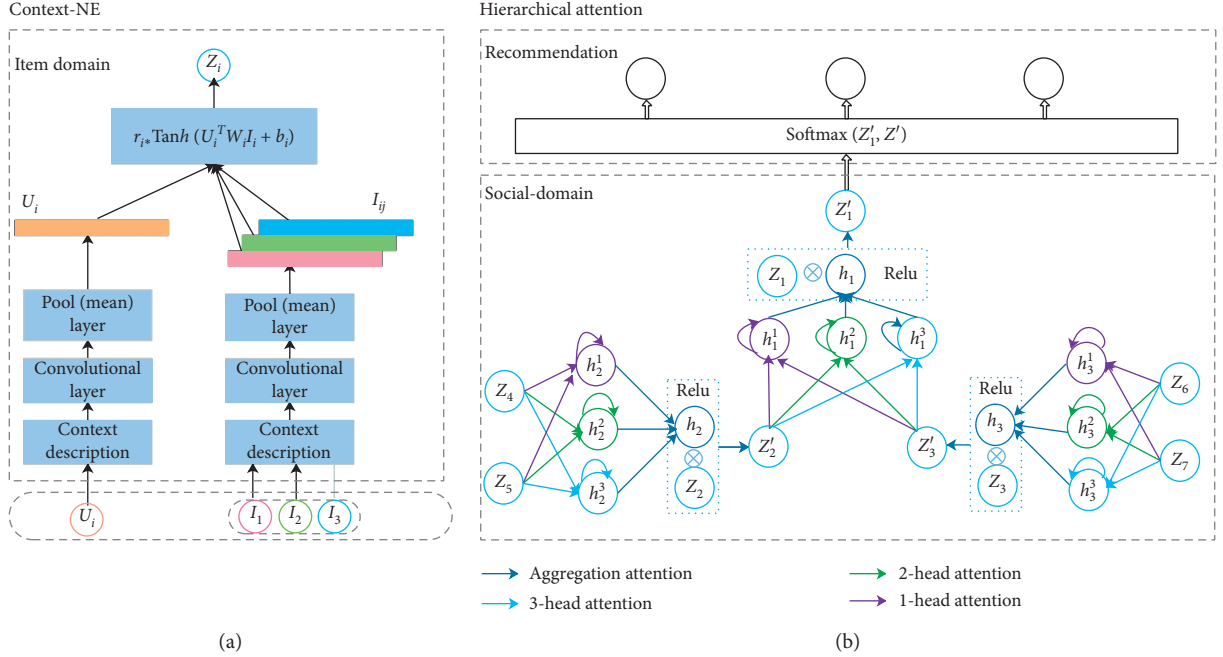


FIGURE 2: The framework of HA-RS model in this paper. (a) The user's node representation model Context-NE in the item domain. (b) The hierarchical attention embedding model, which is used to learn the feature of user friends.

$$N(u) = \begin{cases} N(u), & \text{if } 0 < |N(u)| < T, \\ \text{top}_T[f(u, v)], & \text{else,} \end{cases} \quad (7)$$

where  $\text{top}[\cdot]$  means take  $T$  users ( $v$ ) with the highest similarity to user  $u$ ,  $f(u, v) = u^T v$  is the similarity calculation function between two nodes, and  $N(u)$  represents the neighbors of node  $u$ . Only considering the more critical neighbor nodes can reduce the computation and improve the quality of the user representation, which can be applied to large-scale networks.

In this paper, we assume that users' preferences are related to their  $L$ -order neighbors and use the graph convolution network to aggregate the features of the user's neighbor. Specifically, we assign different weights to the user's neighbors through the hierarchical attention mechanism to indicate the influence degree of different neighbors on users. This process can be likened to the calculation of local reputation in ego-network, where users'  $L$ -order neighbors are user friends in ego-network and the function of attention coefficient is similar to the local reputation. The input of graph convolution network is the embedded set  $Z$  learned in the item domain,  $Z = \{Z_1, Z_2, \dots, Z_N\}$ ,  $Z_i \in R^F$ ,  $N$  represents the number of user nodes, and  $F$  represents the number of features in each node. The output through the layer attention mechanism is represented as  $Z'$ ,  $Z' = \{Z'_1, Z'_2, \dots, Z'_N\}$ ,  $Z'_i \in R^{F'}$ , which can be calculated by the following equation:

$$Z'_u = \text{ReLU}\left(W \cdot \text{Concat}\left(Z_u, \vec{h}_u\right) + b\right), \quad (8)$$

where  $\vec{h}_u$  is the output vector of the aggregate attention layer, which aggregates the output of  $K$  heads of attention in the upper layer, which can be calculated by the following equation:

$$\vec{h}_u = \sigma\left(\sum_{k=1}^K \alpha_h W_h \vec{h}_u^k\right), \quad (9)$$

where  $\alpha_h$  is the attention coefficient of the aggregation layer,  $W_h$  is the weight vector of the aggregation layer, and the output of the  $k$ -head attention is aggregated through the sigmoid function to obtain the output of node  $u$ , the calculation method of  $\alpha_h$  as shown in the following equation:

$$e_h^k = \gamma\left(\beta^T \text{Concat}\left[W_h z_u, W_h \vec{h}_u^k\right]\right) + b_h, \quad (10)$$

$$\alpha_h = \text{softmax}_k(e_h^k) = \frac{\exp(e_h^k)}{\sum_{j=1}^K \exp(e_h^j)},$$

where  $e_h^k$  represents the influence of  $k$ th head attentional output  $\vec{h}_u^k$  on the node when calculating the aggregation expression of node  $u$ .  $\vec{h}_u^k$  is the output of node  $u$  through the  $k$ th head attention, and it can be calculated by (11). Vectors  $z_u$  and  $\vec{h}_u^k$  are multiplied by the Shared weight vector  $W_h$  and then concatenated as the input of the single-layer feedforward neural network;  $b_h$  is the bias.  $\beta$  is the weight vector of the neural network layer, and  $\gamma$  is the LeakyReLU nonlinear function with an input gradient of 0.2. In order to make the coefficients between different neighbor nodes easy to compare, the attention coefficient  $\alpha_h$  of each neighbor node is obtained by using the softmax function:

$$\vec{h}_u^k = \sigma\left(\alpha_{uv}^k W Z_v \mid v \in N(u)\right), \quad (11)$$

where  $\alpha_{uv}^k$  represents the attention coefficient of the first layer,  $W$  is the weight vector,  $Z_u$  and  $Z_v$  are the representation of user learned from item domain, and the calculation method of  $\alpha_{uv}^k$  is as shown in the following equation:

$$e_{uv}^k = \gamma(\beta^T \text{Concat}[W_k z_u, W_k z_v]) + b_k, \quad (12)$$

$$\alpha_{uv}^k = \text{softmax}_v(e_{uv}^k) = \frac{\exp(e_{uv}^k)}{\sum_{j \in N(u)} \exp(e_{uj}^k)},$$

where  $e_{uv}^k$  represents the influence of neighbor  $v$  on the node when calculating the expression of node  $u$ , the calculation process is consistent with  $e_h^k$ ,  $W_k$  is the shared weight vector, and  $b_k$  is the bias. We get the attention vector  $\alpha_{uv}^k$  by normalizing through sigmoid.

**4.3. Mask Mechanism.** The size of the network determines the percentage of the mask; the more extensive the network, the larger the proportion of the nodes in the mask, and there are three ways to calculate the mask node:

- (i) The probability of  $1 - \theta$  is randomly sampling the feature through mask vector from the neighbor node to represent the corresponding feature of the mask node.
- (ii) The probability of  $1 - \theta$  is randomly sampling the feature through mask vector from the neighbor node to represent the corresponding feature of the mask node. The probability of  $\theta/2$  represents the corresponding features of the masked node by randomly sampling the features from other nodes in the network.
- (iii) The probability of  $1 - \theta$  is randomly sampling the feature through mask vector from the neighbor node to represent the corresponding feature of the mask node. The probability of  $\theta/2$  does not deal with the node features of the masked node.

$Z_{imask}$  stands for mask node in item domain and  $Z_{umask}$  stands for mask node in social domain, and the calculation formula is as shown in the following equations:

$$Z_{imask} = \begin{cases} g(Z_{imask}, Z_{irandom}), & \pi = 1 - \theta, \\ Z_{random}, & \pi = \theta/2, \\ Z, & \pi = \theta/2, \end{cases} \quad (13)$$

$$Z_{umask} = \begin{cases} g(Z_{umask}, Z_{urandom}), & \pi = 1 - \theta, \\ Z_{random}, & \pi = \theta/2, \\ Z, & \pi = \theta/2, \end{cases} \quad (14)$$

$g(Z_{imask}, Z_{irandom})$  indicates that, by randomly selecting a neighbor of masked node in the item domain and transmitting the features of neighbor to masked node by mask vector, the mask vector is learned in the training process;  $Z_{random}$  indicates randomly selecting a node in the item domain and propagating the features of the node to masked node;  $Z$  indicates keeping the features of the masked node.  $f(z_{umask}, z_{urandom})$  function indicates randomly selecting a

neighbor of masked node in the social domain and transmitting the features of neighbor to masked node;  $z_{random}$  indicates randomly selecting a node in the social domain and propagating the features of the node to masked node; and  $z$  represents the original embedding of the masked node.  $\pi$  represents the strategy for processing mask nodes.

**4.4. Complexity Analysis.** The time cost of the HA-RS model is composed of two parts. One is to calculate the users' representation in the item domain. Each user needs to be calculated once, the time complexity is  $O(N)$ , and  $N$  is the number of users. The second is aggregating the features of neighbor nodes in the social domain. The time complexity of the aggregation process is  $O(NLT)$ , where  $N$  is the number of users,  $L$  is the maximum neighbor order of the aggregation, and  $T$  is the number of neighbors of each user. After the reduction strategy, the number of neighbors of each user is less than or equal to  $T$ . Since  $T$  is much less than  $N$ , and  $M$  and  $L$  have small values, the time complexity of the HA-RS model is less than  $O(N^2)$ .

## 5. Experiments

In this section, we conduct experiments to evaluate the performance of HA-RS on four datasets. Specifically, we aim to answer the following two research questions: how does HA-RS perform compared with the state-of-the-art social recommendation. How do different components (i.e., the Context-NE model, the attention model, and the mask mechanism) affect HA-RS?

### 5.1. Experimental Settings

**5.1.1. Datasets.** In this section, we conduct experiments that are including transductive learning, inductive learning, ablation experiment, and parameter selection to evaluate the performance of HA-RS on four datasets. Specifically, we aim to answer the following two research questions: how does HA-RS perform compared with the state-of-the-art social recommendation. How do different components (i.e., the Context-NE model, the attention model, and the mask mechanism) affect HA-RS?

### 5.2. Experimental Settings

**5.2.1. Datasets.** We use the citation network datasets for transductive learning, which is a collection composed of references and relationships between references. In this paper, three citation networks (Cora, CiteSeer, and PubMed) were used for experiments. The description of the dataset is shown in Table 1. In each network, nodes and edges are papers and undirected citations, respectively. The node content is constructed by extracting the words from the documents.

We use Yelp for inductive learning that is an online location-based social network. It contains business tables, reviews, tips, user information, and check-in tables. The business table lists the restaurant's name, geographic

TABLE 1: The description of the citation network dataset.

Dataset	No. of nodes	No. of edges	No. of classes	No. of features
CiteSeer	3327	4732	6	3703
Cora	2708	5429	7	1433
PubMed	19717	44338	3	500

location, opening hours, cuisine categories, average star rating, and so forth. The review table lists the restaurant’s star rating, evaluation content, evaluation time, and support rate for the evaluation. In this paper, we transform the ratings that are larger than 3 as the liked items by this user and contain 141804 users and 17625 items.

5.2.2. *Baselines.* For the transductive learning task, we compare HA-RS with various state-of-the-art baselines, including three classical recommendation models: multi-layer perceptron (MLP), label propagation (LP) [8], and graph embedding (DeepWalk) [33]. Besides, we also compare our proposed model with four graph convolutional based recommendation models: graph convolution with Chebyshev filters (Chebyshev) [18], graph convolutional network (GCN) [16], graph attention networks (GAT) [19], and Masked Graph Convolutional Network (Masked GCN) [28].

- (i) MLP: It has multiple layers of neurons, which is a deep neural network, which cannot effectively process the graph data.
- (ii) LP: It uses local information from truncated random walks as input, learning a representation of vertices in a network that encodes social relations in a continuous vector space.
- (iii) Chebyshev: It provides strict control over the local support of filters to extract local and stationary features through graph convolutional layers. The method is computationally more efficient by avoiding an explicit use of the Graph Fourier basis.
- (iv) GCN: It uses an efficient layerwise propagation rule to learn hidden layer representations; the rule is based on a first-order approximation of spectral convolutions on graphs. However, it scales linearly in the number of graph edges.
- (v) GAT: It allows for implicitly assigning different importance to different nodes within a neighborhood while dealing with different sized neighborhoods and does not depend on knowing the entire graph structure upfront.
- (vi) Masked GCN: It only propagates a specific part of the node attributes via a mask vector learned for each node.

For the inductive learning task, we compare HA-RS with three classical recommendation models: BPR [34], feature enhanced latent factor model FM [35], and a state-of-the-art social recommendation model TrustSVD [24]. Besides, we also compare our proposed model with four graph

convolutional based recommendation models: GC-MC [36], PinSage [29], GraphRec [31], and DiffNet [32]. Note that GraphRec and DiffNet consider the heterogeneity of social network.

- (i) BPR: It only considered the user-item rating information for recommendation; FM and TrustSVD improve over BPR by leveraging the node features and social network information.
- (ii) GC-MC: It is one of the first few attempts that directly applied the graph convolutions for a recommendation. GC-MC defines a user-item bipartite graph from user-item interaction behavior and formulating matrix completion as a link prediction task on the bipartite graph
- (iii) PinSage: It is designed for similar item recommendations from an extensive recommender system, combining efficient random walks and graph convolutions to generate embeddings of nodes (i.e., items) that incorporate both graph structure and node feature information. However, this method homogenizes the nodes in the graph without taking into account the differences between user and user interactions and between the user and item interactions.
- (iv) GraphRec: It provides a principled approach to jointly capture interactions and opinions (rate) in the user-item graph, and the framework coherently models two graphs and heterogeneous strengths.
- (v) DiffNet: It is designed to simulate how users are influenced by the recursive social diffusion process for social recommendation. This method considers the heterogeneity of social networks, which separately learns user embedding and item embedding, finally summing the user embedding and user pleasing items embedding for prediction.
- (vi) In order to further validate the improvement obtained by HA-RS, we designed three variant models based on HA-RS.
- (vii) HA-attn has removed the hierarchical attention mechanism from the model, HA-mask has removed the mask mechanism from the model, and RS only considers the heterogeneity of the social network, while removing the hierarchical attention mechanism and the mask mechanism.

5.2.3. *Evaluation Metrics.* On the citation network dataset, the average classification accuracy (with standard deviation) was used to measure the model performance, and we reuse the metrics already reported in Masked Graph Convolutional Network for the state-of-the-art techniques.

On the Yelp dataset, as we focus on recommending top-N items for each user, we use two widely adopted ranking based metrics: hit ratio (HR) and normalized discounted cumulative gain (NDCG). For these two indicators, the larger the value, the better the performance. We reuse the metrics already reported in A Neural Influence Diffusion

Model for Social Recommendation for the state-of-the-art techniques.

HR measures the number of items that the user likes in the test data that has been successfully predicted in the top- $N$  ranking list, and its calculation formula is shown in the following equation:

$$\text{HR@}N = \frac{\sum_{u \in N(u)} \text{HR@}N_u}{T}. \quad (15)$$

The denominator is all the test sets, and the numerator is the sum of the number of each user’s top- $N$  recommendation list in test sets.

NDCG considers the hit positions of the items and gives a higher score if the hit items are in the top positions and its calculation formula is shown in the following equation:

$$\text{NDCG} = \frac{1}{\log_2(1 + \text{rank}_{\text{pos}})}, \quad (16)$$

where  $\text{rank}_{\text{pos}}$  indicates the rank of the positive item (successful prediction item).

**5.3. Parameter Setting.** During training, when learning user representation in the item domain, we apply the grid search to set the hyper-parameters  $M$ ,  $C$ , and  $L_w$ . By using the Masked LM model in Bert [37] for reference, this paper sets  $\theta$  to 20%. We aggregate the feature embedding of  $L$ -layer neighbor in the social domain and set layer  $L = 2$ ; the user’s maximum number of neighbors  $T$  is set to 15 based on the grid search results. L2 regularization was used to prevent overfitting. The attenuation coefficient is set to 0.0005 when the dataset was the citation network and 0.001 when the dataset was Yelp. User nodes in Yelp dataset output embedded dimension  $D$  after passing through the Context-NE model. In the first layer of attention and the aggregation layer of attention, the output of each layer is also set to  $D$  dimension. The first attentional mechanism is composed of  $K = 3$  attentional heads, whose output is the input of the second attentional head. We used Adam as the optimization method for all the models that relied on gradient descent-based methods, with an initial learning rate of 0.001.

## 5.4. Experimental Results

**5.4.1. Accuracy of Transductive Learning on Citation Network.** For the transductive learning, the results of classification accuracy on the three datasets are shown in Table 2. It can be seen from Table 2 that the performance of our model is better than other baselines. More specifically, the accuracy of the model in the three datasets of the citation network in this paper is 1.9%, 2%, and 2.1% higher than the GAT model, respectively, indicating that the HA-RS with mask mechanism and hierarchical attention can improve the performance.

Masked GCN propagates partial attributes instead of the entire ones via a mask vector learned for each node, and the Masked GCN significantly improves the performances compared to GCN and GAT. However, according to the

TABLE 2: Summary of results in terms of classification accuracies, for Cora, CiteSeer, and PubMed.

Methods	Cora (%)	CiteSeer (%)	PubMed (%)
MLP	55.1	46.5	71.4
LP	68.0	45.3	63.0
DeepWalk	67.2	43.2	65.3
Chebyshev	81.2	69.8	74.4
GCN	81.5	70.3	79.0
GAT	83.0	72.5	79.0
Masked GCN	84.4	73.8	80.2
HA-RS	84.9	74.5	81.1

paper of the Masked Graph Convolutional Network, the running time of Masked GCN is 1.24 times compared to that of GAT on average, and the extra time is mainly utilized to learn the parameters of masks. HA-RS only learns mask parameters of a few nodes; it can save much time. Besides, our model HA-RS outperforms Masked GCN.

**5.4.2. Experimental Results of Inductive Learning on Yelp.** For the inductive learning, the results in terms of HR and NDCG are shown in Tables 3 and 4. Table 3 shows the HR@10 and NDCG@10 results with varying output dimension size  $D$ , and we have the following observations:

- (i) On both HR and NDCG, the graph convolutional based recommendation models outperform classical recommendation models. Graph convolution methods considering network heterogeneity, such as GraphRec and DiffNet, have better performance than GC-MC and PinSage. As DiffNet is the best baselines, our model HA-RS improves over DiffNet range of 3% to 3.5% on HR as the output dimension size  $D$  increases from 16 to 64 and range of 1.2% to 5.3% on NDCG. It indicates that this paper divides the heterogeneous social network into the item domain and social domain, and the user representation obtained is more suitable for a recommendation.
- (ii) We find that the performance of all models does not increase as the output latent dimension size  $D$  increases from 16 to 64. The BRP and FM models achieve the best performance at  $D = 32$ , and other models achieve optimal performance at  $D = 64$ . It might be concluded that BPR only considered the user-item rating information for the recommendation; a too large dimension increases the noise in the representation, resulting in a decrease in the recommended performance. Although FM leveraging the node features and social network information, the feature extraction ability of the model is too weak.

Table 4 shows the HR@ $N$  and NDCG@ $N$  results with varying top- $N$  recommendation size  $N$ . From the results, we also find similar observations as in Table 3, with our proposed model HA-RS always showing the best performance. Our model HA-RS improves over DiffNet range of 2.3% to 4.0% on HR as the top- $N$  recommendation size  $N$  increases

TABLE 3: HR@10 and NDCG@10 comparisons for different dimension sizes  $D$ .

Models	Yelp					
	HR			NDCG		
	$D=16$	$D=32$	$D=64$	$D=16$	$D=32$	$D=64$
BPR	0.2443	0.2632	0.2617	0.1471	0.1575	0.1550
FM	0.2768	0.2835	0.2825	0.1698	0.1720	0.1717
TrustSVD	0.2853	0.2880	0.2915	0.1704	0.1723	0.1738
GC-MC	0.2875	0.2902	0.2937	0.1657	0.1686	0.1740
PinSage	0.2952	0.2958	0.3065	0.1758	0.1779	0.1868
DiffNet	0.3366	0.3437	0.3477	0.2052	0.2095	0.2121
HA-RS	0.3487	0.3542	0.3596	0.2122	0.2179	0.2235

TABLE 4: HR@ $N$  and NDCG@ $N$  comparisons for different top- $N$  values.

Models	Yelp					
	HR			NDCG		
	$N=5$	$N=10$	$N=15$	$N=5$	$N=10$	$N=15$
BPR	0.1713	0.2632	0.3289	0.1243	0.1575	0.1773
FM	0.1881	0.2835	0.3463	0.1359	0.1720	0.1895
TrustSVD	0.1906	0.2915	0.3693	0.1385	0.1738	0.1983
GC-MC	0.1932	0.2937	0.3652	0.1420	0.1740	0.1922
PinSage	0.2099	0.3065	0.3873	0.1536	0.1868	0.2130
DiffNet	0.2276	0.3477	0.4232	0.1679	0.2121	0.2331
HA-RS	0.2357	0.3557	0.4404	0.1803	0.2185	0.2451

from 5 to 15 and range of 2.9% to 5.1% on NDCG. Based on the experiment results, we could empirically conclude that our proposed HA-RS model outperforms all the baselines under different output dimensions and different recommendation size  $N$ .

### 5.4.3. Ablation Experiment

(1) *Ablation Experiment and Cold-Start Analysis.* We set the ablation experiment in Yelp dataset to verify the effectiveness of the model, and we set the dimension size as 64 and the  $N$  as 10 in top- $N$  recommendations. The experimental results are shown in Figure 3.

According to Figures 3(a) and 3(b), we observe that HA-RS outperforms the others on HR and NDCG. When removing the hierarchical attention mechanism, the recommendation performance in HR and NDCG decreased slightly, 4.8% and 3.8%, respectively, but there is a great decline in performance of HR and NDCG when removing the mask mechanism, 10.2% and 9%, respectively. It indicates that both the reduced hierarchical attention mechanism and the mask mechanism will affect the performance of the model, and removing the mask mechanism has a more significant impact on the model.

Besides, removing both the hierarchical attention mechanism and the mask mechanism, the performance of the RS was the worst; however, it also outperforms PinSage that uses both social network information and node feature information. The main reason is that PinSage processes all

nodes in the full graph in the same way and does not explicitly distinguish between the user (board) and item (pins) nodes which does not consider the heterogeneity of the social network. It demonstrates the superiority of our model by dividing the heterogeneous social graph into two domains.

We verified the cold-start problem on the Cora dataset, set  $N_c$  of the cold-start users as 0, 50, 100, and 150, respectively, and compared the model HA-RS with its three variants. We regard the cold-start users as the masked nodes and learned by mask vector in HA-RS and HA-attn while, in HA-mask and RS, the cold-start user is randomly initialized; the experimental results are shown in Table 5.

For cold-start users, there is very little information available about a recommendation, and the recommendation result obtained by randomly initializing user embedding is often not ideal. In the method based on social networks, the cold-start problem is alleviated by combining the information of friends. In this paper, cold-start users are regarded as masking nodes, and user embedding is learned by the mask vector. From Table 5, we can conclude that the HA-RS and HA-attn (with mask mechanism) outperform HA-mask and RS (without mask mechanism). Therefore, it can say that the mask mechanism can solve the cold-start problem.

(2) *Parameter Selection.* In this section, we mainly explore the impact of different parameters of HA-RS. We study the impact of L-layer neighbor in the aggregating process in the social domain, the effect of a different number of user neighbors, and the effect of different mask proportions.

We vary the layer of neighbor from 1 to 3 and set the number of user neighbors in every layer from 5 to 20; the performance of HR and NDCG is shown in Figure 4. We can observe from Figure 4 that HA-RS has the best performance when the number of user neighbors is 15; the performance of HA-RS has declined dramatically when the number of neighbors increases to 20. Meanwhile, we can find that aggregating 2-layer neighbor in HA-RS outperforms aggregating 1-layer or 3-layer neighbor; this is because 1-layer cannot capture the higher-order relationships between users in the social domain. Nevertheless, 3-layer may bring massive noise to the model. The above observation has proved that aggregating the features of too many neighbors will increase the noise in the user’s representation, which is described in Section 4.2. Other related studies have empirically found similar trends, with the best layer size set as  $L=2$  [32, 38].

Consider that Masked GCN propagates partial attributes instead of the entire ones via a mask vector learned for each node, and we will discuss that the different mask ratios influence the performance of the model. We apply the mask mechanism of this paper to the original GCN and GAT models, respectively, expressed as GCN-mask and GAT-mask, and compare with HA-RS proposed in this paper.

As mentioned above, the mask mechanism plays a vital role in improving the model. Six groups of different mask proportions are set in the experiment. When the proportion is set to 0%, it is the GCN, GAT, and HA-RS without mask mechanism, and the maximum mask proportion is set to 10%. The accuracy of the three methods in the citation network is shown in Figure 5.

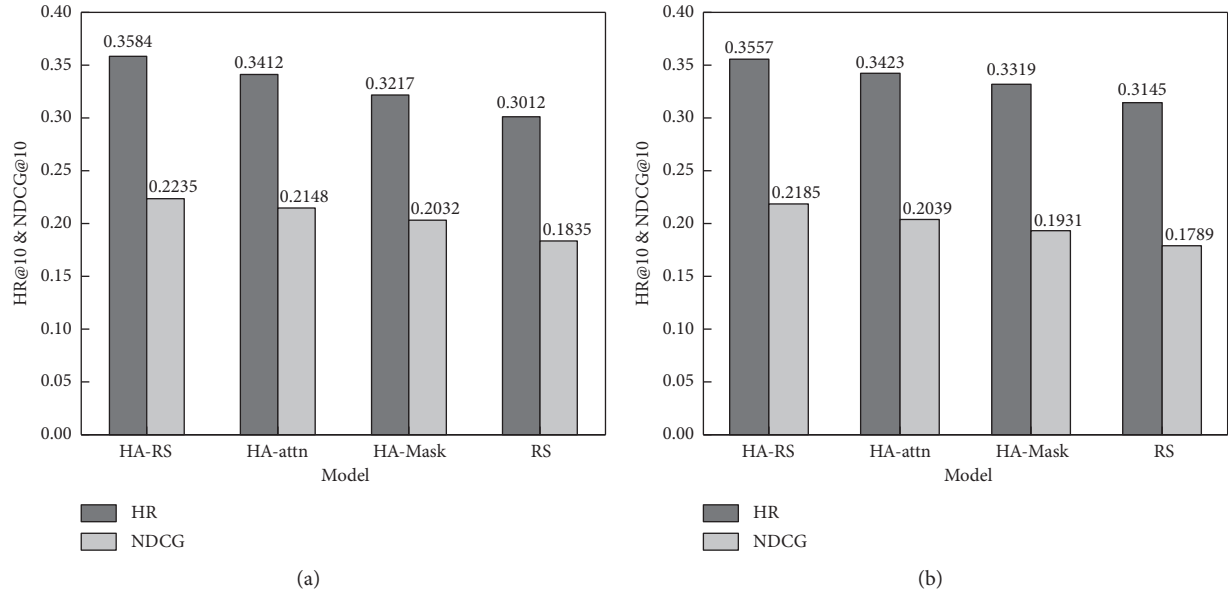


FIGURE 3: The results of the ablation experiment. (a) HR@10 and NDCG@10 comparisons for dimension size = 64. (b) HR@10 and NDCG@10 comparisons for the top-10 values.

TABLE 5: Results of different cold-start users on Cora dataset.

Model	$N_c = 0$ (%)	$N_c = 50$ (%)	$N_c = 100$ (%)	$N_c = 150$ (%)
RS	83.4	83.1	82.6	82.1
HA-attn	84.5	84.3	84.0	84.6
HA-mask	83.7	83.4	83.0	82.5
HA-RS	84.9	84.6	84.3	83.8

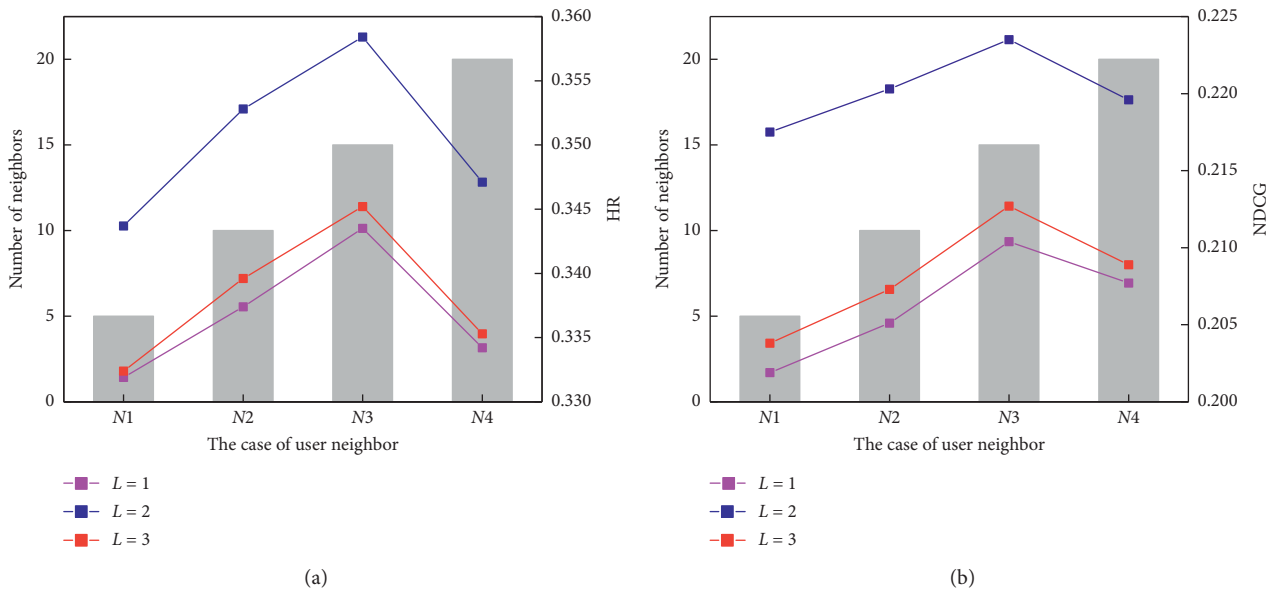


FIGURE 4: The background histograms indicate the four cases of user neighbor; meanwhile, the lines demonstrate the performance of HR/ NDCG with different layers. (a) Performance of HR. (b) Performance of NDCG.

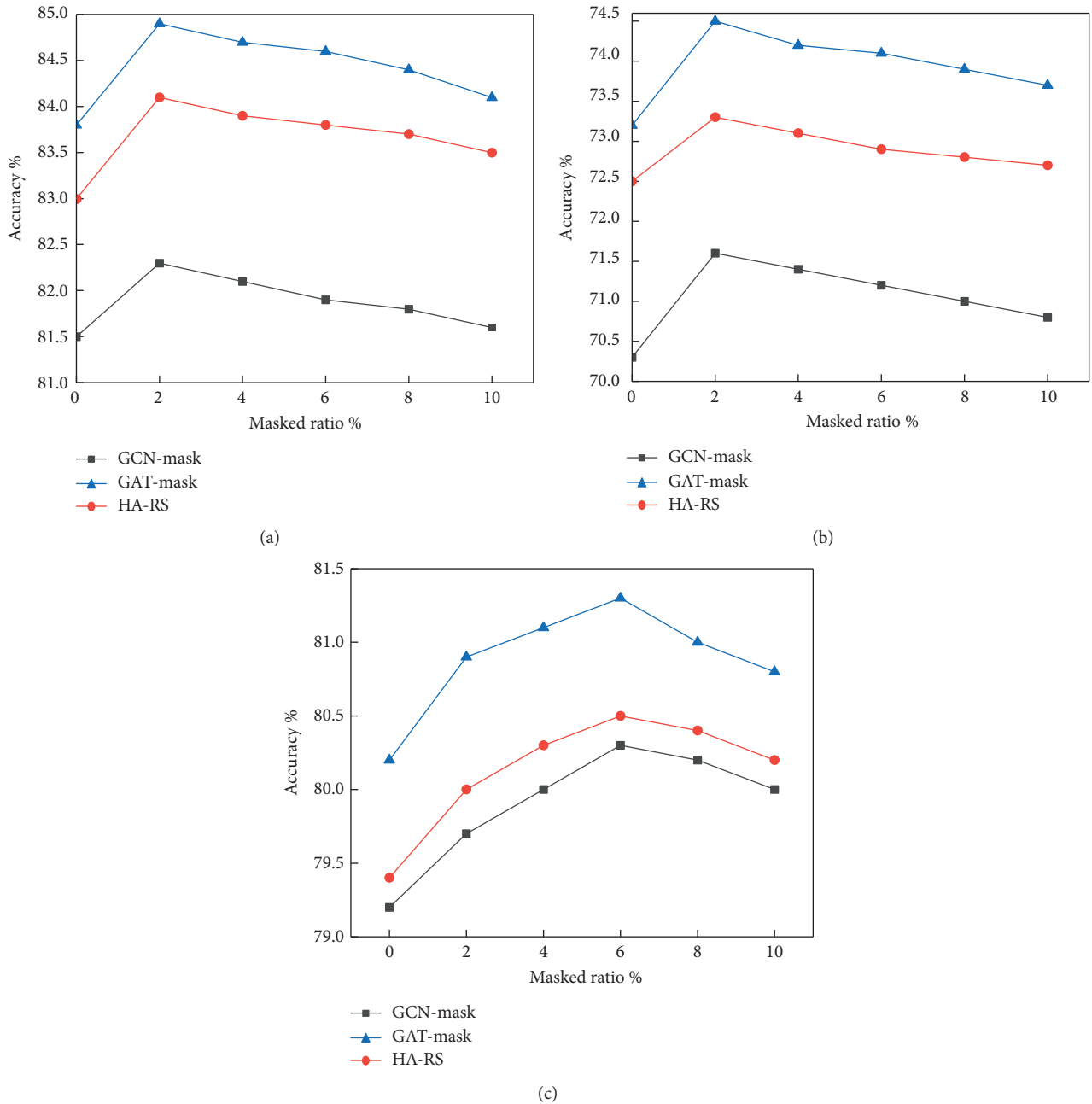


FIGURE 5: Results of different masked ratios on the citation dataset. (a) Cora dataset. (b) CiteSeer dataset. (c) PubMed dataset.

According to Figure 5, we can find that the HA-attn outperforms GCN and GAT when the mask rate is 0%, and it is also proved that dividing the heterogeneous network into the social domain and item domain to learn user representation is better for classification and recommendation. In the dataset of Cora and CiteSeer, the accuracy of the three models reached the highest when the mask proportion was 2%, while the accuracy of the models decreased with the increase of the mask proportion. On PubMed dataset, the mask ratio reached the highest value when it was 6% and then decreased slowly. The main reason is that there are more nodes in PubMed dataset, 6 times more than in Cora and CiteSeer. At a mask ratio of 10% in three datasets, the

three methods perform better than those without a masking mechanism. This shows that the mask mechanism can indeed affect the performance of the model, and the proportion of the mask is related to the network scale. The proportion of the mask that achieves the best performance increases with the increase of the network scale.

## 6. Conclusion

In this work, we explore the problem of extracting more abundant features of user representation for the recommendation in heterogeneous social networks. Considering that the behavior of users in the item domain and the social



domain is inconsistent, we learn that user representation in the HA-RS is divided into two parts. First, we learn the user representation with user behavior in the item domain by the Context-NE model and then through the hierarchical graph attention model to learn user embedding with social impact in the social domain. The mask mechanism in the model also improves the flexibility of the network. As new nodes are added, we can regard it as a masked node. A large number of experiments on real datasets show that our model performs better than the existing techniques above. Future research directions will focus on how to use the geographical location and timing information better to improve the quality of the algorithm further.

## 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

The work was partially supported by the National Natural Science Foundation of China under Grant nos. 81901389 and 71701026; the China Postdoctoral Science Foundation under Grant no. 2019M653400; the Sichuan Science and Technology Program under Grant nos. 2019YFS0236 and 2018GZDZX0039; the Key Project for Soft Science of Meteorology under Grant no. 201704; and the Youth Foundation for Humanities and Social Sciences of Ministry of Education of China under Grant no. 17YJC190035.

## References

- [1] S. Wei, N. Ye, S. Zhang, X. Huang, and J. Zhu, "Item-based collaborative filtering recommendation algorithm combining item category with interestingness measure," in *Proceedings of the 2012 International Conference on Computer Science and Service System*, pp. 2038–2041, IEEE, Nanjing, China, August 2012.
- [2] H. Huo, Z. Wei, L. Liang, and L. Yang, "Collaborative filtering recommendation model based on convolutional denoising auto encoder," in *Proceedings of the 12th Chinese Conference on Computer Supported Cooperative Work and Social Computing*, Kunming, China, August 2017.
- [3] X. Xiong, S. Qiao, Y. Li et al., "ADPDF: a hybrid attribute discrimination method for psychometric data with fuzziness," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 1, pp. 265–278, 2019.
- [4] X. Xiong, C. Xie, R. Zhao, Y. Li, S. Ju, and M. Jin, "A clickthrough rate prediction algorithm based on users' behaviors," *IEEE Access*, vol. 7, pp. 174782–174792, 2019.
- [5] X. Xiong, Y. Li, R. Zhang, Z. Bu, G. Li, and S. Ju, "DGI: recognition of textual entailment via dynamic gate matching," *Knowledge-Based Systems*, vol. 194, p. 105544, 2020.
- [6] X. Xiong, S. Qiao, Y. Li, N. Han, F. Xiong, and L. He, "Affective impression: sentiment-awareness poi suggestion via embedding in heterogeneous lbsns," *IEEE Transactions on Affective Computing*, p. 1, 2020.
- [7] X. Xiong, F. Xiong, J. Zhao, S. Qiao, Y. Li, and Y. Zhao, "Dynamic discovery of favorite locations in spatio-temporal social networks," *Information Processing & Management*, vol. 57, no. 6, Article ID 102337, 2020.
- [8] X. Zhu, Z. Ghahramani, and J. D. Lafferty, "Semi-supervised learning using Gaussian fields and harmonic functions," in *Proceedings of the Machine Learning, Proceedings of the Twentieth International Conference (ICML 2003)*, Washington, DC, USA, August 2003.
- [9] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. S. Olkoff, "Learning with local and global consistency," *Advances in Neural Information Processing Systems*, vol. 16, no. 3, 2004.
- [10] P. De Meo, L. Fotia, F. Messina, D. Rosaci, and G. M. L. Sarné, "Providing recommendations in social networks by integrating local and global reputation," *Information Systems*, vol. 78, pp. 58–67, 2018.
- [11] X. Xiong, S. Qiao, N. Han et al., "Where to go: an effective point-of-interest recommendation framework for heterogeneous social networks," *Neurocomputing*, vol. 373, pp. 56–69, 2020.
- [12] X. Xiong, S. Qiao, Y. Li, N. Han, G. Yuan, and Y. Zhang, "A point-of-interest suggestion algorithm in multi-source geosocial networks," *Engineering Applications of Artificial Intelligence*, vol. 88, Article ID 103374, 2020.
- [13] J. Tang, S. Wang, X. Hu et al., "Recommendation with social dimensions," in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, Phoenix, AZ, USA, February 2016.
- [14] A. M. Elkahky, Y. Song, and X. He, "A multi-view deep learning approach for cross domain user modeling in recommendation systems," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 278–288, Florence, Italy, May 2015.
- [15] C. Wang, M. Niepert, and H. Li, "Recsys-dan: discriminative adversarial networks for cross-domain recommender systems," *IEEE Transactions on Neural Networks and Learning Systems*, 2019.
- [16] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proceedings of the 5th International Conference on Learning Representations, ICLR 2017*, Toulon, France, April 2017.
- [17] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: methods and applications," *IEEE Data Engineering Bulletin*, vol. 40, no. 3, pp. 52–74, 2017.
- [18] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," pp. 3844–3852, 2016, <http://arxiv.org/abs/1606.09375>.
- [19] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proceedings of the 6th International Conference on Learning Representations, ICLR 2018*, Vancouver, BC, Canada, April-May 2018.
- [20] M. J. Pazzani and D. Billsus, "Content-based recommendation systems," in *The Adaptive Web, Methods and Strategies of Web Personalization, Vol. 4321 of Lecture Notes in Computer Science*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds., pp. 325–341, Springer, Berlin, Germany, 2007.
- [21] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," *Communications of The ACM*, vol. 35, no. 12, pp. 61–70, 1992.
- [22] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018.

- [23] B. Yang, Y. Lei, J. Liu, and W. Li, "Social collaborative filtering by trust," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1633–1647, 2017.
- [24] G. Guo, J. Zhang, and N. Yorke-Smith, "Trustsvd: collaborative filtering with both the explicit and implicit influence of user trust and of item ratings," in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, TX, USA, January 2015.
- [25] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 2019.
- [26] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, pp. 1–14, 2019.
- [27] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2008.
- [28] L. Yang, F. Wu, Y. Wang, J. Gu, and Y. Guo, "Masked graph convolutional network," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019*, S. Kraus, Ed., pp. 4070–4077, Macao, China, August 2019.
- [29] R. Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, "Graph convolutional neural networks for web-scale recommender systems," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018*, Y. Guo and F. Farooq, Eds., pp. 974–983, ACM, London, UK, August 2018.
- [30] J. Zhang, X. Shi, S. Zhao, and I. King, "STAR-GCN: stacked and reconstructed graph convolutional networks for recommender systems," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019*, S. Kraus, Ed., pp. 4264–4270pp. 4264–, Macao, China, August 2019.
- [31] W. Fan, Y. Ma, Q. Li et al., "Graph neural networks for social recommendation," in *Proceedings of the 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pp. 417–426, Portland, OR, USA, November 2019.
- [32] L. Wu, P. Sun, Y. Fu, R. Hong, X. Wang, and M. Wang, "A neural influence diffusion model for social recommendation," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 235–244, Paris, France, July 2019.
- [33] 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*, pp. 701–710, New York, NY, USA, August 2014.
- [34] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: bayesian personalized ranking from implicit feedback," 2012, <http://arxiv.org/abs/1205.2618>.
- [35] S. Rendle, "Factorization machines," in *Proceedings of the 2010 IEEE International Conference on Data Mining*, pp. 995–1000, IEEE, Sydney, Australia, December 2010.
- [36] R. van den Berg, T. N. Kipf, and M. Welling, "Graph convolutional matrix completion," 2017, <http://arxiv.org/abs/1706.02263>.
- [37] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," 2018, <http://arxiv.org/abs/1810.04805>.
- [38] L. Hu, C. Li, C. Shi, C. Yang, and C. Shao, "Graph neural news recommendation with long-term and short-term interest modeling," *Information Processing & Management*, vol. 57, no. 2, Article ID 102142, 2020.

## Research Article

# Self-Presentation and Adolescent Altruistic Behaviors in Social Networks

Yaling Zhu,<sup>1</sup> Yue Shen,<sup>2</sup> and Qiang Zhao <sup>2</sup>

<sup>1</sup>International Business School, Shaanxi Normal University, Xi'an 710119, China

<sup>2</sup>School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710064, China

Correspondence should be addressed to Qiang Zhao; 965446620@qq.com

Received 9 June 2020; Revised 10 July 2020; Accepted 15 July 2020; Published 6 August 2020

Guest Editor: Fei Xiong

Copyright © 2020 Yaling 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.

Social networks provide a convenient place for people to interact; members in social networks may create new connections or break existing connections, driving the evolution of complex network structure. Dynamics in social networks, such as opinion formation and spreading dynamics, may result in complex collective phenomena. This paper conducts a survey on 495 students from six schools in Shaanxi, Henan, and Zhejiang provinces and discusses the impact of self-presentation on adolescent network altruistic behaviors, the intermediary role of social ability cognition, and the moderating role of privacy awareness. The results show the following: (1) Self-presentation in social networks can positively predict adolescent network altruistic behaviors. The positive prediction effect of network sharing is the largest, and the positive prediction effect of network support is the least. (2) Social ability cognition plays an intermediary role between self-presentation and adolescent network altruistic behaviors. (3) The moderating effect of privacy awareness is not significant.

## 1. Introduction

According to the China Internet Development Statistical Report, the number of Internet users had reached 904 million by March 2020, with a penetration rate of 64.5%. The Internet has become an important part of modern people's daily life. With the rapid development and popularity of the Internet, negative behaviors such as Internet addiction, cyberbullying, cybercrime, and others have received extensive attention from researchers. However, in addition to these negative behaviors, there are also some positive behaviors such as network altruistic behaviors in network environment [1]. Network altruistic behaviors refer to the conscious and active behaviors of individuals who invade the network, benefiting others or society, and do not expect to be rewarded [2]. Some researchers pointed out that the network environment has the characteristics of anonymity, timeliness, and interactivity [3, 4], which makes altruistic behaviors more likely to occur. Adolescents are a critical stage of life development, and it is of great significance for researchers to study network altruistic behaviors at this stage.

Existing research has shown that altruistic behaviors had positive effects on the academic performance and subjective well-being of adolescent [5]. Considering the importance of network altruistic behaviors in establishing a harmonious network environment and playing active roles in the network [6], it is necessary to study the influencing factors and mechanism of adolescent network altruistic behaviors in order to cultivate adolescent's network altruistic behaviors and promote their positive development.

Self-presentation is a series of actions performed by an individual to present information to others, which will affect the individual's psychology and behavior [7]. From the perspective of objective self-awareness theory, self-presentation in social networks often contains a lot of personal information, which may increase user self-awareness [8]. When self-awareness increases, individuals will become more conscious of social standards and therefore form more pro-social behaviors [9, 10]. At the same time, self-presentation in social network can enable individuals to obtain social support and enhance interpersonal intimacy [11], as well as enhance the individual's sense of social connection

and happiness [12]. From the perspective of reciprocal altruism, the favourable experience brought by self-presentation may cause individuals to generate network altruistic behaviors out of the psychological repayment.

Social ability cognition is a general intermediary mechanism in social psychology changes and an important factor in predicting individual behavior [13]. Social ability cognition refers to an individual's evaluation of his ability to establish and maintain good social relationships [14]. Self-presentation in social networks brings opportunity to interact online with others and can provide individuals with more possibilities to establish and deepen interpersonal relationships. Bandura believed that there were four main sources of social ability cognition: successful experience, alternative experience, verbal persuasion, and emotional state [15]. The online interaction opportunities and positive online feedback that self-presentation brings to individuals can fall into the category of successful experience. Therefore, the self-presentation of social network may have a positive impact on individual's network social ability. At the same time, individuals with high social skills are more willing to connect themselves with others and are more confident in their ability to help others, so they may produce more altruistic behaviors [16]. Empirical studies have shown that online social ability cognition can positively predict network altruistic behaviors [14].

Privacy awareness refers to users' concerns about the collection, control, and excavation of their personal data. Privacy awareness has led users to be more cautious when disclosing personal data [17, 18]. In social network environment, various social activities begin with the provision of personal data, and the continued social activities are accompanied by the contribution of more personal data. Personal data contains private content, and the contribution of personal data leads to user privacy awareness. Obviously, privacy awareness has a negative impact on users' use of social networks. But at the same time, for the purpose of interpersonal relationship management [19], self-presentation [20], and subjective norms [21], despite the disclosure of privacy issues, users voluntarily provide personal data. Thus, it can be said that the user's privacy awareness would affect the user's behaviors in social networks. The moral concept of adolescents has not been finally constructed and has strong plasticity [22], but meanwhile the awareness of privacy protection is not strong. Some adolescents do not pay much attention to their privacy, so their self-presentation on network platforms is mainly true self-presentation, most adolescents hold privacy awareness, and their presence on the network usually hides some private information. It can be said that privacy awareness may affect adolescents' self-presentation in social networks, further changing the adolescents' altruistic behaviors.

Through the above analysis, this paper mainly investigates the influence of adolescent self-presentation on their network altruistic behaviors, the intermediary role of adolescents' social ability cognition, and the moderating role of privacy awareness and attempts to construct the mechanism of adolescent social network self-presentation on their network behaviors. Through the analysis of adolescent self-

presentation, social ability cognition, and privacy awareness, it provides theoretical basis and practical guidance for promoting adolescent's network altruistic behaviors, cultivating their positive psychological quality and forming abundant social networks.

## 2. Theoretical Analysis and Hypothesis

*2.1. Adolescent Self-Presentation and Network Altruistic Behavior.* Social networks provide several platforms for self-presentation. Adolescents have their own personal home-pages on social networking sites, which can be self-presented by posting text, pictures, and videos. Self-presentation has become a major activity in social networks. Since social networks are mainly based on the characteristics of acquaintances [23], individuals will try to present real information [24]. Social network self-presentation, as a special way for individuals to truly reveal themselves to their friends, can enhance interpersonal trust and relationship intimacy, help individuals accumulate social capital, maintain and deepen interpersonal relationships, and obtain social support [19]. Research shows that adolescents who present more personal information on the Internet tend to have more trust in the online environment and are more willing to integrate into the interaction on online platform, which makes it easier for them to identify with the online community and thus more likely to help others online [25]. From the perspective of reciprocal altruism [26], the good experience that self-presentation brings to an individual may cause the individual to generate network altruistic behaviors out of the psychology of repayment. Therefore, we propose Hypothesis 1.

*Hypothesis 1.* Self-presentation in social networks can positively promote adolescent network altruistic behaviors.

According to the form of network altruistic behavior, we draw on the meaning of network altruistic behaviors by scholars such as Wright Li [6] and divide adolescent network altruistic behaviors into network support, network guidance, network sharing, and network warning.

Network support refers to the behaviors such as sharing personal life experiences and perceptions in social networks. It not only contains the behaviors that individuals use their own experience to support others to overcome difficulties, but also includes behaviors that provide network technical support, and these behaviors are often achieved through self-presentation in networks. Adolescents present their experiences and offer support in networks; others can obtain these pieces of information by downloading or contacting adolescents; therefore, adolescents' network support behaviors are strengthened. Thus, we propose H2.

*Hypothesis 2.* Self-presentation in social networks can positively promote adolescents' network support behaviors.

Network guidance is behaviors such as providing guidelines for building groups or running computers. It focuses more on programs. Adolescents' self-presentation includes their experiences, ability, and professional knowledge; these knowledge presentations offer opportunity

for network members to get guidelines of running computer or building group, thus promoting network guidance behavior. Therefore, we propose H3.

*Hypothesis 3.* Self-presentation in social networks can positively promote adolescents' network guidance behaviors.

Researchers have found that the degree of self-presentation in the online social Q&A community can positively predict their online knowledge sharing behavior. Jin et al. pointed out online knowledge sharing behavior was a form of online altruistic behavior and self-presentation was a way of sharing; this meant the contents of self-presentation helped individuals to share their experiences and further facilitated altruistic behavior. But they only targeted specific online communities and did not delve into the internal mechanisms of the impact of self-presentation on other online altruistic behaviors [25]. Based on previous research, we propose H4.

*Hypothesis 4.* Self-presentation in social networks can positively promote adolescents' network sharing behaviors.

Network warning, such as exposing some adverse events in social networks to remind others to pay attention, is attracting adolescents' attention. Due to system vulnerabilities, corruption, and other reasons, some people could not get fairness and they chose silence for fear of retaliation. With the popularity of Internet, networks became a way to unleash emotions and unfairness. This is also a way of self-presentation. Networks provide platforms for self-presentation; people who did not get fairness can warn others in networks. Therefore, we propose H5.

*Hypothesis 5.* Self-presentation in social networks can positively promote adolescents' network warning behaviors.

*2.2. Intermediary Role of Social Ability Cognition.* There are many factors influencing network altruistic behaviors. Previous studies mainly analyzed the generation of network altruistic behaviors from three aspects: altruists, network environment, and recipients [21]. Researchers have found that the occurrence of adolescent network altruistic behaviors was highly correlated with altruistic factors, such as gender, altruism, and other factors [6], but few focused on the cognitive factors of altruistic behaviors. The researchers pointed out that the cognitive factors of altruists had a great influence on the implementation of altruistic behaviors [27]. Therefore, we mainly explore the influence mechanism of self-presentation in social networks on adolescent network altruistic behaviors from the perspective of altruists' cognitive factors. Social ability cognition refers to an individual's evaluation of his ability to establish and maintain social relationships [14]. Adolescents' positive cognition and thinking of self-ability may enhance their confidence in helping others online, thereby promoting the generation of network altruistic behaviors. Adolescents are eager to communicate more with their peers and friends [28], and their evaluation of social networking ability, which we also call social ability perception, can affect their network behaviors. At the same time, self-presentation in social networks can stimulate

adolescents' self-awareness, and self-awareness can promote adolescents' self-exploration, further generating new evaluations and thinking strategies of their abilities [29]. Network social capabilities may enable adolescents to have a more positive view of their social capabilities and interpersonal relationships and then generate more network altruistic behaviors [5, 13]. Therefore, we propose Hypothesis 6.

*Hypothesis 6.* Social ability cognition plays a positive intermediary role between adolescents' self-presentation and altruistic behaviors in social networks.

*2.3. Moderating Role of Privacy Awareness.* Privacy awareness has led users to be more cautious when disclosing data. In social networks, especially online network environment, users are usually reluctant to fill in personal information or present their true situation out of fear that their information will be leaked or abused. In social networks of adolescents, since the moral concept of adolescents has not been finally constructed, their attention to private information is not strong. When self-presenting, they usually perform real self-presentation based on the characteristics of acquaintance socialization. The previous research pointed out that self-presentation may have a positive effect on adolescents' altruistic behaviors. However, under the privacy awareness, due to competition, information protection, and other reasons, adolescents may reduce their behaviors in social networks, such as filling out a questionnaire. If personal information needs to be authorized, many users will refuse to fill out the questionnaire, indicating that the strengthening of privacy awareness may reduce adolescents' altruistic behaviors. In addition, in the process of self-presentation, due to information protection considerations, adolescents will also hide personal information in self-presentation. For example, users in WeChat usually use mosaic to hide core information when they share experience; when speaking on social platforms such as Weibo, users prefer to use nicknames instead of real names. These behaviors will reduce interests for other users [6]; as a result, other users are also reluctant to share experience in network sites. It can be said that, under the effect of privacy awareness, the effect of self-presentation on adolescent network altruistic behaviors may be affected. Through the above analysis, we propose Hypothesis 7.

*Hypothesis 7.* Privacy awareness plays a negative moderating role between adolescents' self-presentation and altruistic behaviors in social networks.

The theoretical framework and hypothesis of adolescents' self-presentation and social network altruistic behaviors is summarized in Figure 1.

### 3. Materials and Methods

*3.1. Subject.* The participants of this study were 575 students from three junior high schools and three senior high schools in Shaanxi, Henan, and Zhejiang provinces. A total of 575 questionnaires were issued and 495 valid questionnaires

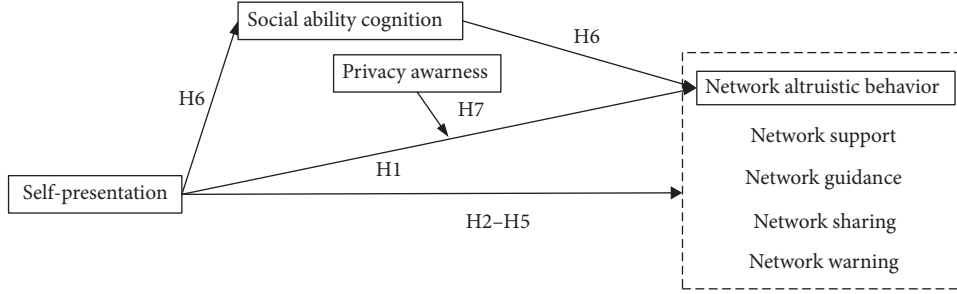


FIGURE 1: Theoretical framework of adolescents' self-presentation and social network altruistic behaviors.

were recovered with an effective recovery rate of 85.91%. Among them, there are 260 boys (52.63%) and 235 girls (47.36%). The average age of the subjects was 14.53 years (SD = 1.68, age range is from 12 to 18).

**3.2. Variable Selection and Measurement.** The network altruistic behaviors were measured by a scale compiled by Wright and Li [6]. A total of 26 projects were divided into four dimensions, which are network support (9 projects, such as sharing with others the experiences and perceptions of their lives in social networks), network guidance (6 projects, such as creating a platform for a person to communicate), network sharing (6 items, such as sharing your successful learning experience with others in social networks), and network warning (5 items, such as exposing some illegal events on the Internet to remind others to pay attention). The scale was scored at 4 points (1 “never” to 4 “always”). The scale has been used to measure the level of network altruistic behaviors among adolescents with good reliability and validity. In this study, the questionnaire’s confirmatory factor analysis fit the indicators well:  $\chi^2/df = 6.34$ , RMSEA = 0.06, NFI = 0.90, GFI = 0.90, CFI = 0.92, the internal consistency coefficient  $\alpha$  of the scale was 0.971, and coefficient  $\alpha$  of each dimension was 0.909 (network support), 0.922 (network guidance), 0.872 (network sharing), and 0.940 (network warning).

Social network self-presentation was measured by a self-presentation questionnaire prepared by Kim and Lee [27]. The self-presentation questionnaire included 4 items to measure the degree to which individuals present their true emotions and thoughts in social networks (e.g., I will post pictures showing my true idea; I don’t mind sharing some of bad feelings in social networks, etc.). The questionnaire used 5 points (1 “completely inconsistent” to 5 “completely consistent”). The scale has been applied to measure the individual’s social network self-presentation strategy with good reliability and validity. In this study, the questionnaire’s confirmatory factor analysis fit well:  $\chi^2/df = 4.36$ , RMSEA = 0.08, NFI = 0.995, GFI = 0.998, CFI = 0.997. The internal consistency coefficient  $\alpha$  of the questionnaire was 0.798.

Social ability cognition was measured by a network social ability questionnaire compiled by Jeong and Kim [14]. The questionnaire in this study had 4 items (for example, in online communication, I easily make friends with others; I often participate in online activity groups, etc.), using 5 points (1 “completely inconsistent” to 5 “completely

consistent”). In this study, the confirmatory factor analysis of the questionnaire fit well:  $\chi^2/df = 2.47$ , RMSEA = 0.04, NFI = 0.998, GFI = 0.999, CFI = 0.999, and the internal consistency coefficient  $\alpha$  was 0.934.

Privacy awareness was measured by a privacy awareness questionnaire prepared by Kishalay et al (2018) [28], including 4 items (for example, I am concerned that the information submitted on the social platform will be misused; I am unwilling to fill in personal information on the social platform because they may be used in an unpredictable way, etc.), using 5 points (1 “completely inconsistent” to 5 “completely consistent”). In this study, the questionnaire’s confirmatory factor analysis fit well:  $\chi^2/df = 2.24$ , RMSEA = 0.02, NFI = 0.998, GFI = 0.999, CFI = 0.999, and the internal consistency coefficient  $\alpha$  was 0.910.

**3.3. Procedures and Data Processing.** In this study, after obtaining the consent of the school leaders, teachers, and students themselves, the rigorously trained investigators explained the questionnaire practices in accordance with standardized guidelines, using a unified questionnaire to conduct group testing in units of classes, and all questionnaires were recovered on the spot. In this study, SPSS22.0 software was used to analyze the data, principal component analysis was used for project packaging and dimensionality reduction, and deviation-corrected percentile Bootstrap method was used to test the intermediary role of social ability cognition.

First, we needed to reduce the dimensionality of the variables. Using the principal component analysis method, we first standardized the original data. In  $n$  objects, apply  $x_1, x_2, \dots, x_m$  to represent the  $m$  variables of the principal component analysis index, and  $a_{ij}$  represents the value of the  $i$ -th evaluation object corresponding to the  $j$ -th index. Convert each index value  $a_{ij}$  into a standardized index  $a_{ij}^*$ , which is

$$a_{ij}^* = \frac{a_{ij} - \mu_j}{s_j}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m. \quad (1)$$

Among it,

$$\mu_j = \frac{1}{n} \sum_{i=1}^n a_{ij},$$

$$s_j = \frac{1}{n-1} \sum_{i=1}^n (a_{ij} - \mu_j)^2. \quad (2)$$

Accordingly, the standardized indicator variable is

$$x_{ij}^* = \frac{x_j - \mu_j}{s_j}, \quad j = 1, 2, \dots, m. \quad (3)$$

Second, calculate the correlation matrix  $R$ ,  $R = (r_{ij})_{m \times m}$  where

$$r_{ij} = \frac{\sum_{k=1}^n a_{kj}^* a_{ki}^*}{n-1}, \quad i, j = 1, 2, \dots, m. \quad (4)$$

Furthermore, the eigenvalues and eigenvectors of the correlation coefficient matrix were calculated. Solving the characteristic equation  $|\lambda \mathbf{I} - \mathbf{R}| = 0$ , we got the eigenvalue  $\lambda_i$ ,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$  and calculated the corresponding eigenvector  $u_i$ . The  $m$  new indexes composed of the eigenvectors are

$$\begin{cases} y_1 = u_{11}x_1^* + u_{21}x_2^* + \dots + u_{m1}x_m^*, \\ y_2 = u_{12}x_1^* + u_{22}x_2^* + \dots + u_{m2}x_m^*, \\ \dots \\ y_m = u_{1m}x_1^* + u_{2m}x_2^* + \dots + u_{mm}x_m^*. \end{cases} \quad (5)$$

Among them,  $y_1$  is the first principal component,  $y_2$  is the second principal component, and  $y_m$  is the  $m$ -th principal component.

Finally,  $p$  ( $p \leq m$ ) principal components were selected according to the information contribution rate and the cumulative contribution rate, and the comprehensive evaluation value was calculated.

The information contribution rate  $b_j$  and the cumulative contribution rate  $a_p$  are

$$\begin{aligned} b_j &= \frac{\lambda_j}{\sum_{k=1}^m \lambda_k}, \\ a_p &= \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^m \lambda_k}. \end{aligned} \quad (6)$$

If  $a_p$  is close to 1 (generally 85%–95%), the first  $p$  index variables are used as  $p$  principal components, and  $b_j$  is used as the weight to calculate the comprehensive score value; that is,

$$Z = \sum_{j=1}^p b_j y_j. \quad (7)$$

SPSS software was used to package adolescent self-presentation, network altruistic behaviors, social ability cognition, and privacy awareness scale items. The scree plots of four major variables are shown in Figure 2. Social network altruistic behaviors were packaged as four indicators, and self-presentation, social ability cognition, and privacy awareness were all packaged into one indicator.

## 4. Results and Discussion

**4.1. Common Method Biases and Descriptive Analysis.** After data collection, this study used Harman single factor analysis to statistically test the deviation of the common

method. The results showed that there were 8 factors with eigenvalues greater than 1 without rotation, explaining 56.53% of the variation, and the first factor explained 29.84% of the variation, which is less than the 40% judgment standard recommended by the predecessors, so this study did not have serious common method biases. The statistical description of the variables is shown in Table 1.

The results show that self-presentation is positively correlated with adolescent network altruistic behaviors; the correlation coefficient is 0.808. Moreover, self-presentation is positively correlated with each dimension of network altruistic behaviors; the correlation coefficients are 0.686, 0.760, 0.959, and 0.713. Social ability cognition and privacy awareness are significantly positively correlated with self-presentation and network altruistic behaviors. Age is significantly negatively correlated with adolescent network sharing and gender is significantly negatively correlated with adolescent altruistic behaviors, network support, network guidance, and network sharing, as well as self-presentation. Weekly online time is positively correlated with network support. Therefore, in the follow-up study, age, gender, and weekly online time were used as control variables.

Furthermore, use P-P plot to observe the data distribution of the main explanatory variables, which are self-presentation and adolescent network altruistic behaviors (Figure 3). The P-P plot shows that the data of self-presentation and adolescent network altruism are approximately concentrated in direct vicinity, indicating that the data follow a normal distribution and satisfy the basic assumption of linear regression.

**4.2. Econometric Model.** The explained variable in this paper is adolescent altruistic behaviors in social networks. This variable belongs to discrete and ordered variables, so the Ordered-Probit model was selected for analysis. The effect of self-presentation on network altruistic behaviors is expressed as follows:

$$Y_i^* = x_i' \beta_i + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (8)$$

Among them,  $Y_i^*$  represents altruistic behaviors,  $x_i'$  represents factors that affect altruistic behaviors, including explanatory variables and control variables,  $\beta_i$  is the coefficient, and  $\varepsilon_i$  is the random error term and follows a standard normal distribution. The set values of  $Y_i^*$  are 1, 2, 3, and 4, and the specific meanings are as follows: if  $Y_i^* < \xi_1$ ,  $Y_i = 1$  means that the individual never performs altruistic behaviors; if  $\xi_1 < Y_i^* < \xi_2$ ,  $Y_i = 2$  means that the individual occasionally engages in altruistic behaviors; if  $\xi_2 < Y_i^* < \xi_3$ ,  $Y_i = 3$  means that the altruistic behaviors are usually performed; if  $Y_i^* > \xi_3$ ,  $Y_i = 4$  means that the individual always engages in altruistic behaviors. Among them,  $\xi_1 < \xi_2 < \xi_3 < \xi_4$  is the parameter to be estimated, which becomes the “cutting point.”

Use  $\varphi$  to represent the cumulative density function of the standard normal distribution, and then the probability of  $Y_i$  at each value is

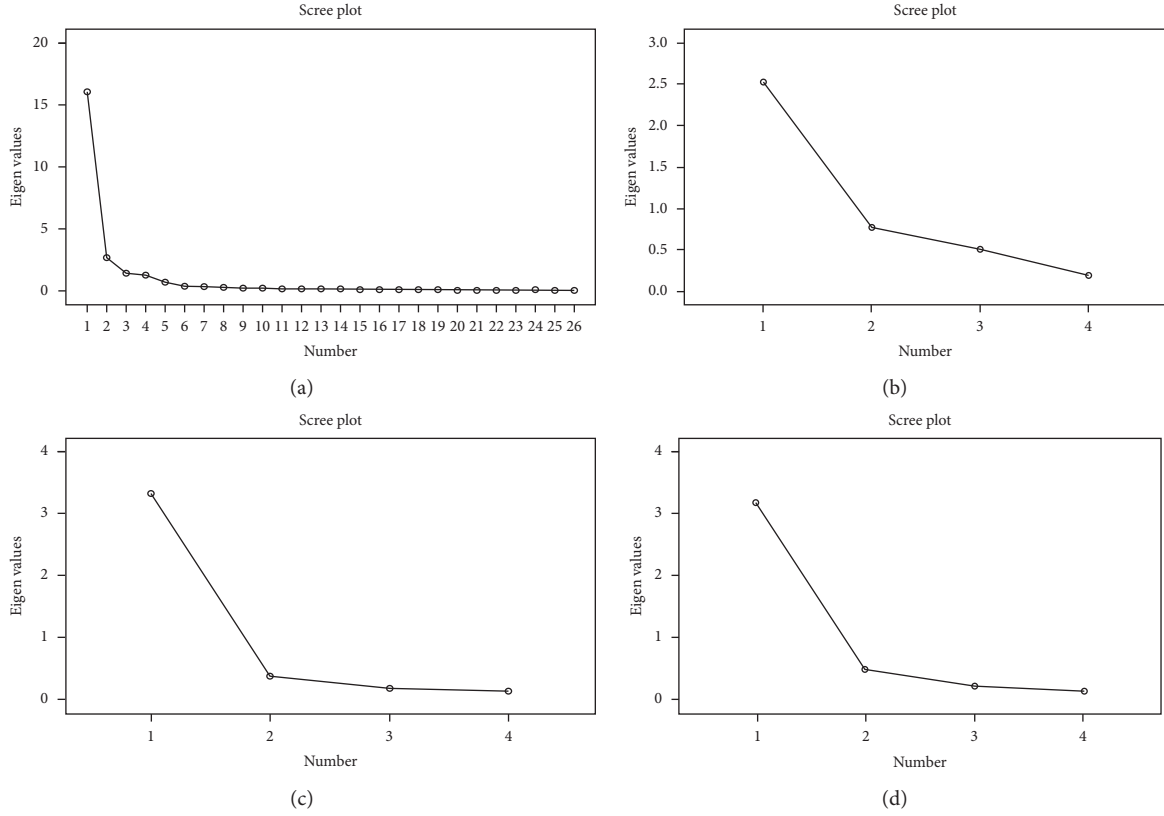


FIGURE 2: Scree plots of adolescent altruistic behavior (a), self-presentation (b), social ability cognition (c), and privacy awareness (d).

TABLE 1: Mean, standard deviation, and correlation coefficient of variables.

	M	SD	1	2	3	4	5	6	7	8	9	10	11
(1) Age	14.53	1.68	1										
(2) Gender	0.53	0.5	-0.018	1									
(3) Weekly online time	1.91	0.91	-0.060	0.078	1								
(4) Altruistic behaviors	2.68	0.85	-0.008	-0.110*	0.047	1							
(5) Network support	2.76	0.67	-0.016	-0.116**	0.033*	0.958**	1						
(6) Network guidance	2.52	0.65	-0.018	-0.115*	0.051	0.976**	0.909**	1					
(7) Network sharing	2.42	0.66	-0.016*	-0.103*	0.082	0.878**	0.764**	0.845**	1				
(8) Network warning	2.51	0.65	0.018	-0.083	0.020	0.959**	0.913**	0.925**	0.764**	1			
(9) Self-presentation	2.33	0.67	0.006	-0.116**	0.076	0.808**	0.686**	0.760**	0.959**	0.713**	1		
(10) Social ability cognition	2.41	0.71	-0.058	-0.078	0.084	0.850**	0.773**	0.840**	0.917**	0.715**	0.780**	1	
(11) Privacy awareness	2.53	0.69	-0.042	-0.072	0.076	0.864**	0.762**	0.867**	0.935**	0.736**	0.809**	0.980**	1

Note:  $N = 495$ , gender is a dummy variable, girls = 0, boys = 1. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

$$\begin{aligned}
 P(Y_{i=1} | x'_i) &= P(PY_i^* < \xi_1 | x'_i) = \varphi(\xi_1 - x'_i\beta_i), \\
 P(Y_{i=2} | x'_i) &= P(\xi_1 < Y_i^* < \xi_2 | x'_i) = \varphi(\xi_2 - x'_i\beta_i) - \varphi(\xi_1 - x'_i\beta_i), \\
 P(Y_{i=3} | x'_i) &= P(\xi_2 < Y_i^* < \xi_3 | x'_i) = \varphi(\xi_3 - x'_i\beta_i) - \varphi(\xi_2 - x'_i\beta_i), \\
 P(Y_{i=4} | x'_i) &= P(Y_i > \xi_3 | x'_i) = 1 - \varphi(\xi_3 - x'_i\beta_i).
 \end{aligned} \tag{9}$$

Furthermore, the likelihood function of the whole sample, the estimated value of the coefficient, and the

Ordered-Probit model could be obtained. The specific form of the benchmark model in this paper is

$$AB_i = \alpha \cdot SP_i + \mathbf{PQ} + \mu_i, \tag{10}$$

where  $AB_i$  represents the altruistic behaviors of the  $i$ -th sample,  $SP_i$  is the self-presentation of the  $i$ -th sample,  $Q$  is the control variable matrix, and  $\mu_i$  is the random error term (subject to normal distribution,  $i = 1, 2, \dots, n$ ),  $\alpha, \beta$  is the parameter to be estimated.



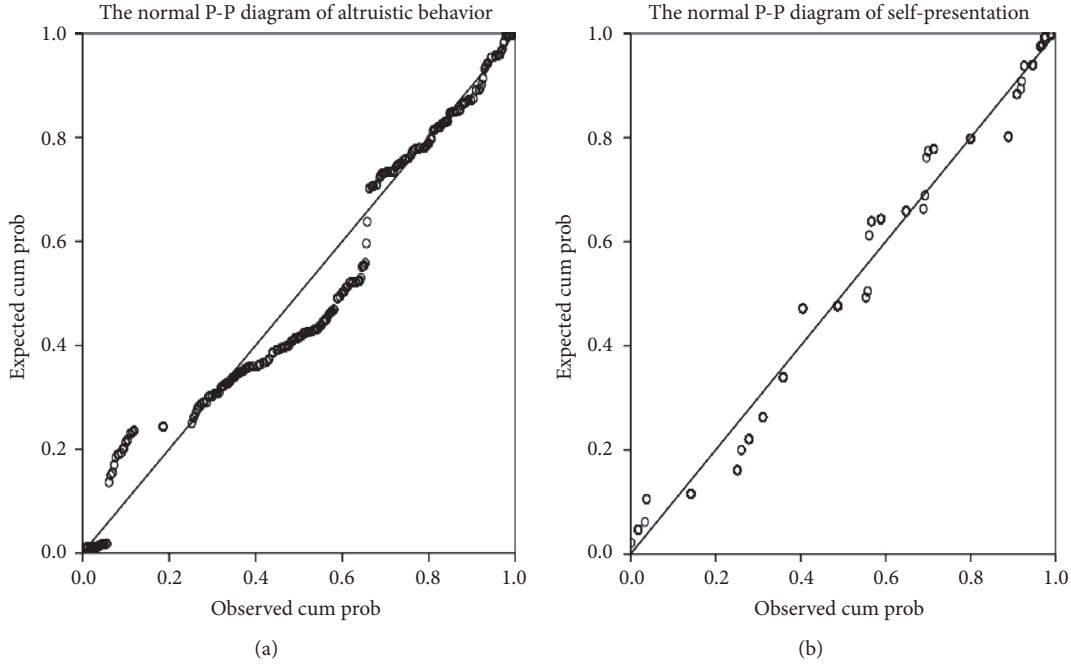


FIGURE 3: P-P plots of adolescent altruistic behaviors (a) and self-presentation (b).

According to the research of Wu et al. [29], testing the intermediary effect is required to test the parameters of the following three regression equations:

$$AB_i = cSP_i + \mathbf{PQ} + e_1, \quad (11)$$

$$SE_i = aSP_i + \mathbf{PQ} + e_2, \quad (12)$$

$$AB_i = c'SP_i + bSE_i + \mathbf{PQ} + e_3. \quad (13)$$

Among them,  $SE$  is social ability cognition,  $Q$  is the control variable matrix, and  $e_1 \sim e_3$  are regression residuals.  $c$  is the total effect of independent variable (self-presentation,  $SP$ ) on dependent variable (adolescent network altruistic behaviors,  $AB$ );  $a$  is the effect of independent variable (self-presentation,  $SP$ ) on intermediary variable (social ability cognition,  $SE$ );  $b$  is the effect of intermediary variable on the dependent variable after controlling the independent variable;  $c'$  is the direct effect of the independent variable on dependent variables after controlling the intermediary variable. If  $a$ ,  $b$ , and  $c$  are all significant, the mediation effect is significant; otherwise, there are other effects. For example, if  $c$  is not significant, there is a masking effect; if  $c'$  is significant, there also exist direct effects. Besides, the non-parametric percentile Bootstrap method with bias correction was used to test the mediation effect; if the 95% confidence interval does not contain zero, the mediation effect is significant [29].

For the moderating role of privacy awareness, the following econometric model is constructed:

$$AB_i = \alpha SP_i + \beta PR_i + \gamma SP_i * PR_i + \mathbf{PQ} + e_4. \quad (14)$$

Among them,  $PR$  is privacy awareness. In order to explore the independent effects among core variables, gender and age were controlled during regression.

#### 4.3. Result

**4.3.1. Main Result.** First, we examined the effects of self-presentation on adolescent network altruistic behaviors. The results are shown in Table 2. Adjusted  $R^2$  shows that except for the fit of self-presentation to the network support, which is slightly lower than 0.5, other regressions have reached the ideal standard, indicating that regression can better reflect the relationship between variables.

From the results in Table 2, self-presentation has a significant positive effect on adolescent network altruistic behaviors ( $\beta = 0.808$ ,  $p < 0.001$ ), and Hypothesis 1 is verified. For specific network altruistic behaviors, self-presentation has a significant positive effect on adolescents' network support, network guidance, network sharing, and network warning, with coefficients of 0.684, 0.757, 0.959, and 0.716, all significant at 1% level; Hypotheses 2–5 are supported.

#### 4.3.2. Results of Intermediary Effect and Moderating Effect

**(1) Result of Intermediary Effect.** In order to test the mediating effect of social ability cognition on self-presentation and adolescent network altruistic behaviors, three models were constructed and equations (11)–(13) were regressed (Table 3). The Adjusted  $R^2$  shows that the model has a good fit. In equation (11), self-presentation has a significant effect on adolescent network altruistic behaviors,  $c = 0.808$ ,  $p < 0.001$ . In equation (12), self-presentation has a significant influence on social ability cognition,  $a = 0.780$ ,  $p < 0.001$ . In equation (13), independent variable (self-presentation) and intermediary variable (social ability cognition) have significant effects on network altruistic behavior,  $c' = 0.369$ ,  $b = 0.565$ ,  $p < 0.001$ . According to the intermediary effect test

TABLE 2: Regression results of self-presentation and adolescent network altruistic behaviors.

Variables	Altruistic behaviors	Network support	Network guidance	Network sharing	Network warning
Constant	0.169 (0.698)	0.266 (0.892)	0.243 (0.911)	0.162 (1.387)	-0.026 (-0.090)
Age	-0.008 (-0.533)	-0.013 (-0.675)	-0.014 (-0.802)	-0.013* (-1.668)	0.007 (0.347)
Gender	-0.031 (-0.571)	-0.071 (-1.071)	-0.054 (-0.915)	0.015 (0.597)	0.008 (0.119)
Weekly online time	-0.016 (-0.547)	-0.020 (-0.550)	-0.006 (-0.194)	0.007 (0.528)	-0.038 (-1.087)
Self-presentation	0.808*** (30.089)	0.684*** (20.629)	0.757*** (25.560)	0.959*** (74.140)	0.716*** (22.394)
Adj-R <sup>2</sup>	0.651	0.469	0.575	0.919	0.506

*t* value is in parentheses. Symbols \*, \*\*, and \*\*\* indicate that the coefficient is significant at 10%, 5%, and 1% levels.

TABLE 3: Regression results of intermediary effects of social ability cognition.

Variables	Equation (11), altruistic behaviors	Equation (12), social ability cognition	Equation (13), altruistic behaviors
Constant	0.169 (0.698)	0.475* (1.858)	-0.100 (-0.513)
Age	-0.008 (-0.533)	-0.036** (-2.179)	0.012 (0.956)
Gender	-0.031 (-0.571)	0.020 (0.354)	-0.042 (-0.975)
Weekly online time	-0.016 (-0.547)	0.022 (0.703)	-0.029 (-1.204)
Self-presentation	0.808- (30.089)	0.780** (27.472)	0.369*** (10.682)
Social ability cognition			0.565*** (16.50)
Adj-R <sup>2</sup>	0.651	0.613	0.778

*t* value is in parentheses. Symbols \*\* and \*\*\* indicate that the coefficient is significant at 10%, 5%, and 1% levels.

procedure of Wu et al. [29], it can be concluded that the intermediary effect of social ability cognition is significant.

Furthermore, use bias-corrected Bootstrap method to test the intermediary effect of social ability cognition. 5000 Bootstrap samples were randomly selected from the original samples for direct and indirect effect estimation. Table 4 shows the direct effect of self-presentation on network altruistic behaviors and the indirect effect of social ability cognition. Meanwhile, Table 4 shows the 95% confidence interval. If the 95% confidence interval does not contain zero, the intermediary effect is significant. It can be seen from Table 4 that the coefficient of indirect effect caused by social ability cognition is 0.4411, and the 95% confidence interval is [0.3825, 0.4995], excluding zero, so the intermediary effect of social ability cognition is significant; Hypothesis 2 is supported.

(2) *Moderating Effect.* A stepwise regression method was used to further examine the moderating role of privacy awareness in the relationship between self-presentation and adolescent network altruistic behaviors. The regression results are shown in Table 5, which shows that although both self-presentation and privacy awareness have a significant positive effect on adolescent network altruistic behaviors, the interaction of self-presentation and privacy awareness has no significant effect, indicating that the moderating effect of privacy awareness is not significant.

In addition, this study also found that age had a negative predictive effect on adolescent network sharing; that is, the older the age, the lower the possibility of adolescent network sharing. Sex and weekly online time had no significant effect on the prediction of adolescent network altruistic behaviors.

#### 4.4. Discussion

*4.4.1. Influence of Self-Presentation in Social Networks on Adolescent Altruistic Behaviors.* This study shows that self-presentation in social networks can significantly predict adolescent network altruistic behaviors. Adolescents often have a strong desire to show themselves, and the popularity of social networking sites provides them with space to express and explore themselves [18]. Adolescents can present themselves in social networks by sharing photos and updating their dynamics. They can also display positive personal information and shape ideal self-images through active self-presentation, or they can show their true life status and ideas to others and conduct in-depth self-disclosure through real self-presentation. Adolescents with high levels of self-presentation in social networks may be more willing to participate in online socialization and treat themselves as a member of the network community [3], and this sense of belonging to the social network environment can promote altruistic behaviors. At the same time, self-presentation in social networks can enable adolescents to experience social support and intimate network interpersonal relationships in interactions with others [19], which may inspire adolescents to repay and make them willing to use the altruistic behaviors to reward the perceived positive network atmosphere. From the perspective of self-awareness theory, self-presentation in social networks can improve adolescents' self-awareness and make adolescents more sensitive to social norms and more inclined to do something helpful to society and others [8]; this can also promote the generation of network altruism.

Based on coefficients, adolescent self-presentation has the largest positive prediction effect on network sharing, and the least positive prediction effect on network support,

TABLE 4: Bootstrap significance test of social ability cognition.

Direct effect of self-presentation on altruistic behavior					
	Effect	se	$t$ ( $p$ )	LLCI	ULCI
	0.3668	0.0343	10.6832 (0.0000)	0.2993	0.4343
Indirect effect of self-presentation on altruistic behavior					
	Effect	Boot standard error		Boot LLCI	Boot ULCI
Social ability cognition	0.4411	0.0298		0.3825	0.4995

TABLE 5: Regression results of moderating effect of privacy awareness.

Variables	Equation (14), altruistic behavior			
Constant	-0.150 (-0.726)	-0.047 (-0.242)	-0.035 (-0.183)	
Age	0.016 (1.165)	0.008 (0.655)	0.008 (0.644)	
Gender	-0.094 (-2.053)	-0.056 (-1.318)	-0.054 (-1.248)	
Weekly online time	-0.015 (-0.595)	-0.023 (-0.965)	-0.024 (-1.003)	
Self-presentation		0.312*** (8.631)	0.314*** (8.651)	
Privacy awareness	0.862*** (37.858)	0.611*** (16.961)	0.614*** (16.899)	
Self-presentation * Privacy awareness			-0.011 (-0.666)	
Adj- $R^2$	0.747	0.780	0.780	

$t$  value is in parentheses. Symbols \*, \*\*, and \*\*\* indicate that the coefficient is significant at 10%, 5%, and 1% levels.

which may be related to the adolescents' knowledge reserve and network motivation in reality. Since most adolescents are in junior high school or senior high school stage, under the practical pressure of China's high school entrance examination and college entrance examination, the knowledge in this stage is mainly based on mathematical, physical, and chemical knowledge and is not much connected with the computer network and other aspects of knowledge. Thus, there are not many support behaviors in social networks [30]. For example, adolescents rarely upload useful programs. But this does not mean that adolescents are unwilling to provide online support; some research results pointed out that as long as adolescents have sufficient knowledge reserves, they are extremely willing to provide network support in the process of self-presentation. In addition, judging from the motivation of adolescents, Chinese adolescents mainly share on social network platforms, such as personal experience sharing on WeChat and Weibo. Because the motivation of adolescents to go online is mainly sharing activities, self-presentation has the greatest effect on the positive prediction of network sharing.

*4.4.2. Intermediary Role of Social Ability Cognition and the Moderating Role of Privacy Awareness.* The results of this study show that social ability cognition plays an intermediary role between self-presentation and network altruistic behaviors. This intermediary role may be explained from the following aspects. First, self-presentation in social networks can bring adolescents a successful experience of network socialization to a certain extent, because self-presentation can not only provide adolescents with the opportunity to exercise social skills, but also bring positive network feedback and social support [11, 24]. Successful experience is one of the most important factors for an individual to gain a sense of efficacy [15], so self-presentation in social networks

has an important positive predictive effect on adolescents' social ability cognition. Secondly, adolescents with a higher level of social ability cognition are usually more confident in their social ability; they believe that they can help others and may show more altruistic behaviors in the network environment. Finally, the behaviors of self-presentation in social networks can affect adolescents' cognitive factor of social ability evaluation, which in turn affects their network altruistic behaviors. Specifically, active self-presentation can shape a positive self-image and help adolescents focus on their positive aspects, so they have a more positive evaluation and understanding of themselves [31], which may make adolescents more positive about their social ability cognition. When adolescents have positively evaluated their abilities, they may be more willing to use their abilities to implement altruistic behaviors. At the same time, according to the expansion-construction theory of positive emotions, self-presentation in social networks brings positive emotional experiences that can make adolescents more positively evaluate their abilities and thinking, which may promote the generation of online altruistic behaviors [32].

The moderating role of privacy awareness between self-presentation and adolescent network altruistic behaviors is not significant. This may be related to the generally low privacy awareness of adolescents. In the process of using social networks, adolescents, despite perceiving the risks of privacy, still disclose a lot of their personal privacy when meeting the needs of communication, self-expression, and the possible rewards of revealing personal information, and rarely consciously protecting personal privacy information [33]. Although the strengthening of privacy awareness strengthens adolescents' confidence in social networks and may promote network altruistic behaviors, the overall low privacy awareness will not significantly change the self-presentation of adolescent in networks and thus will not

significantly regulate the impact of self-presentation on network altruistic behaviors.

## 5. Conclusions

This study examined the impact of adolescents' self-presentation on social network altruistic behaviors and explored its internal mechanisms to enrich previous research. Furthermore, this study explored the intermediary role of social ability cognition and moderating role of privacy awareness in the relationship between self-presentation and adolescent network altruistic behaviors. The research results support the self-awareness theory: self-presentation in social networks not only directly affects adolescent altruistic behaviors, but also promotes the generation of altruistic behaviors by influencing adolescents' social ability cognition. The moderating effect of privacy awareness during the relationship between self-presentation and altruistic behavior is not significant, but privacy awareness significantly affects adolescent altruistic behaviors. The results further indicate the importance of altruistic cognitive factors in network altruistic behaviors. From a practical point, in order to promote adolescent network altruistic behaviors and cultivate their positive psychological qualities, adolescents should be encouraged to properly present themselves in social networks and actively integrate into the network environment. At the same time, related mental health counselling courses can also be set up to guide adolescents to establish positive self-awareness and strengthen privacy awareness.

Though we have analyzed and tested the adolescent altruistic behaviors in social networks, this study also has some shortcomings. First of all, the research mainly used cross-sectional sample data and analyzed from the perspective of individuals; it is hard to obtain the dynamic process of variables in the process of individual development. Secondly, this study found that the privacy awareness did not have a moderating role in the relationship between self-presentation and network altruistic behaviors, which may be affected by the sample data or other factors. Future research can expand the sample for empirical study or explore other roles of privacy awareness. Finally, the data in this study came from the self-reporting of adolescents; future research can analyze the real data in the network environment, for example, based on the big data of online social media, through text analysis and other methods to determine the type of individual self-presentation on the social networking site and collect information about the individual's network altruistic behaviors, thereby obtaining more reliable results.

## Data Availability

The survey 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 study was supported by the National Natural Science Foundation of China (71974157) and the Postdoctoral Science Foundation Project of China (2019M653536).

## References

- [1] Y. Amichai-Hamburger and A. Furnham, "The positive net," *Computers in Human Behavior*, vol. 23, no. 2, pp. 1033–1045, 2007.
- [2] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 25, no. 3, pp. 1–13, 2019.
- [3] J. M. Fatkin, "Pro"social Media: Using Key Social Psychological Theories to Increase Prosocial Engagement on Social Media Sites, Heriot-Watt University, Edinburgh, Scotland, 2015.
- [4] J. B. Walther, "Computer-mediated communication," *Communication Research*, vol. 23, no. 1, pp. 3–43, 1996.
- [5] G. V. Caprara, B. P. L. Kanacri, M. Gerbino et al., "Positive effects of promoting prosocial behavior in early adolescence," *International Journal of Behavioral Development*, vol. 38, no. 4, pp. 386–396, 2014.
- [6] M. F. Wright and Y. Li, "The associations between young adults' face-to-face prosocial behaviors and their online prosocial behaviors," *Computers in Human Behavior*, vol. 27, no. 5, pp. 1959–1962, 2011.
- [7] R. F. Baumeister, "A self-presentational view of social phenomena," *Psychological Bulletin*, vol. 91, no. 1, pp. 3–26, 1982.
- [8] A. L. Gonzales and J. T. Hancock, "Mirror, mirror on my facebook wall: Effects of exposure to Facebook on self-esteem," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 1–2, pp. 79–83, 2011.
- [9] S. Pfattheicher and J. Keller, "The watching eyes phenomenon: The role of a sense of being seen and public self-awareness," *European Journal of Social Psychology*, vol. 45, no. 5, pp. 560–566, 2015.
- [10] Y. Shin, M. Kim, C. Im, and S. C. Chong, "Selfie and self: The effect of selfies on self-esteem and social sensitivity," *Personality and Individual Differences*, vol. 111, pp. 139–145, 2017.
- [11] H.-Y. Huang, "Examining the beneficial effects of individual's self-disclosure on the social network site," *Computers in Human Behavior*, vol. 57, pp. 122–132, 2016.
- [12] S. Utz, "The function of self-disclosure on social network sites: not only intimate, but also positive and entertaining self-disclosures increase the feeling of connection," *Computers in Human Behavior*, vol. 45, pp. 1–10, 2015.
- [13] L. Guo, Q. Zhang, W. Hu, Z. Sun, and Y. Qu, "Learning to complete knowledge graphs with deep sequential models," *Data Intelligence*, vol. 1, no. 3, pp. 289–308, 2019.
- [14] E. J. Jeong and D. H. Kim, "Social activities, self-efficacy, game attitudes, and game addiction," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 4, pp. 213–221, 2011.
- [15] A. Bandura, "Self-efficacy: toward a unifying theory of behavioral change," *Psychological Review*, vol. 84, no. 2, pp. 191–215, 1977.
- [16] A. Bandura, C. Pastorelli, C. Barbaranelli, and G. V. Caprara, "Self-efficacy pathways to childhood depression," *Journal of Personality and Social Psychology*, vol. 76, no. 2, pp. 258–269, 1999.
- [17] M. J. Culan, "How did they get my name?": An exploratory investigation of consumer attitudes toward secondary

- information use,” *MIS Quarterly*, vol. 17, no. 3, pp. 341–361, 1993.
- [18] D. Tobias and M. J. Metzger, “An extended privacy calculus model for SNS: Analyzing self-disclosure and self-withdrawal in a representative U. S. sample,” *Journal of Computer-Mediated Communication*, vol. 21, no. 5, pp. 368–383, 2016.
- [19] N. Park, B. Jin, and S.-A. Annie Jin, “Effects of self-disclosure on relational intimacy in facebook,” *Computers in Human Behavior*, vol. 27, no. 5, pp. 1974–1983, 2011.
- [20] J. G. Proudfoot, D. Wilson, J. S. Valacich, and M. D. Byrd, “Saving face on facebook: privacy concerns, social benefits, and impression management,” *Behaviour & Information Technology*, vol. 37, no. 1, pp. 16–37, 2018.
- [21] M. M. Al-Debei, E. Al-Lozi, and A. Papazafeiropoulou, “Why people keep coming back to facebook: Explaining and predicting continuance participation from an extended theory of planned behaviour perspective,” *Decision Support Systems*, vol. 55, no. 1, pp. 43–54, 2013.
- [22] L. M. Padilla-Walker, G. Carlo, and M. K. Memmott-Elison, “Longitudinal change in adolescents’ prosocial behavior toward strangers, friends, and family,” *Journal of Research on Adolescence*, vol. 28, no. 3, pp. 698–710, 2018.
- [23] P. D. Ellis, *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*, Cambridge University Press, Cambridge, UK, 2010.
- [24] C.-c. Yang and B. Bradford Brown, “Online self-presentation on facebook and self development during the college transition,” *Journal of Youth and Adolescence*, vol. 45, no. 2, pp. 402–416, 2016.
- [25] J. Jin, Y. Li, X. Zhong, and L. Zhai, “Why users contribute knowledge to online communities: an empirical study of an online social Q&A community,” *Information & Management*, vol. 52, no. 7, pp. 840–849, 2015.
- [26] R. L. Trivers, “The evolution of reciprocal altruism,” *The Quarterly Review of Biology*, vol. 46, no. 1, pp. 35–57, 1971.
- [27] J. Kim and J.-E. R. Lee, “The facebook paths to happiness: effects of the number of Facebook friends and self-presentation on subjective well-being,” *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 6, pp. 359–364, 2011.
- [28] K. Adhikari and R. K. Panda, “Users’ information privacy concerns and privacy protection behaviors in social networks,” *Journal of Global Marketing*, vol. 31, no. 2, pp. 96–110, 2018.
- [29] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A comprehensive survey on graph neural networks,” *IEEE Transactions on Neural Networks and Learning*, pp. 1–21, 2020.
- [30] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, “Topological influence-aware recommendation on social networks,” *Complexity*, vol. 2019, Article ID 6325654, Article ID 6325654, 12 pages, 2019.
- [31] C. L. Toma, “Affirming the self through online profiles: beneficial effects of social networking sites,” in *Proceedings of the International Conference on Human Factors in Computing Systems*, pp. 1749–1752, Atlanta, GA, USA, April 2010.
- [32] F. T. Villavicencio and A. B. I. Bernardo, “Beyond math anxiety: positive emotions predict mathematics achievement, self-regulation, and self-efficacy,” *The Asia-Pacific Education Researcher*, vol. 25, no. 3, pp. 415–422, 2016.
- [33] S. Ji, S. Pan, E. Cambria et al., “A survey on knowledge graphs: Representation, acquisition and applications,” 2020, <http://arxiv.org/abs/2002.00388>.

## Research Article

# Attention with Long-Term Interval-Based Deep Sequential Learning for Recommendation

Zhao Li <sup>1</sup>, Long Zhang,<sup>1</sup> Chenyi Lei,<sup>1</sup> Xia Chen,<sup>1</sup> Jianliang Gao <sup>2</sup>, and Jun Gao<sup>3</sup>

<sup>1</sup>Alibaba Group, Hangzhou, China

<sup>2</sup>Central South University, Changsha, China

<sup>3</sup>School of EECS, Peking University, Beijing, China

Correspondence should be addressed to Zhao Li; [lizhao.lz@alibaba-inc.com](mailto:lizhao.lz@alibaba-inc.com)

Received 19 May 2020; Revised 15 June 2020; Accepted 19 June 2020; Published 13 July 2020

Guest Editor: Shirui Pan

Copyright © 2020 Zhao 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.

Modeling user behaviors as sequential learning provides key advantages in predicting future user actions, such as predicting the next product to purchase or the next song to listen to, for the purpose of personalized search and recommendation. Traditional methods for modeling sequential user behaviors usually depend on the premise of Markov processes, while recently recurrent neural networks (RNNs) have been adopted to leverage their power in modeling sequences. In this paper, we propose integrating attention mechanism into RNNs for better modeling sequential user behaviors. Specifically, we design a network featuring Attention with Long-term Interval-based Gated Recurrent Units (ALI-GRU) to model temporal sequences of user actions. Compared to previous works, our network can exploit the information of temporal dimension extracted by time interval-based GRU in addition to normal GRU to encoding user actions and has a specially designed matrix-form attention function to characterize both long-term preferences and short-term intents of users, while the attention-weighted features are finally decoded to predict the next user action. We have performed experiments on two well-known public datasets as well as a huge dataset built from real-world data of one of the largest online shopping websites. Experimental results show that the proposed ALI-GRU achieves significant improvement compared to state-of-the-art RNN-based methods. ALI-GRU is also adopted in a real-world application and results of the online A/B test further demonstrate its practical value.

## 1. Introduction

Due to the increasing abundance of information on the Web, helping users filter information according to their preferences is more and more required, which emphasizes the importance of personalized search and recommendation [42–45]. Traditional methods for providing personalized content, such as item-item collaborative filtering [33], did not take into account the dynamics of user behaviors, which are recently recognized as important factors. For example, to predict the user's next action such as the next product to purchase, the profiling of both long-term preferences and short-term intents of user is required, where modeling the user's behaviors as sequences provides key advantages. Nonetheless, modeling sequential user behaviors with temporal dimension raises even more challenges than

modeling them without the temporal dimension. How to identify the correlation and dependence among actions is one of the difficult issues. This problem has been studied extensively, and many methods based on the Markov process assumption, such as Factorizing Personalized Markov Chain [32] and Hierarchical Representation Model [41], have been designed and adopted in different tasks [11, 15]. These methods usually focus on factor models, which decompose the sparse user-item interaction matrix into low-dimensional matrices with latent factors. However, for modeling sequential information, it is often not clear how to integrate the dynamics of user intents into the framework of factor models.

Recently, neural-network-based algorithms have received much attention from researchers [6, 9, 18, 30, 48]. For instance, many different kinds of graph neural networks

(GNNs), instead of matrix factorization (MF) based algorithms [36], have been proposed to learn graph embedding due to their ability of learning from non-Euclidean space [30]. Many different kinds of recurrent neural networks (RNNs) have been proposed to model user behaviors due to their powerful descriptive ability for sequential data [14, 29, 39, 47, 52]. For example, Hidasi et al. [14] propose an approach based on a number of parallel RNNs with rich features to model sequential user behaviors. Wu et al. [47] endow both users and movies with a long short-term memory (LSTM) autoregressive model to predict user future behaviors. Furthermore, to utilize the temporal information better, Zhu et al. [52], Neil et al. [29], and Vassøy et al. [39] introduce time intervals between sequential actions into RNN cells to update and forget information rather than only considering the order of actions.

Despite the success of abovementioned RNN-based methods, there are several limitations that make it difficult to apply these methods into the wide variety of applications in the real world. One inherent assumption of these methods is that the importance of historical behaviors decreases over time (e.g., equation (15) in [52]), which is also the intrinsic property of RNN cells such as gated recurrent units (GRU) and long- and short-term memory (LSTM). However, this assumption does not always apply in practice, where the sequences may have complex cross-dependence [46]. For example, user’s online actions are not straightforward but contain much noise and randomness. See Figure 1 for illustration, which shows a real sequence of clicked items of a user in one of the largest online shopping websites. We can conjecture the user intended to buy a T-shirt in honor of LeBron James and he/she finally bought  $Item_6$ . But the user also viewed items of different kind (shoes as  $Item_3$  and  $Item_4$ ). Clearly,  $Item_1$  and  $Item_2$  are more important than  $Item_3$  and  $Item_4$  to predict the final deal, although the former two items are earlier than the latter two items in the temporal dimension. This example displays the difficulty in analyzing sequential user behaviors, where the simple assumption of time interval-based correlation between actions is not enough to cope with.

In this paper, we are inspired by the attention mechanism proposed for natural language processing [1, 49] which achieves remarkable progress in the past few years. Attention mechanism introduced into deep networks provides the functionality to focus on portions of input data or features to fulfill the given task. Similarly, we expect that a trained attention mechanism helps to identify the important correlated actions from sequential user behaviors to make prediction. However, the existing attention mechanism is inefficient in modeling sequential user behaviors. Hence, we design a new attention mechanism specifically for our purpose.

Specifically, we propose a network featuring Attention with Long-term Interval-based Gated Recurrent Units (ALI-GRU) for modeling sequential user behaviors to predict user’s next action. The network is depicted in Figure 2. We adopt a series of bidirectional GRU to process the sequence of items that user had accessed. The GRU cells in our network consist of not only normal GRU but also time interval-based GRU, where the latter reflects the short-term

information of time intervals. In addition, the features extracted by bidirectional GRU are used as the input of attention model, and the attention distribution is calculated at each timestamp rather than as single vector as in Seq2Seq model [1, 46]. Therefore, this attention mechanism is able to consider the long-term correlation along with short-term intervals. Our designed attention mechanism is detailed in Section 4.

We have performed a series of experiments using both well-known public datasets (LastFM and CiteULike [52]) and the dataset collected by ourselves, built from real-world data. Extensive results show that our proposed ALI-GRU outperforms the state-of-the-art methods by a significant margin on these datasets. Moreover, ALI-GRU is adopted online and we have performed online A/B test; test results further demonstrate the practical value of ALI-GRU in comparison with the well-optimized baseline in a real-world e-commerce search engine.

This paper makes the following contributions:

- (i) First, we propose a bidirectional time interval-based GRU to model the long- and short-term information of user actions for better capturing temporal dynamics between actions. Time interval-based GRU is able to effectively extract short-term dynamics of user intents as driven signals of attention function and refine the long-term memory by contextual information.
- (ii) Second, we design a new attention mechanism to encode long- and short-term information and identify complex correlation between actions, which attends to the driven signals at each time step along with the embedding of contextual information. This mechanism is less affected by the noise in the historical actions and is robust to extract the important correlated information between sequential user behaviors to make a better prediction.
- (iii) Third, we conduct a series of experiments on two well-known public datasets and a large-scale dataset constructed from a real-world e-commerce platform. Extensive experimental results show that our proposed ALI-GRU obtains significant improvement compared to state-of-the-art RNN methods. In addition, ALI-GRU is adopted and we conducted online A/B test, and the results further demonstrate the practical value in a real-world e-commerce search engine.

The remainder of this paper is organized as follows. Section 2 discusses related works. The problem of modeling sequential user behaviors is formulated in Section 3, followed by detailed description of our proposed ALI-GRU in Section 4. Experimental results are reported in Section 5 and concluding remarks in Section 6.

## 2. Related Work

We give brief overview of related work at two aspects, modeling of sequential user behaviors and attention mechanism.

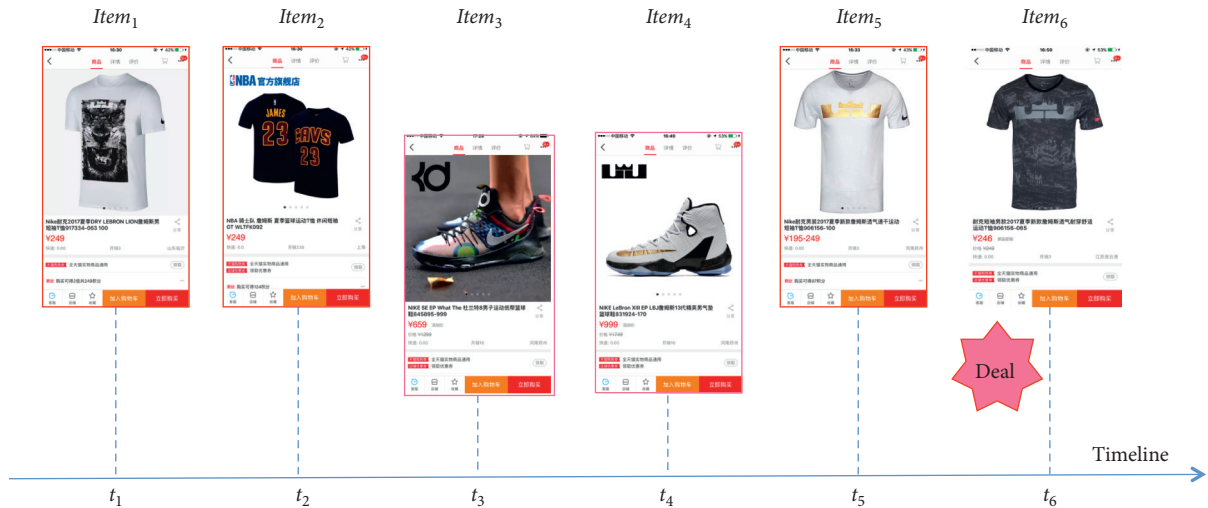


FIGURE 1: An example of sequential actions of a user in one of the largest online e-commerce platforms. The user clicked several items sequentially and finally purchased *Item*<sub>6</sub>. The red frame around each item indicates the corresponding (estimated) contribution of that item to the final deal; darker color stands for more contribution.

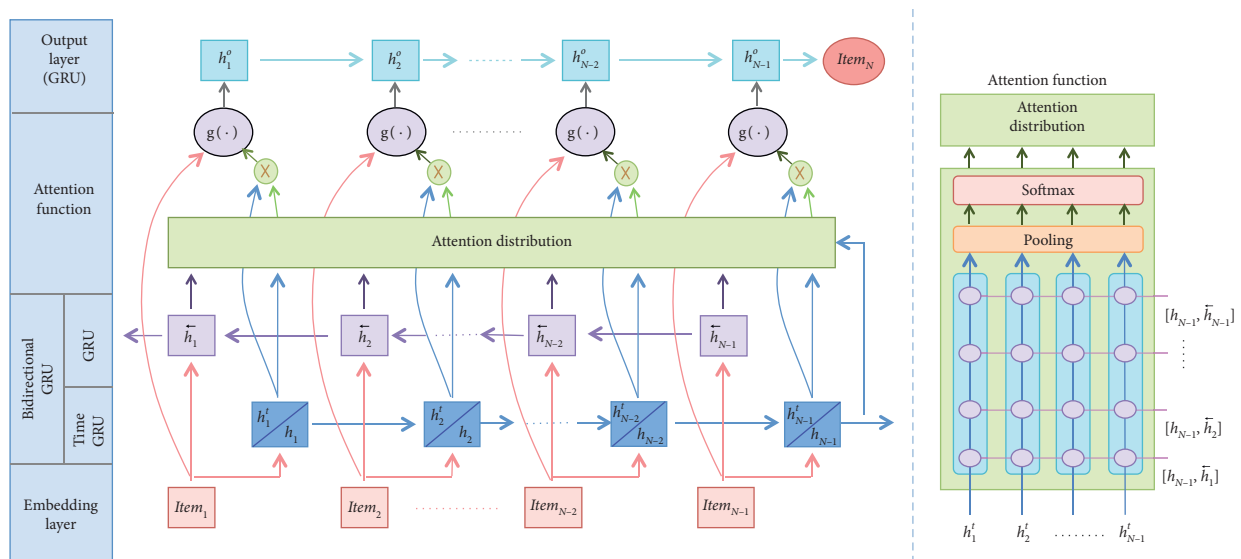


FIGURE 2: The proposed framework for modeling sequential user behaviors (left) and the designed attention mechanism (right).

2.1. *Modeling Sequential User Behaviors.* Due to the significance of user-centric tasks such as personalized search and recommendation, modeling sequential user behaviors has attracted great attention in both industry and academia. Most of the pioneering work relies on model-based Collaborative Filtering (CF) to analyze user-item interaction matrix. There have been a variety of such algorithms including Bayesian methods [28] and matrix factorization (MF) methods [31, 50]. Due to the characteristics of sequential information, several CF works take the temporal dynamics into account, often based on the assumption of Markov processes [11, 21, 32]. For the task of sequential recommendation, Rendle et al. [32] propose Factorizing Personalized Markov Chain (FPMC) to combine matrix factorization of user-item matrix with Markov chains. He and McAuley [11] further integrate similarity-based

methods [20] into FPMC to tackle the problem of sequential dynamics.

The major problems of the abovementioned work are that these methods independently combine several components, rely on low-level hand-crafted features of user or item, and have difficulty to handle long-term behaviors. On the contrary, along with the development of deep neural networks, Lei et al. [22] and Zheng et al. [51] employ deep learning to learn effective representations of user/item automatically. Furthermore, with the success of recurrent neural networks (RNNs) in the past few years, a paucity of work has made attempts to utilize RNNs [14, 25, 47]. For example, Liu et al. [25] consider jointly the contextual information such as weather into the RNN architecture to improve modeling performance. The insight that RNN-based solutions achieve success in modeling sequential user



behaviors is that RNN has well-demonstrated ability of capturing patterns in the sequential data. Recent studies [10, 29, 39, 52] also indicate that time intervals within sequential signal are a very important clue to update and forget information in RNN architecture. Zhu et al. [52] design several time gates in LSTM units to improve modeling performance. He et al. [10] embed items into a “transition space,” where users are modeled as translation vectors operating on item sequences. Liu et al. [26] employ adaptive context-specific input matrices and adaptive context-specific transition matrices to capture external situations and how lengths of time intervals between adjacent behaviors in historical sequences affect the transition of global sequential features, respectively. But, in practice, there is complex dependence and correlation between sequential user behaviors, which requires deeper analysis of relation among behaviors rather than simply modeling the presence, order, and time intervals. To summarize, how to design an effective RNN architecture to model sequential user behaviors effectively is still a challenging open problem.

**2.2. Attention Mechanism.** Attention mechanism is now a commonly adopted ingredient in various deep learning tasks such as machine translation [1, 27], image captioning [24], question answering [38], and speech recognition [5], which has been shown to be effective against capturing the contribution and correlation between different components in the network. The success of attention mechanism is mainly due to the reasonable assumption that human beings do not tend to process the entire signal at once; instead, they only focus on selected portions of the entire perception space when and where needed [17]. To avoid the limitation of normal networks that the entire source must be encoded to one hidden layer, the attention-based network contains a set of hidden representations that scale with the size of the source. The network learns to assign attention weights to perform a soft selection of these representations.

With the development of attention mechanism, recent researches start to leverage different attention architectures to improve performance of related tasks [1, 3, 34, 40, 46, 49]. For example, Bahdanau et al. [1] conjecture that the use of a fixed-length vector is a bottleneck in improving the performance of this basic encoder-decoder architecture; therefore, they design a model to automatically search for parts of a source sentence which are relevant to predicting a target word. Yang et al. [49] propose a hierarchical attention network at word and sentence level, respectively, to capture contributions of different parts of a document. Vaswani et al. [40] utilize multihead attention mechanism to improve performance. Wang et al. [46] propose a coverage strategy to combat the misallocation of attention caused by the memorylessness of traditional attention mechanism. Nevertheless, most of the previous work calculates attention distribution according to the interaction of every source vector with a single embedding vector of contextual or historical information (such as translated words in the sentence), which may lead to information loss caused by early summarization and noise caused by incorrect previous

attention. In particular, Shen et al. [35] propose an attention-based language understanding method without any other network structure (e.g., RNN). In [35], the input sequence is processed by directional (forward and backward) self-attentions to model context dependency and produce context-aware representations for all tokens. Then, a multidimensional attention computes a vector representation of the entire sequence.

Indeed, the attention mechanism is very important to the task of modeling sequential user behaviors. However, to the best of our knowledge, there are a few works concentrating on this paradigm. Chen et al. [2] consider the attention mechanism into a multimedia recommendation task with multilayer perceptron. Song et al. [37] propose a recommender system for online communities based on a dynamic-graph-attention neural network. They model dynamic user behaviors with a recurrent neural network and context-dependent social influence [23] with a graph-attention neural network, which dynamically infers the influencers based on users’ current interests. In this paper, an effective solution with attention mechanism for better modeling sequential user behaviors is to be investigated.

### 3. Problem Formulation

We start our discussion with the definitions of some notations. Let  $\mathcal{U}$  be a set of users and let  $\mathcal{I}$  be a set of items in a specific service such as products in online shopping websites. For each user  $u \in \mathcal{U}$ , his/her historical behaviors are given by  $\mathcal{H}^u = \{(i_k^u, t_k^u) \mid i_k^u \in \mathcal{I}, t_k^u \in \mathcal{R}^+, k = 1, 2, \dots, N_u\}$ , where  $k$  denotes the user’s action and  $(i_k^u, t_k^u)$  denotes the interaction between user  $u$  and item  $i_k^u$  at time  $t_k^u$ ; interaction has different forms in different services, such as clicking, browsing, and adding to favorites. The objective of modeling sequential user behaviors is to predict the conditional probability of the user’s next item  $p(i_{N_u+1}^u \mid \mathcal{H}^u, t_{N_u+1}^u)$  for a certain given user  $u$ .

We take RNN as the basic model, which generates the conditional probability in multiple steps sequentially. At step  $k$ , the  $k$ -th item  $i_k^u$  is vectorized into  $x_k$  and then fed into RNN units by nonlinear transformation, e.g., multilayer perceptron. Then, it updates the hidden state of RNN units, i.e.,  $h_k = \text{RNN}(x_k, h_{k-1})$ , as well as the output of RNN units. The representations of hidden state and output are trained to predict the next item vectorized as  $x_{k+1}$  given  $h_k$ . To train the RNN, we aim to maximize the likelihood of historical behaviors of a set of users  $\mathcal{U}$ :

$$p(\mathcal{U}) = \prod_u \prod_{k=1}^{N_u-1} p(i_{k+1}^u \mid \mathcal{H}_k^u, t_{k+1}^u) = \prod_u \prod_{k=1}^{N_u-1} p(i_{k+1}^u \mid h_k^u, t_{k+1}^u), \quad (1)$$

where  $i_{k+1}^u$  is the target item for the given user  $u$ . In other words, we aim to minimize the negative logarithmic likelihood, that is, the objective function:

$$\min_{\theta} \mathcal{L}(\mathcal{U}) = - \sum_u \sum_{k=1}^{N_u-1} \log p(i_{k+1}^u \mid h_k^u, t_{k+1}^u), \quad (2)$$

where  $\theta$  is the set of parameters in the RNN model.

To fulfill this learning, it requires us to design an effective RNN architecture including the inner functions of RNN cells and the overall network structure, which together approximate a highly nonlinear function for obtaining the probability distribution of next item. In this process, RNN usually suffers from the complex dependency problem, especially when we deal with user actions that have much noise and randomness. Attention mechanism is a possible solution, which constructs a pooling layer on top of the RNN cells at each step to characterize the dependence between the current intent and all of the historical actions. We will describe our designed network architecture with attention mechanism in the next section.

#### 4. ALI-GRU

As illustrated in the left part of Figure 2, our designed network features an attention mechanism with long-term interval-based gated recurrent units for modeling sequential user behaviors. This network architecture takes the sequence of items as raw signal. There are four stages in our network. The embedding layer maps items to a vector space to extract their basic features. The bidirectional GRU layer is designed to capture the information of both long-term preferences and short-term intents of user; it consists of normal GRUs and time interval-based GRUs (see Figure 3). The attention function layer reflects our carefully designed attention mechanism, which is illustrated in the right part of Figure 2. Finally, there is an output layer to integrate the attention distribution and the extracted sequential features and utilize normal GRUs to predict the conditional probability of next item.

**4.1. Embedding Layer.** The purpose of the embedding layer is to map the raw data of items into a rectified vector space, where the vectorized representations of items still keep the semantics of the items; e.g., semantically relevant items have small distance in the vector space. Usually items can be first represented as one-hot vectors and then processed by several fully connected layers [52]. If however the number of items is too large, pretrained encoding network is useful to process the items, which encodes not only the basic properties such as category of items but also the crowdsourcing properties such as sales of items [4]. In this paper, we adopt these two strategies for different datasets, respectively.

**4.2. Bidirectional GRU Layer with Time-GRU.** This layer is designed to extract driven signals from input sequence and to refine the long-term memory by contextual information. We now detail our method for these two targets.

In the previous work for natural language processing tasks, the attention function is driven by a single vector of input [1, 27, 40]. That model works well because of the relatively stable syntax and semantics of input words. However, sequential user behaviors contain much noise and randomness that make the simple model problematic. We propose a new network structure with time-GRU to extract

short-term dynamics of user intents as driven signal of the attention function.

The structure of time-GRU in comparison with normal GRU is shown in Figure 3, where the black lines denote the network links of normal GRU and the red lines denote new links of time-GRU. The normal GRU are as follows:

$$z_N = \sigma(W_z I_N + U_z h_{N-1} + b_z), \quad (3)$$

$$r_N = \sigma(W_r I_N + U_r h_{N-1} + b_r), \quad (4)$$

$$\tilde{h}_N = \tanh(W_h I_N + r_N \odot (U_h h_{N-1}) + b_h), \quad (5)$$

$$h_N = (1 - z_N) \odot h_{N-1} + z_N \odot \tilde{h}_N, \quad (6)$$

where  $I_N$  denotes the  $N$ -th sequence item vector.  $h_{N-1}$  denotes the  $(N-1)$ -th hidden state vector.  $\tilde{h}_N$  is the candidate activation.  $z_N$  represents the update gate, which decides how much the unit updates its activation.  $r_N$  is the reset gate to control how much the last state contributes to the current activation.  $\sigma$  represents the sigmoidal nonlinearities function and  $\tanh$  represents the tanh nonlinearities function, and  $\odot$  is an element-wise multiplication. Weight parameters  $W_z, W_h, W_r$  and  $U_z, U_h, U_r$  connect different inputs and gates; parameters  $b_z, b_h, b_r$  are biases.

The above equations imply that normal GRU is good at capturing the general sequential information. Since GRU is originally designed for NLP tasks, there is no consideration of time intervals within inputs, which are very important for modeling sequential user behaviors. To include the short-term information, we augment the normal GRU with a time gate  $T_N$ :

$$T_N = \sigma(W_t \Delta t_N + U_t I_N + b_t) \quad (7)$$

s.t.  $W_t < 0$ ,

where  $\Delta t_N$  is the time interval between adjacent actions. The constraint  $W_t < 0$  is to utilize the simple assumption that smaller time interval indicates larger correlation. Moreover, we generate a time-dependent hidden state  $h_N^t$  in addition to the normal hidden state  $h_N$ ; that is,

$$h_N^t = (1 - z_N \odot T_N) \odot h_{N-1}^t + z_N \odot T_N \odot \tilde{h}_N^t, \quad (8)$$

where we utilize the time gate as a filter to modify the update gate  $z_N$  so as to capture short-term information more effectively.

In addition, we want to utilize contextual information to extract long-term information with as little information loss as possible. Recent methods usually construct a bidirectional RNN and add or concatenate two output vectors (forward and backward) of bidirectional RNN. Bidirectional RNN outperforms unidirectional one but still suffers from the embedding loss, since the temporal dynamics are not considered enough. On the contrary, we propose to combine the output of forward normal GRU ( $h_N$  in equation (6)) with all the outputs of backward GRU at different steps (the output of backward GRU at step  $k$  is denoted by  $\overleftarrow{h}_k$  in Figure 2). Specifically, we produce concatenated vectors  $[h_{N-1}, \overleftarrow{h}_{N-1}, [h_{N-1}, \overleftarrow{h}_{N-2}], \dots, [h_{N-1}, \overleftarrow{h}_1]]$ , as shown in the

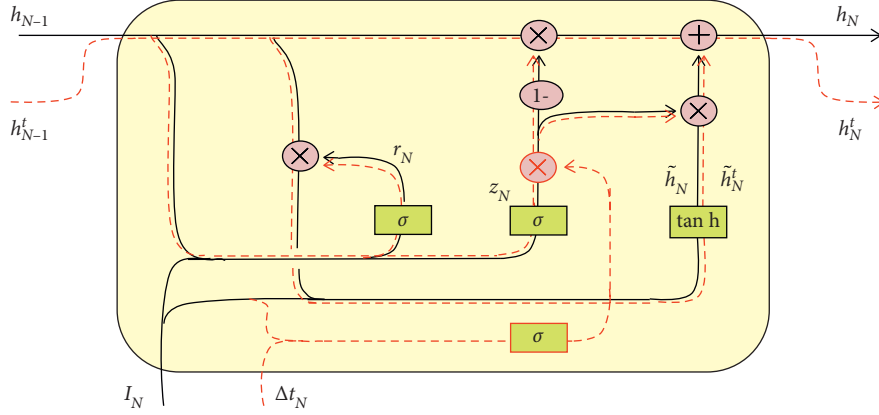


FIGURE 3: The structure of our proposed time-GRU. In addition to the network links of normal GRU (shown in black lines), we design network links and a new gate to capture time interval information (shown in red-dotted lines).

right part of Figure 2, where  $[,]$  stands for concatenation of vectors. This design effectively captures the contextual information as much as possible.

**4.3. Attention Function Layer.** Attention function layer is responsible for linking and analyzing the dependence and contribution over driven signals and contextual long-term information provided by the previous layers. Unlike previous attention mechanisms, we do not simply summarize the contextual long-term information into individual feature vectors, e.g., using  $p(i_1, h_{N-1})$  for calculating the attention weight of item  $i_1$  to the hidden state at the  $(N-1)$ -th step [46]. Instead, we design to attend to the driven signals at each time step along with the embedding of contextual information.

Specifically, as shown in the right part of Figure 2 and as already discussed in the last subsection, we use  $\mathbf{H}_k = [h_{N-1}, \overleftarrow{h}_k] \in \mathcal{R}^{2d}$ ,  $k = 1, 2, \dots, N-1$ , where  $d$  is the dimension of GRU states, to represent the contextual long-term information.  $h_k^t \in \mathcal{R}^d$  denotes the short-term intent reflected by item  $i_k$ . We then construct an attention matrix  $A \in \mathcal{R}^{(N-1) \times (N-1)}$ , whose elements are calculated by

$$A_{ij} = \alpha(\mathbf{H}_i, h_j^t) \in \mathcal{R}, \quad (9)$$

where the attention weight,

$$\alpha(\mathbf{H}_i, h_j^t) = v^T \tanh(W_a \mathbf{H}_i + U_a h_j^t), \quad (10)$$

is adopted to encode the two input vectors.  $v^T$  are the weight parameters. There is a pooling layer, for example, average or max pooling, along the direction of long-term information, and then there is a Softmax layer to normalize the attention weights of each driven signal. Let  $a_k$  be the normalized weight on  $h_k^t$ ; then the attended short-term intent vector is  $\hat{h}_k = a_k h_k^t \in \mathcal{R}^d$ . At last, we use  $g(i_k, \hat{h}_k) = [i_k, \hat{h}_k, |i_k - \hat{h}_k|, i_k \odot \hat{h}_k] \in \mathcal{R}^{4d}$  as the output to the next layer, where  $i_k$  is the embedded vector of the item at the  $k$ -th step.

We want to emphasize the insight of our carefully designed attention mechanism described above, which is different from the existing methods, to reduce the loss of

contextual information caused by early summarization. Furthermore, since driven signals are attended to the long-term information at different steps, the attentions can obtain the trending change of user's preferences, being more robust and less affected by the noise in the historical actions.

**4.4. Output Layer.** Given  $g(i_k, \hat{h}_k)$  produced by the attention function layer, we use a layer of normal GRUs to produce the embedding vector ( $h_k^o$  in Figure 2), which is expected to contain the contextual long-term information about all of the user's historical actions with respect to the single item and short-term intents. The embedding vector is then decoded to produce the final result. For example, we use a Softmax function after a fully connected layer to obtain the probability distribution of different items in the next action:

$$p(i_N) = \text{softmax}(w_p^T h_{N-1}^o). \quad (11)$$

If the number of candidate items is too big, we shall use a slightly different decoding function, which will be detailed in Section 5.3.

## 5. Experiments

In this section, we first describe the used datasets and several state-of-the-art approaches that were compared as baselines in this paper. Then, we report and discuss experimental results on different datasets.

**5.1. Datasets.** To verify our proposed ALI-GRU, we conduct a series of experiments on two well-known public datasets (LastFM (<http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/lastfm-1K.html>) and CiteULike (<http://www.citeulike.org/faq/data.adp>)). Additionally, we also perform offline and online experiments on the real data from one of the largest online shopping websites. Table 1 shows the statistics of LastFM and CiteULike:

- (i) LastFM contains  $\langle user\_id, timestamp, artist\_id, song\_id \rangle$  tuples collected from Last.fm API (<https://www.last.fm/api/>). It represents the whole listening

TABLE 1: Statistics of LastFM and CiteULike.

	LastFM	CiteULike
Number of users	987	1625
Number of items	5,000	5,000
Number of behaviors	818,767	35,834

habits (till 5 May 2009) for 1000 users. We extract tuples  $\langle user\_id, song\_id, timestamp \rangle$  from the original dataset to conduct experiments, where each  $song\_id$  represents an item and each tuple represents the action or behavior that the user  $user\_id$  listens to the song  $song\_id$  at time  $timestamp$ .

- (ii) CiteULike consists of the tuples  $\langle user\_id, paper\_id, timestamp, tag \rangle$ , where each tuple represents that the user  $user\_id$  annotates the paper  $paper\_id$  with  $tag$  at time  $timestamp$ . One user annotating one research paper (i.e., item) at a certain time may have several records, in order to distinguish different tags. We merge them as one record and extract tuples  $\langle user\_id, paper\_id, timestamp \rangle$  to construct dataset like in [52].

**5.2. Compared Approaches.** We compare ALI-GRU with the following state-of-the-art approaches for performance evaluation:

- (i) *Factorized Sequential Prediction with Item Similarity Models (Fossil)* [11]. This is a state-of-the-art factorized sequential prediction method based on Markov processes. Fossil also considers the similarity of explored items to those already consumed/liked by user, which achieves a certain success to handle the long-tail problem. We have used the implementation provided by the authors (<https://drive.google.com/file/d/0B9Ck8jw-TZUEeEhSWXU2WWloc0k/view>).
- (ii) *Basic GRU/Basic LSTM* [7]. This method directly uses normal GRU/LSTM as the primary network. For fair comparison, we set the network to use the same embedding layer and the same decoding function as our method.
- (iii) *Session RNN* [13]. Hidasi et al. propose an RNN-based method to capture the contextual information according to sessions of user behaviors. In our experiments, we use a commonly adopted method described in [16] to identify sessions so as to adopt this baseline.
- (iv) *Time-LSTM* [52]. This method utilizes LSTM to model the pattern of sequential user behaviors. Compared to normal LSTM units, Time-LSTM considers the time intervals within sequential signal and designs several time gates in LSTM units similarly as our time-GRU.
- (v) *Simplified Version 1 (SV1)*. This approach is designed to verify the effectiveness of our designed attention mechanism. The SV1 approach is identical

to ALI-GRU, with the only difference being that SV1 uses an attention mechanism provided in [1, 46] which simply summarizes the contextual long-term information into individual feature vectors. Specifically, user contextual behaviors are modeled as a vector  $[h_{N-1}, \bar{h}_1]$ , and all driven signals are attended to this vector.

- (vi) *Simplified Version 2 (SV2)*. This approach is designed to verify the effectiveness of our proposed time-GRU for generating driven signals according to short-term information. Compared to ALI-GRU, the only difference is that SV2 uses single item at each step (the embedded vector) to attend to contextual information.

All RNN-based models are implemented with the open source deep learning platform TensorFlow (<https://www.tensorflow.org/>). Training was done on a single GeForce Tesla P40 GPU with 8 GB graphical memory.

**5.3. Experiments on LastFM and CiteULike.** We first evaluate our method on two well-known public datasets for the task of sequential recommendation.

**5.3.1. Datasets.** In this experiment, we use the same datasets as those adopted in [52], i.e., LastFM and CiteULike. Table 1 presents the statistics of these two datasets. Both datasets can be formulated as a series of tuples  $\langle user\_id, item\_id, timestamp \rangle$ . Our target is to recommend songs in LastFM and papers in CiteULike for users according to their historical behaviors.

For fair comparison, we follow the segmentation of training set and test set as described in [52]. Specifically, 80% users are randomly selected for training. The remaining users are for testing. For each test user  $u$  with  $k$  historical behaviors, there are  $k - 1$  test cases, where the  $k$ -th test case is to perform recommendations at time  $t_{k+1}^u$  given the user's previous  $k$  actions, and the ground-truth is  $i_{k+1}^u$ . The recommendation can also be regarded as a multiclass classification problem. For more details, please refer to [52].

**5.3.2. Implementation.** Following the method in [52], we use one-hot representations of items as inputs to the network and one fully connected layer with 8 nodes for embedding. The length of hidden states of GRU-related layers including both normal GRU and time-GRU is 16. A Softmax function is used to generate the probability prediction of next items. For training, we use the AdaGrad [8] optimizer, which is a variant of Stochastic Gradient Descent (SGD). Parameters for training are minibatch size of 16 and initial learning rate of 0.001 for all layers. The training process takes about 8 hours.

**5.3.3. Evaluations.** In the test stage, following the evaluation method in [52], we select 10 items with top probabilities as final recommendations. We use Recall@10 to measure

whether the ground-truth item is in the recommendation list. Recall@10 is defined as

$$\text{Recall@10} = \frac{n_{\text{hit}}}{n_{\text{testcase}}}, \quad (12)$$

where  $n_{\text{hit}}$  is the number of test cases where  $i_g$  is in the recommendation list and  $n_{\text{testcase}}$  is the number of all test cases. We further use MRR@10 (Mean Reciprocal Rank) to consider the rank of ground truth in the recommendation list. This is the average of reciprocal ranks of  $i_g$  in the recommendation list. The reciprocal rank is set to 0 if the rank is above 10.

**5.3.4. Overall Performance.** The results of sequential recommendation tasks on LastFM and CiteULike are shown in Table 2. It can be observed that our approach performs the best on both LastFM and CiteULike for all metrics, which demonstrates the effectiveness of our proposed ALI-GRU. Specifically, ALI-GRU obtains significant improvement over Time-LSTM, which is the best baseline, averagely by 4.70% and 6.55% for Recall@10 and MRR@10, respectively. It owes to the superiority of introducing attention mechanism into RNN-based methods, especially in capturing the contribution of each historical action.

**5.3.5. Performance of Cold-Start.** Cold-start refers to the lack of enough historical data for a specific user, which often decreases the efficiency of making recommendations. We analyze the influence of cold-start on the LastFM dataset and the results are given in Figure 4. In this figure, test cases are separately counted for different numbers of historical actions, and small number refers to cold-start. We can observe that, for cold users with only 5 actions, ALI-GRU performs slightly worse than the state-of-the-art methods. This is because ALI-GRU considers short-term information as driven signals, which averages source signal to some extent and leads to less accurate modeling for cold users. Along with the increase of historical actions, ALI-GRU achieves significantly better performance than the baselines, which indicates that bidirectional GRU and attention mechanism can better model the long-term preferences for making recommendations.

**5.4. Offline Experiments.** We have collected a large-scale dataset from a real-world e-commerce website for further performance evaluation. ALI-GRU is also adopted online and results of online A/B test will be reported in the next section.

**5.4.1. Dataset.** User behaviors of this dataset are randomly sampled from the logs of clicking and purchasing in seven days (a beginning week of July 2017) on a real-world e-commerce website. The dataset is again formulated as a series of tuples  $\langle \text{user\_id}, \text{item\_id}, \text{timestamp} \rangle$ .

We focus on the task of the personalized search of e-commerce websites. So we define positive cases as those

TABLE 2: Comparison results on LastFM and CiteULike.

	LastFM		CiteULike	
	Recall@10	MRR@10	Recall@10	MRR@10
Basic-LSTM	0.2451	0.0892	0.6824	0.2889
Session RNN	0.3405	0.1573	0.7129	0.2997
Time-LSTM	0.3990	0.2657	0.7586	0.3660
ALI-GRU	<b>0.4752</b>	<b>0.2697</b>	<b>0.7764</b>	<b>0.4930</b>

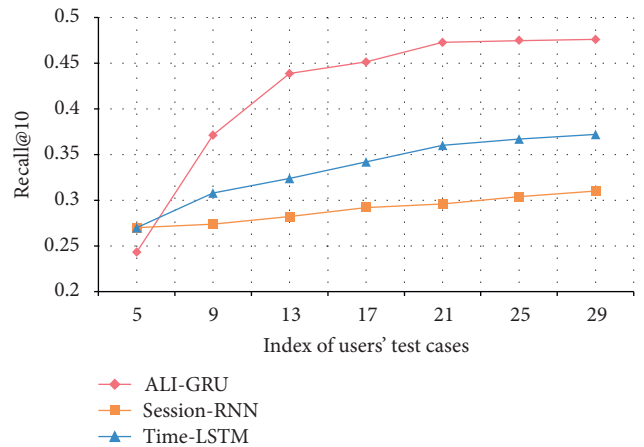


FIGURE 4: Recall@10 evaluated on different indexes of users' test cases in LastFM.

purchasing behaviors led by the e-commerce search engine mentioned above, while negative cases are those clicks without purchases (if there is no purchase around the click within 5 actions). Finally, we have 24,282,032 positive cases, 84,322,922 negative cases, 30,602,427 users, and 10,808,463 items. We randomly select 80% users for training, and the remaining users are for testing. For each positive or negative case  $i_k$  in the sequence, our target is to predict whether the user would purchase  $i_k$  according to his/her historical behaviors, which is a typical binary classification problem.

**5.4.2. Implementation.** Since the amount of items is too huge in this dataset, it is inconvenient to employ the one-hot representations as inputs to RNN-based models. Instead, we use pretrained embedding vectors of items as inputs and additionally use two fully connected layers, both with 128 nodes, to reembed the item vectors. Also, we have followed the wide & deep learning approach in [4] to convert the outputs of the final fully connected layer, whose size is 48, into representations of corresponding items. For fair comparison, all RNN-based approaches employ the pre-trained item representations as inputs. Hidden state size of GRU-related layers is 128. We finally use the sigmoid function to predict whether the user would purchase  $i_k$ . For training, the loss function is cross-entropy and AdaGrad optimizer is employed with a minibatch size of 256 and an initial learning rate of 0.001 for all layers. The entire training process takes about 50 hours.

TABLE 3: Comparison results on offline dataset.

	Fossil	Basic-GRU	Session RNN	Time-LSTM	SV1	SV2	ALI-GRU
AUC	0.7017	0.6609	0.7543	0.7712	0.7833	0.8062	<b>0.8370</b>

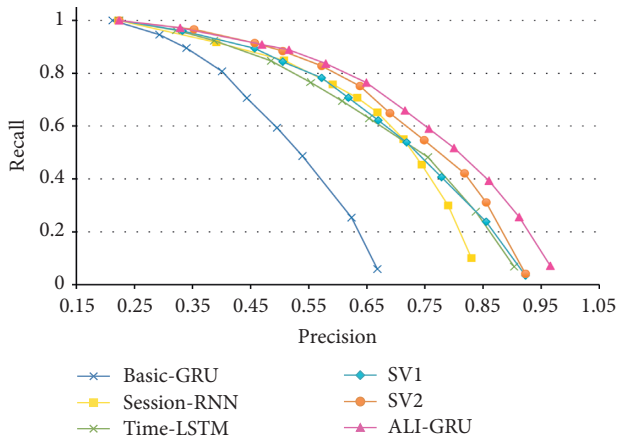


FIGURE 5: Precision-Recall curves of different methods on offline dataset.

5.4.3. *Evaluations.* In the test stage, we use Precision-Recall of positive cases to measure the performance. Moreover, AUC (Area under ROC Curve) is also adopted, which is widely used in imbalanced classification tasks [12]. The larger the value of AUC is, the better the performance is.

5.4.4. *Overall Performance.* Table 3 shows the AUC results for measuring the overall performance. We can observe that all RNN-based methods except Basic-GRU outperform Fossil that is based on matrix factorization with Markov processes, which indicates the advantage of RNN for modeling sequential data. Furthermore, for different views of the capabilities of different RNN-based approaches, we also report Precision-Recall curves shown in Figure 5 and make some comparisons and summarize our findings as follows.

5.4.5. *Basic GRU versus Session RNN versus Time-LSTM.* Session RNN and Time-LSTM achieve significant improvement compared to Basic GRU, which is consistent with the previous results on public datasets. This is due to the limitation of Basic-GRU/Basic-LSTM in modeling complex long-term sequential data. Compared to Session RNN, Time-LSTM achieves better performance for high precision range (precision larger than about 0.73), which owes to the advantage of short-term intents for predicting highly confident items. On the contrary, Session RNN outperforms Time-LSTM for low precision range (precision lower than about 0.73), since Session RNN introduces session view to better model contextual information and then benefits from recalling items based on long-term preferences of user.

5.4.6. *Session RNN versus Time-LSTM versus ALI-GRU.* By adopting time gate, which is strong at modeling short-term dynamics, and bidirectional RNN, which leads to advantages for modeling long-term information, ALI-GRU

better analyzes the complex dependence among items and user intents, together with a novel matrix-form attention mechanism to enhance the performance. ALI-GRU outperforms Session RNN and Time-LSTM by as much as 10.96% and 8.53% for AUC (Table 3), respectively. Observing Precision-Recall curves, we found that ALI-GRU beats Session RNN and Time-LSTM over the entire range, and the improvement is more significant for the high precision range. The superior performance of ALI-GRU on various datasets and views demonstrates its efficacy to handle long-term sequential user behaviors with dynamical short-term intents.

5.4.7. *SV1 and SV2 versus Others.* We also show results of SV1 and SV2 for ablation analyses (Figure 5). Observing the curve of SV1, we can find that the previous attention mechanism with bidirectional GRU only achieves slight improvement compared to Time-LSTM, which indicates the limitation of the previously studied attention mechanism for capturing dynamical importance of items in sequential user behaviors. On the contrary, SV2 outperforms both Session RNN and Time-LSTM consistently, especially for low precision range. It suggests that our proposed matrix-form attention mechanism with bidirectional GRU has superior capacity in distinguishing items’ importance for modeling long-term preferences of user. Nevertheless, the curve of SV2 drops a lot when the precision is larger than about 0.82, where it is comparable to SV1 and Time-LSTM. This is because user behaviors and intents are dynamic with a certain randomness, and single item is not robust enough for calculating attention distribution and capturing short-term intents. Last but not least, we can find that ALI-GRU leads to performance boost against SV1 and SV2 consistently. It demonstrates the advantages of our carefully designed matrix-form attention with long-term interval-based GRU framework for modeling sequential user behaviors.

5.4.8. *Case Study and Insights.* We present three cases in Figure 6 for comprehensive study to give some insights of our proposed approach. Each case consists of one user’s historical items ordered by click time and shows the attention heat map for each item and the final item (click or purchase) with prediction and ground truth.

*Case A.* The user clicked items of several classes such as watch, handbag, and dress and finally purchased the watch that was clicked long before. We make a few observations as follows: (1) ALI-GRU gives higher weights to most watches than other items, which is consistent with the user’s contextual intent (purchasing a watch). It suggests that our proposed approach has capacity to capture user’s real intents from historical behaviors. (2) The 1st, 3rd, and 5th watches, which are the same as or similar to the final purchased watch, have higher weights than the other watches, especially for the



FIGURE 6: Case study of making prediction whether the user purchases the next item according to historical sequential behaviors. Each case consists of one user’s clicked items ordered by click time and shows the attention heat map for each item and the final item (clicked or purchased) with prediction and ground truth. Darker color in the heat map indicates higher attention weight.

1st watch, though it was clicked earliest for a long time ago. More interestingly, we can observe that the 6th watch is for women, and the user is probably a woman (according to the dresses he/she clicked), but the 6th watch has the relative lowest weight in all watches. These observations indicate that ALI-GRU successfully distinguishes the user’s current intent for purchasing a watch for men.

*Case B.* If items were inherently important or repeated or had low frequency, models without attention mechanism might work well since such model could automatically assign low weights to irrelevant items and vice versa. However, the importance of items and user intents is highly dependent on context and is consistent to a certain degree. In Case B, the user finally purchased a coat hanger that belongs to the class he/she never clicked. Nevertheless, ALI-GRU looks at the context of his/her recent behaviors, conjectures the intent is possible for something about laundry, and correctly figures out this is a positive case.

*Case C.* It is not a correct prediction case according to ground truth. Observing the historical behaviors and attention distribution of the user, we can find that ALI-GRU chooses to ignore various items before the first suit; these actions had a long time interval (about two days) from the latter actions. Furthermore, ALI-GRU conjectures that the user wants to buy something for formal wearing. Therefore, ALI-GRU predicts it is a negative case for purchasing a USB cable, which is purchased by the user finally. In such cases, there exist choppy and decisive intents of user, which is a great challenge left for future exploration.

**5.5. Online Test.** Online test with real-world e-commerce users is carried out to study the effectiveness of our proposed method. In particular, we integrate ALI-GRU into the e-commerce search engine mentioned above, which has billions of clicks per day. A standard A/B test is conducted online. Users of the search engine are randomly divided into multiple buckets, and we randomly select two buckets for experiments. For users in bucket A, we use the existing

highly optimized ranking solution of the search engine, which performs Learning to Rank (LTR) and Reinforce Learning (RL) with several effective algorithms such as wide & deep learning and CF prediction. For users in bucket B, we further integrate the results produced by ALI-GRU. Specifically, for a given user, his/her sequential behaviors (clicked items and timestamps) are collected from the entire service, and the user’s intent vector is predicted by ALI-GRU in real time. When the user provides a query, we combine the calculated user intent vector with all the retrieved items to calculate purchasing probability, which is similar to the method of offline experiments. Finally, we integrate the purchasing probability into the existing ranking strategy.

Measures for the online A/B test include Gross Merchandise Volume (GMV), user Click Through Rate (uCTR), Click Conversion Rate (CVR), Per Customer Transaction (PCT), and Unique Visitor Value (UV\_Value), which are all frequently used metrics in e-commerce [19]:

$$\begin{aligned}
 \text{uCTR} &= \frac{\text{number of user clicking}}{\text{number of user browsed items}}, \\
 \text{CVR} &= \frac{\text{number of user trading}}{\text{number of user clicking}}, \\
 \text{PCT} &= \frac{\text{gross merchandise volume}}{\text{number of user trading}}, \\
 \text{UV\_Value} &= \frac{\text{gross merchandise volume}}{\text{number of unique visitor}}.
 \end{aligned} \tag{13}$$

The test was performed within one week in July 2017. Comparative results are given in Table 4, where the absolute values are omitted for business confidentiality. The results show that ALI-GRU achieves better performance for all the metrics. As we expected, uCTR and CVR are improved, which means that users are more likely to click the reranked items, and there is higher probability to purchase these

TABLE 4: The improvements (%) of online test.

GMV	uCTR	CVR	PCT	UV_value
+3.08	+1.71	+1.28	+1.29	+2.57

items. More interesting is the improvement of PCT and UV\_Value, which is due to the increase of number of transactions per user with purchasing actions. This result suggests that our model provides kinds of recommendation functionality into search engine, such as case B in Figure 6. In summary, our proposed ALI-GRU consistently improves the high-quality baseline of one of the largest online e-commerce platforms that has been optimized for several years. Such improvements are very important for e-commerce search engine systems and have significant business value. ALI-GRU has been adopted into the search engine before this paper is prepared.

## 6. Conclusions

Modeling user behaviors as sequential learning plays an important role for predicting future user actions, such as personalized search and recommendation. However, most of RNN-based methods assume that the importance of historical behaviors decreases over time and fail to consider the cross-dependence in sequences, which makes it difficult to apply to the real-world scenarios. To address these problems, we propose a novel and efficient approach called Attention with Long-term Interval-based Gated Recurrent Units (ALI-GRU) for better modeling sequential user behaviors. We first propose a bidirectional time interval-based GRU to identify complex correlation between actions and capture both long-term preferences and short-term intents of users as driven signals. Then, we design a new attention mechanism to attend to the driven signals at each time step for predicting the next user action. The empirical evaluations on two public datasets for sequential recommendation task show that ALI-GRU achieves better performance than state-of-the-art solutions. Specifically, ALI-GRU outperforms Session RNN and Time-LSTM by as much as 10.96% and 8.53% in terms of AUC. In addition, online A/B tests in a real-world e-commerce search engine further demonstrate its practical value. As GRU cannot be calculated in parallel, it takes a lot of time to train the model. In the future, we will adopt a parallel approach to solve this problem.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61873288), Alibaba-PKU Joint Program, and Zhejiang Lab (nos. 2019KE0AB01 and 2019KB0AB06).

## References

- [1] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, <https://arxiv.org/abs/1409.0473>.
- [2] J. Chen, H. Zhang, X. He, L. Nie, W. Liu, and T.-S. Chua, "Attentive collaborative filtering: multimedia recommendation with item-and component-level attention," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 335–344, Tokyo, Japan, August 2017.
- [3] H. Cheng, H. Fang, X. He, J. Gao, and Li Deng, "Bi-directional attention with agreement for dependency parsing," 2016, <https://arxiv.org/abs/1608.02076>.
- [4] H.-T. Cheng, L. Koc, J. Harmsen et al., "Wide & deep learning for recommender systems," 2016, <https://arxiv.org/abs/1606.07792>.
- [5] J. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, "Attention-based models for speech recognition," 2015, <https://arxiv.org/abs/1506.07503>.
- [6] C. Chu, L. Zhao, B. Xin et al., "Deep graph embedding for ranking optimization in e-commerce," in *Proceedings of the ACM International Conference on Information and Knowledge Management*, Turin, Italy, 2018.
- [7] J. Chung, C. Gulcehre, K. H. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, <https://arxiv.org/abs/1412.3555>.
- [8] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, pp. 2121–2159, 2011.
- [9] L. Guo, Q. Zhang, W. Hu, Z. Sun, and Y. Qu, "Learning to complete knowledge graphs with deep sequential models," *Data Intelligence*, vol. 1, no. 3, pp. 224–2243, 2019.
- [10] R. He, W.-C. Kang, and J. McAuley, "Translation-based recommendation: a scalable method for modeling sequential behavior," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 5264–5268, Stockholm, Sweden, 2018.
- [11] R. He and J. McAuley, "Fusing similarity models with Markov chains for sparse sequential recommendation," in *Proceedings of the International Conference on Data Mining*, pp. 191–200, Barcelona, Spain, 2016.
- [12] R. He and J. McAuley, "VBPR: visual bayesian personalized ranking from implicit feedback," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 144–150, Phoenix, AZ, USA, 2016.
- [13] B. Hidasi, A. Karatzoglou, L. Baltrunas, and T. Domonkos, "Session-based recommendations with recurrent neural networks," 2015, <https://arxiv.org/abs/1511.06939>.
- [14] B. Hidasi, M. Quadrana, A. Karatzoglou, and T. Domonkos, "Parallel recurrent neural network architectures for feature-rich session-based recommendations," in *Proceedings of the ACM Conference on Recommender Systems*, pp. 241–248, Boston, MA, USA, 2016.
- [15] B. Hidasi and T. Domonkos, "Fast ALS-based tensor factorization for context-aware recommendation from implicit feedback," in *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 67–82, Bristol, UK, September 2012.
- [16] X. Huang, F. Peng, A. An, and S. Dale, "Dynamic web log session identification with statistical language models," *Journal of the American Society for Information Science and Technology*, vol. 55, no. 14, pp. 1290–1303, 2004.



- [17] R. Hübner, M. Steinhauser, and C. Lehle, "A dual-stage two-phase model of selective attention," *Psychological Review*, vol. 117, no. 3, pp. 759–784, 2010.
- [18] S. Ji, S. Pan, E. Cambria, P. Marttinen, and S. Y. Philip, "A survey on knowledge graphs: representation, acquisition and applications," 2020, <https://arxiv.org/abs/2002.00388>.
- [19] P. Jiang, Y. Zhu, Yi Zhang, and Y. Quan, "Life-stage prediction for product recommendation in e-commerce," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2015.
- [20] S. Kabbur, X. Ning, and K. George, "FISM: factored item similarity models for top-n recommender systems," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 659–667, Chicago, IL, USA, 2013.
- [21] Y. Koren, "Collaborative filtering with temporal dynamics," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 447–456, Paris France, June 2009.
- [22] C. Lei, L. Dong, W. Li, Z.-J. Zha, and H. Li, "Comparative deep learning of hybrid representations for image recommendations," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2545–2553, Las Vegas, NV, USA, 2016.
- [23] S. Li, T. Cai, Ke Deng, X. Wang, T. Sellis, and F. Xia, "Community-diversified influence maximization in social networks," *Information Systems, Amsterdam, The Netherlands*, vol. 92, Article ID 101522, 2020.
- [24] C. Liu, J. Mao, F. Sha, and A. Yuille, "Attention correctness in neural image captioning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 4176–4182, San Francisco, CA, USA, 2017.
- [25] Q. Liu, S. Wu, D. Wang, Z. Li, and L. Wang, "Context-aware sequential recommendation," in *Proceedings of the IEEE 16th International Conference on Data Mining (ICDM)*, pp. 1053–1058, Barcelona, Spain, 2016.
- [26] Q. Liu, S. Wu, D. Wang, Z. Li, and L. Wang, "Context-aware sequential recommendation," *IEEE International Conference on Data Mining*, pp. 1053–1058, 2016.
- [27] M.-T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," 2015, <https://arxiv.org/abs/1508.04025>.
- [28] K. Miyahara and M. J. Pazzani, "Collaborative filtering with the simple Bayesian classifier," in *Proceedings of the Pacific Rim International Conference on Artificial Intelligence*, pp. 679–689, Melbourne, Australia, August 2000.
- [29] D. Neil, M. Pfeiffer, and S.-C. Liu, "Phased LSTM: accelerating recurrent network training for long or event-based sequences," 2016, <https://arxiv.org/abs/1610.09513>.
- [30] S. Pan, R. Hu, S.-F. Fung, G. Long, J. Jiang, and C. Zhang, "Learning graph embedding with adversarial training methods," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2475–2487, 2020.
- [31] A. Paterek, "Improving regularized singular value decomposition for collaborative filtering," *Proceedings of KDD Cup and Workshop*, vol. 2007, pp. 5–8, 2007.
- [32] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, "Factorizing personalized Markov chains for next-basket recommendation," in *Proceedings of the International Conference on World Wide Web*, pp. 811–820, Raleigh, CA, USA, 2010.
- [33] B. Sarwar, K. George, K. Joseph, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the tenth international conference on World Wide Web*, pp. 285–295, Hong Kong, China, 2001.
- [34] M. Seo, A. Kembhavi, F. Ali, and H. Hajishirzi, "Bidirectional attention flow for machine comprehension," 2016, <https://arxiv.org/abs/1611.01603>.
- [35] T. Shen, T. Zhou, G. Long, J. Jiang, S. Pan, and C. Zhang, "DiSAN: directional self-attention network for RNN/CNN-free language understanding," 2018, <https://arxiv.org/abs/1709.04696>.
- [36] X. Shen, S. Pan, W. Liu, Y.-S. Ong, and Q.-S. Sun, "Discrete network embedding," in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 3549–3555, Stockholm, Sweden, July 2018.
- [37] W. Song, Z. Xiao, Y. Wang, L. Charlin, M. Zhang, and J. Tang, "Session-based social recommendation via dynamic graph attention networks," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pp. 555–563, Melbourne, Australia, 2019.
- [38] S. Sukhbaatar, J. Weston, R. Fergus et al., "End-to-end memory networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 2440–2448, Montreal, Canada, 2015.
- [39] B. Vassøy, M. Ruocco, E. d. s. d. Silva, and E. Aune, "Time is of the essence: a joint hierarchical RNN and point process model for time and item predictions," in *Proceedings of the ACM International Conference on Web Search and Data Mining*, pp. 591–599, Los Angeles, CA, USA, 2019.
- [40] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 5998–6008, Long Beach, CA, USA, 2017.
- [41] P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng, "Learning hierarchical representation model for nextbasket recommendation," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 403–412, Santiago, Chile, 2015.
- [42] P. Wang, L. Zhao, X. Pan, D. Ding, X. Chen, and Y. Hou, "Density matrix based preference evolution networks for e-commerce recommendation," in *Proceedings of the International Conference on Database Systems for Advanced Applications*, pp. 366–383, Chiang Mai, Thailand, 2019.
- [43] P. Wang, L. Zhao, Y. Zhang, Y. Hou, and L. Ge, "QPIN: a Quantum-inspired preference interactive network for e-commerce recommendation," in *Proceedings of the ACM International Conference on Information and Knowledge Management*, pp. 2329–2332, Beijing, China, 2019.
- [44] X. Wang, Y. Liu, and F. Xiong, "Improved personalized recommendation based on a similarity network," *Physica A: Statistical Mechanics and Its Applications*, vol. 456, pp. 271–280, 2016.
- [45] X. Wang, Y. Liu, G. Zhang, Y. Zhang, H. Chen, and J. Lu, "Mixed similarity diffusion for recommendation on bipartite networks," *IEEE Access*, vol. 5, no. 2017, pp. 21029–21038, 2017.
- [46] Y. Wang, H. Shen, S. Liu, J. Gao, and X. Cheng, "Cascade dynamics modeling with attention-based recurrent neural network," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, pp. 2985–2991, Melbourne, Australia, 2017.
- [47] C.-Y. Wu, A. Ahmed, A. Beutel, A. J. Smola, and H. Jing, "Recurrent recommender networks," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 495–503, Cambridge, UK, 2017.

- [48] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang, "Connecting the dots: multivariate time series forecasting with graph neural networks," 2020, <https://arxiv.org/abs/2005.11650>.
- [49] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1480–1489, San Diego, CA, USA, 2016.
- [50] H.-F. Yu, C.-J. Hsieh, S. Si, and I. Dhillon, "Scalable coordinate descent approaches to parallel matrix factorization for recommender systems," in *Proceedings of the IEEE 12th International Conference on Data Mining*, pp. 765–774, Brussels, Belgium, 2012.
- [51] L. Zheng, V. Noroozi, and S. Y. Philip, "Joint deep modeling of users and items using reviews for recommendation," in *Proceedings of the ACM International Conference on Web Search and Data Mining*, pp. 425–434, Cambridge, UK, 2017.
- [52] Yu Zhu, H. Li, Y. Liao et al., "What to do next: modeling user behaviors by time-LSTM," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 3602–3608, Melbourne, Australia, August 2017.

## Research Article

# Recommendation Algorithm in Double-Layer Network Based on Vector Dynamic Evolution Clustering and Attention Mechanism

Jianrui Chen <sup>1,2</sup> Zhihui Wang <sup>2</sup> Tingting Zhu <sup>2</sup> and Fernando E. Rosas <sup>3</sup>

<sup>1</sup>Key Laboratory of Modern Teaching Technology, Ministry of Education, Xi'an, China

<sup>2</sup>School of Computer Science, Shaanxi Normal University, Xi'an, China

<sup>3</sup>Data Science Institute and Department of Brain Science, Imperial College London, London, UK

Correspondence should be addressed to Jianrui Chen; [jianrui\\_chen@snnu.edu.cn](mailto:jianrui_chen@snnu.edu.cn)

Received 13 March 2020; Accepted 23 May 2020; Published 7 July 2020

Guest Editor: Liang Wang

Copyright © 2020 Jianrui Chen 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 purpose of recommendation systems is to help users find effective information quickly and conveniently and also to present the items that users are interested in. While the literature of recommendation algorithms is vast, most collaborative filtering recommendation approaches attain low recommendation accuracies and are also unable to track temporal changes of preferences. Additionally, previous differential clustering evolution processes relied on a single-layer network and used a single scalar quantity to characterise the status values of users and items. To address these limitations, this paper proposes an effective collaborative filtering recommendation algorithm based on a double-layer network. This algorithm is capable of fully exploring dynamical changes of user preference over time and integrates the user and item layers via an attention mechanism to build a double-layer network model. Experiments on Movielens, CiaoDVD, and Filmtrust datasets verify the effectiveness of our proposed algorithm. Experimental results show that our proposed algorithm can attain a better performance than other state-of-the-art algorithms.

## 1. Introduction

Information overload is a pervasive problem in our era of big data, being a consequence of the rapid development of the Internet and other information technologies. Recommendation algorithms are one of the most widespread approaches to address this problem [1], whose purpose is to help users to find information quickly and conveniently. Additionally, recommendation systems usually suggest new items using information from previous searches, including product or media recommendation [2, 3]. Recommendation systems play a key role in the digital economy, as they allow web services to improve user's experience, increase product sales, and help products to realize their commercial value.

While the literature of recommendation systems is vast, most algorithms can be classified in three categories: content-based [4], collaborative filtering [5], and hybrid recommendation systems [6]. Among these methods, collaborative filtering recommendation algorithms are the most popular in both research and industry, as they can

exploit social information better. Moreover, collaborative filtering algorithms can be further divided into three categories: memory-based [7], model-based [8], and hybrid filtering [9]. Among model-based recommendation algorithms, matrix factorization models stand out for their superior speed and strong scalability. In this literature, many matrix factorization models were proposed following the seminal work of Billsus and Pazzani [10]. Recent contributions include nonnegative matrix factorization methods for community detection [11] and dynamic networks [12]. Matrix factorization methods are particularly attractive as they can consider the influence of various factors, while bringing good performance and scalability. Unfortunately, conventional collaborative filtering approaches are not well-suited to deal with various problems related with the explosive development of information technologies, which often involve sparse data, cold start, or multidimensional data. These issues are likely to become widespread in the near future due to the continuous increase of online users and items, and hence new approaches that can address these

challenges are required. Since the community detection method can be applied to the recommendation system to find the interest communities of the target users, it can effectively achieve personalized recommendation.

On the above basis, the aim of this paper is to establish a double-layer network, apply attention mechanism to connect the double-layer network, and use the vector dynamic evolution clustering method to detect the community of users and items. The main challenges are as follows:

- (i) Constructing the double-layer network: in this paper, the recommender system is modeled as a double-layer network, and they are the user layer and item layer. Similarity between nodes is the weighted edge between nodes in their corresponding layer.
- (ii) Applying attention mechanism to connect the double-layer network: this paper puts forward a novel approach to carry out real-time recommendations, which is based on an attention mechanism and forgetting function that are used to fit scores and build relationships between users and items. The attention mechanism allows the algorithm to focus on particularly relevant factors while ignoring others, hence surpassing previous approaches based on limited scores.
- (iii) Evolving the state vectors to detect the community of users and items: as the community detection method that can be used to find users' interest communities, it can be used to effectively achieve personalized recommendation. Hence, this paper proposes a community detection procedure for users and items, which uses an evolutionary clustering method based on vectorial dynamics. This clustering procedure enables a more accurate representation of the state value than other approaches based on scalar quantities. Our approach then leverages the community structure of users and items, finding neighbor sets of target users via a Cosine similarity method.

The efficiency of our approach is confirmed via various simulation experimental results, which show that the attained similarity between attribute and rating information is better than that of other state-of-the-art approaches.

The rest of the paper is organized as follows. The state of the art and related work is given in Section 2. Then, the proposed algorithm is introduced in Section 3. Section 4 presents experimental results, and finally Section 5 summarises our main conclusions. The convergence analysis of dynamical evolution clustering method in the double-layer network is shown in Appendix.

## 2. Related Work

Contemporary recommendation algorithms usually have to deal with challenges including sparse data, cold start, or

multidimensional data. To deal with these challenges, Ling et al. proposed a recommendation algorithm for solving cold user problems by applying character capture and clustering methods [13]. Also, West et al. [14] illustrate how clustering technology can be combined with collaborative filtering to improve the recommendation performance. Importantly, as the clustering method divides users and items into several categories and then carries out collaborative filtering recommendation within the class, the recommendation time is greatly reduced. Building up on the well-known k-means clustering algorithm [15], Zahra et al. proposed the different kinds of recommendation algorithms for random selection of initial center of improved k-means clustering algorithm [16]. Because the community detection method can cluster the users with similar interest into the same community, push the users with different interest into different communities, and one can then find the nearest neighbor set in the similar interest community carrying out collaborative filtering. By performing recommendations within the community of the target user not only improves the recommendation accuracy but also reduces the complexity of the algorithm. So, community detection methods have been considered, including the community detection algorithm based on the similarity of paths proposed by Wu et al. [17], and the community detection algorithm is based on the simplification of complex network proposed by Bai et al. [18].

Other focus of study has been scenarios where networks are not static but evolve in time. In these cases, dynamic clustering algorithms are needed in order to obtain an adequate clustering effect. Wu et al. proposed a method for clustering based on dynamic synchronization [19, 20], and then we developed community detection approaches based on evolutionary clustering [21, 22]. Both methods brought similar users together while making dissimilar items distant, which improved the performance of the clustering algorithms. Bu et al. proposed a dynamic clustering strategy based on the attribute clustering graph [23]. One of the main findings of these works is that dynamical clustering methods tend to enable more efficient recommendation algorithms than traditional clustering methods.

Additionally, researchers have proposed different recommendation algorithms for various scenarios, including recommendation algorithm based on graph network models, attention networks, and multilayer networks. These are reviewed as follows:

- (i) Graph network models: graph network has a flexible topology and can express complex relationships. This method treats users and items as nodes and represents relationships as edges between them. Edges are usually weighted and might be directed or undirected. To deal with sparse data and cold start, Moradi et al. applied trust information to the collaborative filtering method by graph clustering algorithm [24]. For cases with strong time constraints, such as financial news, Ren et al. proposed a graph embedding recommendation algorithm based on a heterogeneous graph model [25]. The performance

of recommendation can be greatly improved by exploring the graph network to recommend suitable items to users. Therefore, the relationship between the user and the item is represented by a graph network in this paper, so as to dig out more information in recommendation.

- (ii) Attention mechanism: attention mechanisms are crucial in modulating the user experience and have been leveraged in various engineering applications including image processing and natural language processing. Extending previous recommendation algorithms were driven on user preferences but did not consider user attention. It is a good idea to incorporate the attention mechanism into the recommendation algorithm, which can make the recommendation algorithm more practical. Liang et al. proposed a mobile application for feature interaction through an attention mechanism [26]. At the same time, Feng et al. proposed a recommendation algorithm based on an attention network [27]. The abovementioned work has improved the performance to a certain extent by incorporating the attention mechanism into the recommendation system. Therefore, this paper uses the attention mechanism to connect the user layer and the item layer, so as to obtain a large double-layer graph network.
- (iii) Multilayer networks: while most recommendation algorithms do not consider interactions that can take place between users and items, recent works have proposed to encode them in multilayer networks. Shang et al. proposed a video recommendation algorithm based on a hyperlink multilayer graph model [28]. Yasami proposed a new knowledge-based link recommendation approach using a nonparametric multilayer model of dynamic complex networks [29]. And the methods proposed by them all improve the algorithm based on single-layer network to some extent.

### 3. Proposed Algorithm

**3.1. Scenario.** Sometimes users do not know what clothes to buy, what movies to watch, what songs to listen, and so on. At this time, users can look at the items recommended by the recommendation system for users. Recommendation system is an information filtering scheme and it is based on the user's historical behavior data. The main task of the recommendation system is to predict the rating of the items by users and recommend the items to the users. However, given user history evaluation information (evaluation, rating, and timing), user attribute information (gender, age, occupation, zip code, etc.), and item attribute information (comedy, science, war, etc.), it is a problem to recommend appropriate and interested items to users.

Suppose there are  $m$  users and  $n$  items in a recommendation system. Most recommendation methods work in two ways. On the one hand, a recommendation might be

made based on the attributes of the items that users like. For example, if a user likes comedies, then the system will recommend her to choose comedies within the available films. On the other hand, the system might look for the users who have similar preferences to a target user and recommend to the target user what similar users have chosen before.

**3.2. Algorithm Detail Description.** Here, we give the detail demonstration of our proposed method. Among them, Section 3.2.1 presents the construction of the double-layer network, and Section 3.2.2 presents the vector dynamic evolution clustering method and main convergence analysis. Furthermore, the predicted scores are obtained in each cluster, as shown in Section 3.2.3. Finally, the complete pseudocode and flow chart of our proposed algorithm are shown in Section 3.2.4.

The notations and their explanations in this paper are summarized in Table 1.

**3.2.1. Network Model Construction.** In this paper, all users and items are treated as nodes in networks, and the state of each node is represented as a vector. The user layer is set up by all users in the system, and the item layer is constructed by all items. The similarity between users is regarded as the edge weight of the user layer. The similarity between items is regarded as the edge weight of the item layer. The user and item layers are connected through attention mechanisms and ratings. In this way, the double-layer network model is formed and a simple example is shown in Figure 1.

**(1) Constructing the User Layer Network.** Everyone observes the same thing in different angles. In the same way, interests of everyone will be various. According to MovieLens dataset, from the perspective of gender, male prefers action movies and female prefers romantic movies. Based on this, to a large extent, people who have similar attribute information have similar interests and preferences. We integrate the score information into the calculation of similarity. Firstly, the ages are divided into three stages: younger than 18, 18 to 55, and older than 55. The occupations are divided into three classes: culture class, leisure class, and management class. The sexes are categorized into male and female. If the user has this attribute, it is 1; and if the user does not have this attribute, it is 0. Thus, the attribute vector of the  $i$ th user is 8-dimensional 0-1 vector, denoted as  $U_i$ .

Then, attribute similarity between users  $S_U^{\text{attribute}}$  is defined as follows:

$$S_U^{\text{attribute}}(i, j) = \frac{1}{1 + \|U_i - U_j\|}, \quad (1)$$

where  $U_i$  denotes the attribute vector of the  $i$ th user.

Define rating similarity between users  $S_U^{\text{rating}}$  as follows:

$$S_U^{\text{rating}}(i, j) = \frac{1}{1 + \|Ur_i - Ur_j\|}, \quad (2)$$

TABLE 1: Notations and explanations.

Notations	Explanations
$m$	The number of users
$n$	The number of items
$U_i$	Attribute vector of user $i$
$I_i$	Attribute vector of item $i$
$R_{ui}$	Original score of user $u$ to item $i$
$t_{ui}$	Rating time of user $u$ to item $i$
$S_U^{\text{attribute}}(i, j)$	Attribute similarity between users $i$ and $j$
$S_U^{\text{rating}}(i, j)$	Rating similarity between users $i$ and $j$
$S_U$	Similarity matrix between users
$S_I^{\text{attribute}}(i, j)$	Attribute similarity between items $i$ and $j$
$S_I^{\text{rating}}(i, j)$	Rating similarity between items $i$ and $j$
$S_I$	Similarity matrix between items
$A_U$	Attention matrix of user to item
$A_I$	Attention matrix of item to user
$\mathbf{x}_i(t)$	State vector of user node $i$ at time $t$
$\mathbf{y}_i(t)$	State vector of item node $i$ at time $t$

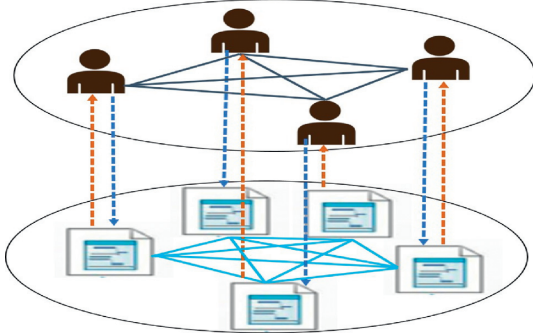


FIGURE 1: Double-layer network model.

where  $U_{r_i}$  and  $U_{r_j}$  denote the score vectors of users  $i$  and  $j$  with sharing scores on items.

In order to get more consistent with the actual user similarity, the final user similarity obtained by associating user attribute similarity and common rating similarity is defined as follows:

$$S_U = \alpha \cdot S_U^{\text{attribute}} + (1 - \alpha) \cdot S_U^{\text{rating}}. \quad (3)$$

Here,  $\alpha \in [0, 1]$  is a mixture parameter, which is the convex combination of two similarity with different angles. So, the adjacent matrix of user layer is obtained, which can represent the coupling relation between users.

(2) *Constructing the Item Layer Network.* Similar to the construction of the user layer, this paper integrates the score information and item attribute into the calculation of similarity between items. In MovieLens dataset, there are 19 attributes of movies. If movie has this attribute, it is 1; if movie does not have this attribute, it is 0. Thus, the attribute vector of the  $i$ th item can be represented as a 19 dimensional vector, denoted as  $I_i$ .

Define attribute similarity between items  $S_I^{\text{attribute}}$  is defined as follows:

$$S_I^{\text{attribute}}(i, j) = \frac{1}{1 + \|I_i - I_j\|}, \quad (4)$$

where  $I_i$  denotes the attribute vector of the  $i$ th item.

Define rating similarity between items  $S_U^{\text{rating}}$  as follows:

$$S_I^{\text{rating}}(i, j) = \frac{1}{1 + \|I_{r_i} - I_{r_j}\|}, \quad (5)$$

where  $I_{r_i}$  and  $I_{r_j}$  denote the score vectors of items  $i$  and  $j$  which are evaluated jointly by the users.

In order to get more accurate relation between items, the final item similarity obtained by associating item attribute similarity and common rating similarity is defined as follows:

$$S_I = \beta \cdot S_I^{\text{attribute}} + (1 - \beta) \cdot S_I^{\text{rating}}. \quad (6)$$

Here,  $\beta \in [0, 1]$  is a mixture parameter, which is the convex combination of two similarities with different angles. So, the adjacent matrix of the item layer is obtained, which can represent the coupling relation between items.

(3) *Connections between the User Layer and the Item Layer.*

In addition, to establish the relationship in the user layer and the item layer, the connection between the two network layers plays a critical role. German psychologist *Hermann Ebbinghaus* found that the human brain forgetting rule is as follows: the process of forgetting is very fast, and forget quickly at first, and then slowly [30]. Inspired by *Ebbinghaus*, we believe that interests of people will also change with time. The closer the score to the current time is, the better it can express current interests of users. According to fitting the *Ebbinghaus* forgetting curve, we obtain a forgetting curve more in line with interests and hobbies of the people. The forgetting function proposed in this paper is as follows:

$$f(t_{ui}) = a \cdot \exp\left(b \cdot \frac{t_{ui} - t_u^{\min}}{t_u^{\max} - t_u^{\min}}\right) + c \cdot \exp\left(d \cdot \frac{t_{ui} - t_u^{\min}}{t_u^{\max} - t_u^{\min}}\right). \quad (7)$$

Moreover, the fitted coefficients are  $a = 0.6368$ ,  $b = -1.947$ ,  $c = 0.3623$ , and  $d = 0.0025$ . Additionally,  $\exp(\cdot)$  is the exponential function.  $f(t_{ui})$  represents the forgetting degree of  $u$ th user to the  $i$ th item.  $t_{ui}$  represents the time of the  $u$ th user rating for the  $i$ th item.  $t_u^{\min}$  represents the earliest rating time of user  $u$ .  $t_u^{\max}$  represents the latest rating time of user  $u$ .

Then, we use the attention mechanism to connect the user layer to the item layer on the processed score. In daily life, everyone pays attention to different things differently. For example, young girls pay more attention to romantic movies than to war movies, and no one can pay attention to everything. Based on this, we define the attention of the user to the item as follows:

$$A_U(u, i) = \frac{\exp(R_{ui} \times f(t_{ui}))}{\sum_{n \in N_u} \exp(R_{ni} \times f(t_{ni}))}. \quad (8)$$

Similarly, every item is not designed for everyone. Some items target different types of users. So, define the attention of the item to the user as follows:

$$A_I(i, u) = \frac{\exp(R_{iu}^T \times f(t_{iu}))}{\sum_{n \in N_i} \exp(R_{in}^T \times f(t_{in}))}. \quad (9)$$

In equations (8) and (9),  $R_{ui}$  represents the initial score of the  $i$ th item rated by the  $u$ th user.  $R_{iu}^T$  is the score that item  $i$  was rated by user  $u$ .  $N_u$  represents the neighbor user set of user  $u$ . We view the items evaluated by user  $u$  as the item neighbor set of user  $u$ .  $N_i$  represents the neighbor set of item  $i$ . We treat all users who have evaluated item  $i$  as neighbors of item  $i$ .  $A_U(u, i)$  represents the attention of the  $u$ th user to the  $i$ th item.  $A_I(i, u)$  represents the attention of the  $i$ th item to the  $u$ th user. Obviously,  $A_U$  and  $A_I$  are not mutually symmetric matrices.

To illustrate the previous definitions, we give a simple example with three users  $U_1$ – $U_3$  and four items  $I_1$ – $I_4$ .  $R$  represents the original scoring matrix:

$$R = \begin{matrix} & I_1 & I_2 & I_3 & I_4 \\ U_1 & \begin{pmatrix} 5 & 3 & 0 & 3 \end{pmatrix} \\ U_2 & \begin{pmatrix} 3 & 3 & 5 & 1 \end{pmatrix} \\ U_3 & \begin{pmatrix} 5 & 0 & 3 & 0 \end{pmatrix} \end{matrix}. \quad (10)$$

$A_U$  represents the user attention matrix, obtained by equation (8):

$$A_U = \begin{matrix} & I_1 & I_2 & I_3 & I_4 \\ U_1 & \begin{pmatrix} 0.0140 & 0.0012 & 0 & 0.0003 \end{pmatrix} \\ U_2 & \begin{pmatrix} 0.0008 & 0.0008 & 0.0057 & 0.0001 \end{pmatrix} \\ U_3 & \begin{pmatrix} 0.0048 & 0 & 0.0008 & 0 \end{pmatrix} \end{matrix}. \quad (11)$$

In the original score matrix  $R$ , we can find that  $R_{12} = 3$  and  $R_{14} = 3$ . They have the same scores, but they are  $A_U(1, 2) = 0.0012$  and  $A_U(1, 4) = 0.0003$  in the attention matrix of users. That means the attention mechanism is meaningful in this paper.

$A_I$  represents the item attention matrix, obtained by equation (9):

$$A_I = \begin{matrix} & U_1 & U_2 & U_3 \\ I_1 & \begin{pmatrix} 0.0109 & 0.0012 & 0.0016 \end{pmatrix} \\ I_2 & \begin{pmatrix} 0.0088 & 0.0104 & 0 \end{pmatrix} \\ I_3 & \begin{pmatrix} 0 & 0.0251 & 0.0009 \end{pmatrix} \\ I_4 & \begin{pmatrix} 0.0031 & 0.0020 & 0 \end{pmatrix} \end{matrix}. \quad (12)$$

Besides, in the original score matrix  $R$ , we can find that  $R_{11} = 5$  and  $R_{31} = 5$ , but by processing,  $A_I(1, 1) = 0.0109$  and  $A_I(1, 3) = 0.0016$  in the attention matrix of items. The attention values of the users to the items are the directed edges from the user layer to the item layer, and the attention values of the items to the users are the directed edges from the item layer to the user layer. Experiments show that attention mechanism can greatly improve the recommendation performance.

**3.2.2. Dynamic Evolution Clustering in Double-Layer Network.** In recent years, due to the explosion of data,

clustering methods emerge endlessly. The clustering method can not only greatly reduce the recommendation time but also improve the recommendation performance. Cluster analysis finds different communities, gathers similar things into one cluster, and pushes dissimilar things in different clusters. Among them, dynamic clustering is more in line with the real situation, so it is applied in various scenarios [19, 21, 22, 31–35]. The phase of the previous dynamic evolution clustering method is only a scalar, which cannot express the interest of users better. In order to grasp the changing rules of interest in different periods, we propose a vector dynamic evolution clustering method.

In this paper, we propose a vector dynamic evolution clustering method in the user layer as follows:

$$\begin{aligned} \mathbf{x}_i(t+1) = & \mathbf{x}_i(t) + K_1 \sum_{j \in DU_i} S_{U_{ij}} \cdot \sin(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \\ & + K_2 \sum_{j \notin DU_i} S_{U_{ij}} \cdot \sin(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \\ & + K_3 \sum_{j \in AU_i} A_{U_{ij}} \cdot \sin(M \cdot \mathbf{y}_j(t) - \mathbf{x}_i(t)) \\ & + K_4 \sum_{j \notin AU_i} A_{U_{ij}} \cdot \sin(M \cdot \mathbf{y}_j(t) - \mathbf{x}_i(t)), \end{aligned} \quad (13)$$

$$i, = 1, 2, \dots, m.$$

Here,  $\mathbf{x}_i(t)$  and  $\mathbf{x}_i(t+1)$  represent the state vectors of user node  $i$  at time  $t$  and time  $t+1$ .  $DU_i \triangleq \{j \mid S_{U_{ij}} \geq \text{Ave}(S_{U_i})\}$ , where  $\text{Ave}(S_{U_i})$  represents the average edge weight of user node  $i$ , which is the average similarity between user node  $i$  and other users.  $AU_i \triangleq \{j \mid A_{U_{ij}} \geq h_1\}$ , where  $h_1$  represents the average value of nonzero elements in attention matrix  $A_U$ .  $K_1$ ,  $K_2$ ,  $K_3$ , and  $K_4$  are the clustering coefficients.  $K_1$  and  $K_3$  are the positive coupling coefficients, and  $K_2$  and  $K_4$  are the negative coupling coefficients. The matrix  $M$  is defined as follows:

$$M_{ij} = \frac{|V_j \cap W_i|}{|V_j|}. \quad (14)$$

Here,  $M_{ij}$  represents the influence degree of the  $i$ th user attribute on the  $j$ th item attribute.  $V_j$  represents the set of users who have evaluated items in the  $j$ th category.  $W_i$  represents the set of users with the  $i$ th attribute. The purpose of adding matrix  $M$  to the evolution equation is to emphasize the different influences of different user attributes.

In the same way, we propose the vector dynamic clustering method in the item layer as follows:

$$\begin{aligned} \mathbf{y}_i(t+1) = & \mathbf{y}_i(t) + K_5 \sum_{j \in DI_i} S_{I_{ij}} \cdot \sin(\mathbf{y}_j(t) - \mathbf{y}_i(t)) \\ & + K_6 \sum_{j \notin DI_i} S_{I_{ij}} \cdot \sin(\mathbf{y}_j(t) - \mathbf{y}_i(t)) \\ & + K_7 \sum_{j \in AI_i} A_{I_{ij}} \cdot \sin(M^T \cdot \mathbf{x}_j(t) - \mathbf{y}_i(t)) \\ & + K_8 \sum_{j \notin AI_i} A_{I_{ij}} \cdot \sin(M^T \cdot \mathbf{x}_j(t) - \mathbf{y}_i(t)), \end{aligned} \quad (15)$$

$$i = 1, 2, \dots, n.$$

Among them,  $\mathbf{y}_i(t)$  and  $\mathbf{y}_i(t+1)$  represent the state vectors of item node  $i$  at time  $t$  and time  $t+1$ .  $DI_i \triangleq \{j | S_{I_{ij}} \geq \text{Ave}(S_{I_i})\}$ , where  $\text{Ave}(S_{I_i})$  represents the average edge weight of user node  $i$ , that is, the average similarity between item node  $i$  and other items.  $AI_i \triangleq \{\{j\} A_{I_{ij}} \geq h_2\}$ , where  $h_2$  represents the average value of nonzero elements in attention matrix  $\mathbf{A}_I$ .  $K_5$ ,  $K_6$ ,  $K_7$ , and  $K_8$  are the clustering coefficients.  $K_5$  and  $K_7$  are the positive coupling coefficients, and  $K_6$  and  $K_8$  are the negative coupling coefficients.  $M^T$  is the transpose of the matrix  $M$ . The purpose of adding matrix  $M^T$  to the evolution equation is to emphasize the different influences of different item attributes.

Besides, this community detection evolution process can be stable after some iterations. The convergence results are obtained from the following theorems according to *Lya-punov* theory.

**Theorem 1.** *Vector dynamic evolution process equations (13) and (15) can be converted into the following forms:*

$$\begin{aligned} \mathbf{x}(t+1) &\leq \mathbf{x}(t) + \mathbf{S}\mathbf{x}(t) + \mathbf{T}\mathbf{y}(t), \\ \mathbf{y}(t+1) &\leq \mathbf{y}(t) + \mathbf{G}\mathbf{y}(t) + \mathbf{H}\mathbf{x}(t). \end{aligned} \quad (16)$$

*Proof.* See it in the Appendix, in the end of this paper.  $\square$

**Theorem 2.** *If appropriate parameters  $K_1$ ,  $K_2$ ,  $K_3$ ,  $K_4$ ,  $K_5$ ,  $K_6$ ,  $K_7$ , and  $K_8$  make the  $\theta < 0$  and  $\omega < 0$ , then the fixed points in equations (13) and (15) are uniformly stable.*

*Proof.* See it in the Appendix, in the end of this paper.

The convergence analysis shows that our community detection algorithm will be stable after some iterations. Finally, the user nodes with similar state vectors and item nodes with similar state vectors are assigned to the same community, that is, all users and all items with higher similarity are assigned to their community with similar interest.  $\square$

**3.2.3. Score Prediction and Recommendation.** In order to sort the user similarity of the same community and obtain the nearest neighbor set of the target user for collaborative filtering recommendation, Cosine similarity is used in this paper to calculate the community similarity:

$$\text{Sim}_{uv}^{\text{COS}} = \frac{\sum_{i \in S_{uv}} R_{ui} R_{vi}}{\sqrt{\sum_{i \in S_{uv}} R_{ui}^2} \sqrt{\sum_{i \in S_{uv}} R_{vi}^2}} \quad (17)$$

To compare different similarity indexes, Pearson correlation coefficient and adjusted Cosine similarity [36, 37] are adopted:

$$\text{Sim}_{uv}^{\text{PCC}} = \frac{\sum_{i \in S_{uv}} (R_{ui} - \bar{R}_u)(R_{vi} - \bar{R}_v)}{\sqrt{\sum_{i \in S_{uv}} (R_{ui} - \bar{R}_u)^2} \sqrt{\sum_{i \in S_{uv}} (R_{vi} - \bar{R}_v)^2}} \quad (18)$$

$$\text{Sim}_{uv}^{\text{ACOS}} = \frac{\sum_{i \in S_{uv}} (R_{ui} - \bar{R}_u)(R_{vi} - \bar{R}_v)}{\sqrt{\sum_{i \in S_u} (R_{ui} - \bar{R}_u)^2} \sqrt{\sum_{i \in S_v} (R_{vi} - \bar{R}_v)^2}}$$

Here,  $R_{ui}$  represents the score of the  $i$ th item is rated by the  $u$ th user,  $S_{uv}$  represents the set of items both scored by users  $u$  and  $v$ ,  $\bar{R}_u$  represents the average score of user  $u$ , and  $S_u$  represents the set of items scored by user  $u$ . Through the above methods, the similarity ranking of target users within the community can be obtained, and the score can be predicted.

Previous prediction methods did not take into account the item community and averaged all item scores, which could not better express the score of target users in the community of target items. In order to better reflect the role of the community, this paper proposes the following methods to predict the score:

$$P_{ui} = \bar{R}_{uC_i} + \frac{\sum_{v \in N_u} \text{Sim}_{uv} \times (R_{vi} - \bar{R}_u)}{\sum_{v \in N_u} \text{Sim}_{uv}} \quad (19)$$

Here,  $C_i$  is the community of item  $i$ ,  $\bar{R}_{uC_i}$  is the average score of user  $u$  to the  $i$ th item community,  $N_u$  represents the nearest neighbor set of user  $u$ ,  $P_{ui}$  represents the prediction score of target user  $u$  for item  $i$ ,  $\text{Sim}_{uv}$  represents the similarity value between users  $u$  and  $v$ , and  $\bar{R}_u$  represents the average score of user  $u$ .

**3.2.4. Algorithm Flowchart.** This paper presents a recommendation algorithm in Double-layer Network based on Vector dynamic evolution Clustering and Attention mechanism (denoted as DN-VCA). The pseudocode is shown in Algorithm 1.

Firstly, we use node attribute information to construct an undirected network in the layer. Secondly, connect double layers of networks with attention mechanism so that a double-layer network connection is established as directed relationship. Thirdly, community detection is carried out according to vector dynamic evolution clustering. Fourthly, according to the new prediction method proposed in this paper, score prediction is carried out. Finally, we give a list of recommendations.

The flow chart of the algorithm proposed in this paper is shown in Figure 2.

## 4. Experiments and Results

**4.1. Datasets.** To verify the effectiveness of the model and our proposed algorithm, Movielens-100k, CiaoDVD, and Filmtrust datasets are tested in this paper.

- (i) **Movielens-100k**<sup>1</sup>: MovieLens is a set of movie ratings. The dataset contains 100,000 ratings provided by 943 users for 1682 movies, and scores are 1–5. Each score has its corresponding time. The dataset also has user attribute information and movie category information. Each user has at least 20 score records. The data sparsity is 93.7%.
- (ii) **CiaoDVD**<sup>2</sup>: the dataset contains 278,483 ratings provided by 7375 users for 99746 DVDs, and scores are 1–5. As the sparsity of this dataset was 99.97%, we retain the users with more than 20 evaluation DVD and the DVD with more than 20 evaluation values so that the sparsity after



**Input:** Training set  $u_1$ , Test set  $u_2$ . Parameters  $K_1, K_2, \dots, K_8, \epsilon_1, \epsilon_2$ .  
**Output:** Prediction score  $P_{ui}$ , MAE, RMSE, Recommendation list.  
 //Constructing the double-layer network  
 Compute the User **similarity** matrix  $S_U$ , Item similarity matrix  $S_I$ , User-item attention matrix  $A_U$ , Item-user attention matrix  $A_I$  according to equations (3), (6), (8) and (9).  
 //Community detection in double-layer network  
**for** each node **do**  
   Apply vector dynamic evolution clustering equations (13) and (15) to find the appropriate community for each node.  
   **If**  $\|x_i(t+1) - x_i(t)\| < \epsilon_1$  and  $\|y_j(t+1) - y_j(t)\| < \epsilon_2$  **then** all nodes stop iterating  
   **end if**  
**end for**  
 //Calculate the dynamic similarity  
**for** each user community and item community **do**  
   Apply equation (17) to calculate the similarity matrix in the same community.  
**end for**  
 //Prediction score  
**for** each target user and target item **do**  
   Find the neighbor set of the target user by similarity sort.  
   Compute the prediction score according to equation (19)  
**end for**  
 //Select the Top-N items as recommendation list for target user

ALGORITHM 1: DN-VCA.

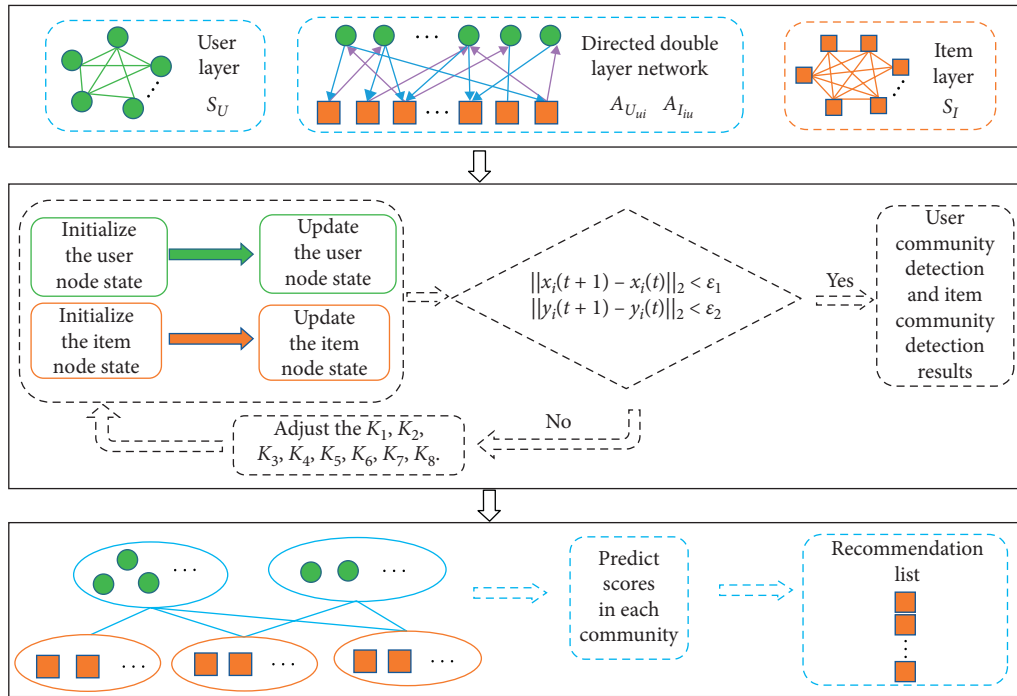


FIGURE 2: Flowchart of DN - VCA.

processing is 97.01% and the processed dataset is denoted as **CiaoDVD-1**.

- (iii) **Filmtrust**<sup>3</sup>: the dataset contains 35,497 ratings provided by 1508 users for 2071 movies, and scores are 0.5–4 in intervals of 0.5. The data sparsity is 98.86%. Moreover, most items have not been evaluated by users, so we only keep the items that have been evaluated more than 3 times. At the same

time, we delete the users whose evaluation times are less than 3 times. Finally, the data sparsity was slightly reduced to 96.61% and the processed dataset is denoted as **Filmtrust-1**.

The original dataset is randomly divided into 80% training set and 20% test set to verify the proposed algorithm on three datasets. In addition, all the algorithms in this paper have conducted five cross experiments, and finally, the

average of five cross experiments are taken as the results. Table 2 gives the statistics of three datasets.

**4.2. Evaluation Indexes.** In order to verify the accuracy of the algorithm proposed in this paper, we use the following five evaluation indicators:

$$\begin{aligned} \text{MAE} &= \frac{\sum_{i=1}^M |R_{ui} - P_{ui}|}{M}, \\ \text{RMSE} &= \sqrt{\frac{M \sum_{i=1}^M (R_{ui} - P_{ui})^2}{M}}. \end{aligned} \quad (20)$$

Here, MAE is the mean absolute error, RMSE is the root mean squared error,  $R_{ui}$  represents the true score of item  $i$  by user  $u$  in the test set,  $P_{ui}$  represents the prediction score generated by the algorithm, and  $M$  represents the number of scores be predicted in the test set. Fixing integer processing is conducted on Movielens and CiaoDVD-1 datasets, but the scores on Filmtrust-1 are floating point numbers with 0.5 as the interval, so no integer processing is performed on this dataset.

Since we present recommendations to target users after the prediction, the following three indexes are used to verify the efficiency of recommendations:

$$\begin{aligned} \text{Recall} &= \frac{\sum_u |R_u \cap T_u|}{\sum_u |T_u|}, \\ \text{Precision} &= \frac{\sum_u |R_u \cap T_u|}{\sum_u |R_u|}, \\ F_1 &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \end{aligned} \quad (21)$$

Here,  $R_u$  represents the set of items recommended by the algorithm for the target user  $u$  and  $T_u$  represents the set of items that target user  $u$  really likes in the test set. Precision indicates the proportion of the items the user really likes in the recommendation list to the total number of recommendations. Recall represents the ratio of the favorite items of users in the recommendation list to the favorite items of the user.  $F_1$  value is a comprehensive indicator of Precision and Recall.

**4.3. Parameter Analysis.** There are several parameters in our proposed algorithm and we will discuss their influences in the recommendation performance.

**The influence of convex combination parameters  $\alpha$  in equation 3 and  $\beta$  in equation 6:** here, on the Movielens dataset, the selected sets of  $\alpha$  and  $\beta$  both are [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]. We tested the effect of  $\alpha$  and  $\beta$  values on MAE and RMSE.

As shown in Figure 3(a), when  $\alpha = 0$  and  $\beta = 0$ , that is, user similarity and item similarity only contain score information, MAE = 0.7085. When  $\alpha = 1$  and  $\beta = 1$ , that is, user similarity and item similarity only contain attribute information, MAE = 0.7637. When  $\alpha = 1$  and  $\beta = 0$ , that is, user similarity only contains attribute information and item similarity only contains score information, MAE = 0.7087.

When  $\alpha = 0$  and  $\beta = 1$ , that is, user similarity only contains score information and item similarity only contains attribute information, MAE = 0.7634. From the results, we can see that MAE = 0.7045 is the lowest, as  $\alpha = 0.3$  and  $\beta = 0.4$ , in other words, the experimental results show the best when the similarity includes both score information and attribute information. Similarly, as shown in Figure 3(b), RMSE = 0.9947 is lowest when  $\alpha = 0.3$  and  $\beta = 0.4$ . In summary,  $\alpha = 0.3$  and  $\beta = 0.4$  on Movielens dataset in the following experiment. Because there is no attribute information in CiaoDVD-1 and Filmtrust-1 datasets, so  $\alpha = 0$  and  $\beta = 0$ .

**The influence of thresholds  $\varepsilon_1$  and  $\varepsilon_2$  in termination criteria:** in order to obtain stable state vectors of nodes more accurately and faster, we limit the number of iterations of dynamic evolution clustering equations. If the state difference between the front and back vectors is less than the threshold, iteration is terminated. We selected several pairs of user layer threshold  $\varepsilon_1$  and item layer threshold  $\varepsilon_2$  for experiments.

As shown in Figure 4, on the Movielens dataset, both MAE and RMSE are the lowest when  $\varepsilon_1 = 0.2$  and  $\varepsilon_2 = 0.1$ . In other words, when  $\varepsilon_1 = 0.2$  and  $\varepsilon_2 = 0.1$ , the algorithm achieves a better prediction accuracy. Therefore, on the Movielens dataset, we choose  $\varepsilon_1 = 0.2$  and  $\varepsilon_2 = 0.1$  for experiments.

As shown in Figure 5, on the CiaoDVD-1 dataset, both MAE and RMSE are the lowest when  $\varepsilon_1 = 2.1$  and  $\varepsilon_2 = 2.1$ , except for the prediction when the number of neighbors is 90. So, on the CiaoDVD-1 dataset, we choose  $\varepsilon_1 = 2.1$  and  $\varepsilon_2 = 2.1$  for following experiments.

As it can be seen from Figure 6, MAE = 0.6164 when  $\varepsilon_1 = 0.01$  and  $\varepsilon_2 = 0.01$ , and the number of neighbors is 60. When  $\varepsilon_1 = 0.001$  and  $\varepsilon_2 = 0.001$ , MAE = 0.6155. When the neighbor number is taken into other values, both MAE and RMSE of  $\varepsilon_1 = 0.01$  and  $\varepsilon_2 = 0.01$  are the lowest, that is, the prediction effect is the best. As a result, the two threshold values selected in the experiment on Filmtrust-1 dataset in this paper are  $\varepsilon_1 = 0.01$  and  $\varepsilon_2 = 0.01$ .

In this paper, only four pairs of parameters are selected for comparison on each dataset, and finally the threshold values with the best effect are selected. In fact, thresholds other than the above can lead to better predictions.

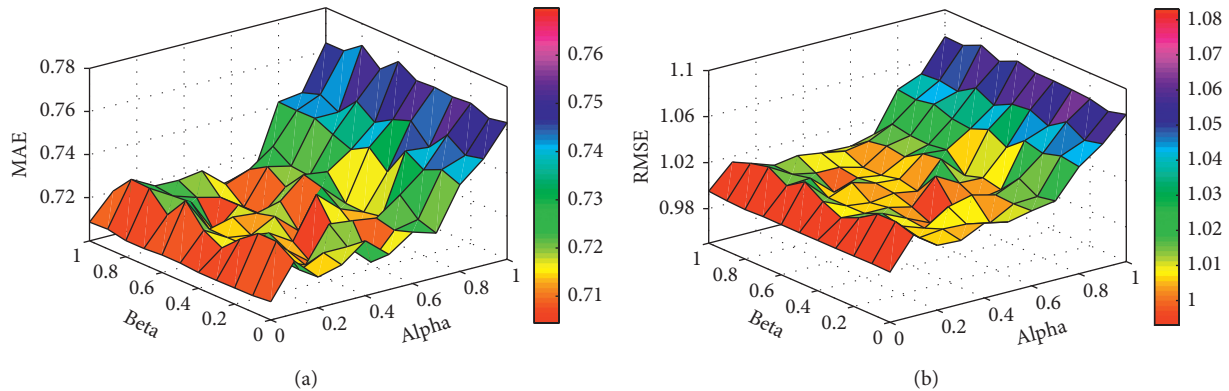
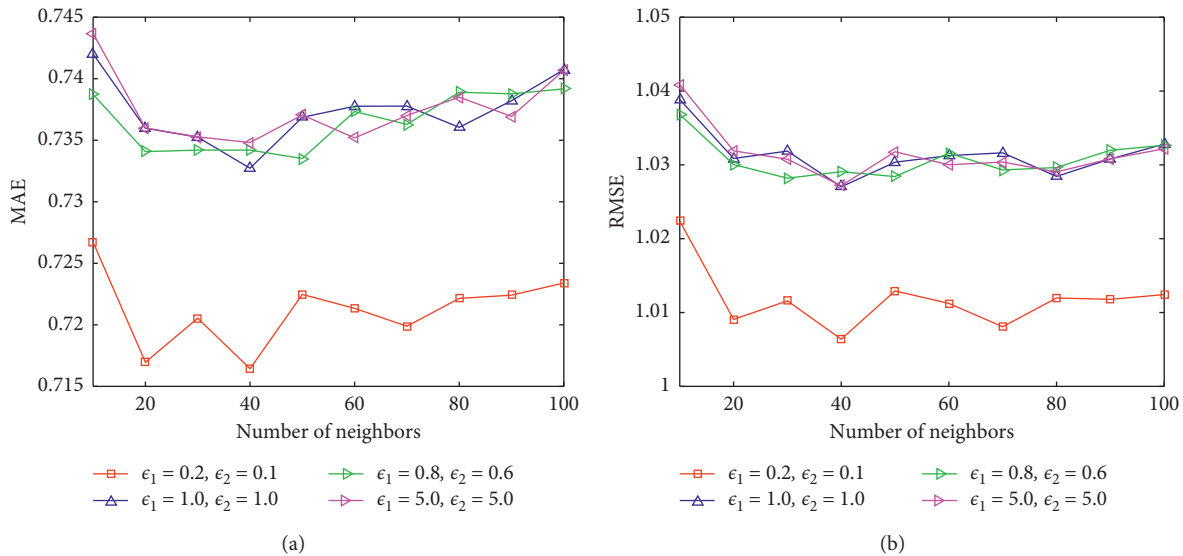
**4.4. Comparing Similarity.** In order to get better results, we choose one of Cosine similarity, Pearson correlation coefficient, and adjusted Cosine similarity, so we conducted experiments on the Movielens dataset.

As shown in Figure 7, because MAE and RMSE of Cosine similarity are lower than Pearson correlation coefficient and adjusted Cosine similarity regardless of the number of neighbors, we select Cosine as the measure of dynamic similarity within the community.

**4.5. Comparing Results of Different Recommendation Algorithms.** In this paper, our proposed algorithm (DN-VCA) is compared with three existing algorithms from Collaborative Filtering recommendation algorithm based on  $K$ -means clustering (named as  $K$ -CF), Ref. [38] (named as DTNM), and Ref. [32] (named as EHC-CF). The selected neighbor set is [10, 20, 30, 40, 50, 60, 70, 80, 90, 100].

TABLE 2: Statistics of datasets.

	Movielens-100k	CiaoDVD	CiaoDVD-1	Filmtrust	Filmtrust-1
User	943	7375	471	1508	1227
Item	1682	99746	689	2071	793
Rating	100000	278483	9707	35497	33009
Scale	[1-5]	[1-5]	[1-5]	[0.5-4]	[0.5-4]
Sparse degree (%)	93.71	99.97	97.01	98.86	96.61

FIGURE 3: Movielens. (a) MAE results of different values of  $\alpha$  and  $\beta$ . (b) RMSE results of different values of  $\alpha$  and  $\beta$ .FIGURE 4: Movielens. (a) MAE results between different values of  $\epsilon_1$  and  $\epsilon_2$ . (b) RMSE results between different values of  $\epsilon_1$  and  $\epsilon_2$ .

4.5.1. *Movielens-100k*. In order to compare the accuracy of our proposed algorithm with other algorithms, a number of experiments are carried out on the *Movielens* dataset.

Figure 8 shows the clustering results based on our double-layer network evolutionary clustering method. Six communities are formed for *Movielens-100k*.

As shown in Figure 9, regardless of the number of neighbors, the MAE values and RMSE values of our DN-VCA are lower than the other three compared algorithms.

The reason our DN-VCA has a good performance is that double-layer network can better represent the relations between users and items. Moreover, our prediction method is no longer based on the user community, but adds the item community into the prediction method, highlighting the role of the item community.

Next, we present the Top-N recommendation list for the user, and select the set of recommended number as [2, 4, 6, 8, 10, 12, 14, 16, 18, 20]. The comparison results of *Precision*, *Recall*, and  $F_1$  are shown in Table 3.

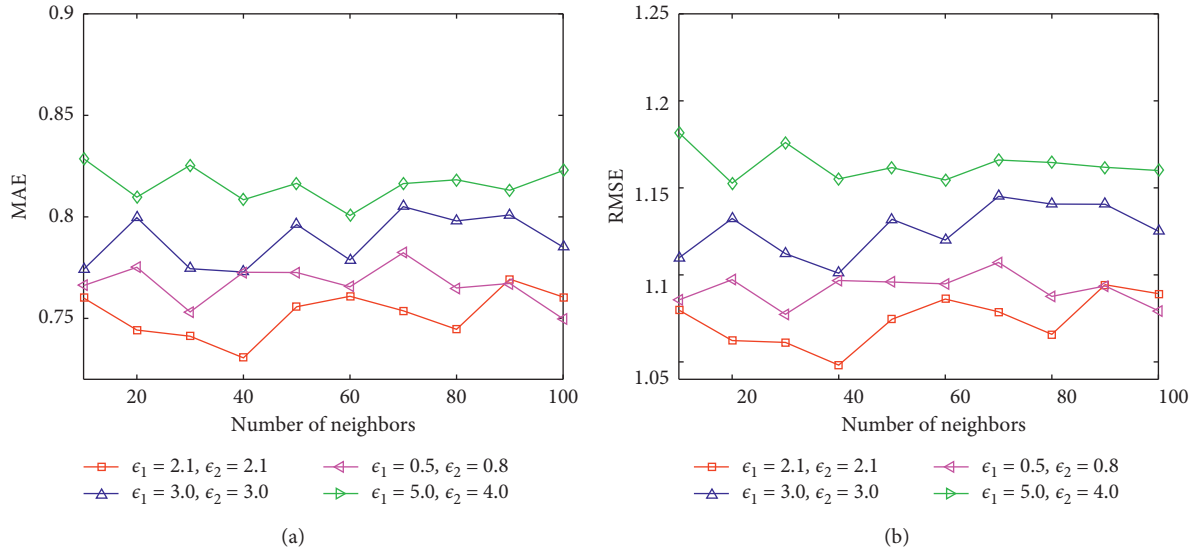


FIGURE 5: CiaoDVD-1. (a) MAE results between different values of  $\epsilon_1$  and  $\epsilon_2$ . (b) RMSE results between different values of  $\epsilon_1$  and  $\epsilon_2$ .

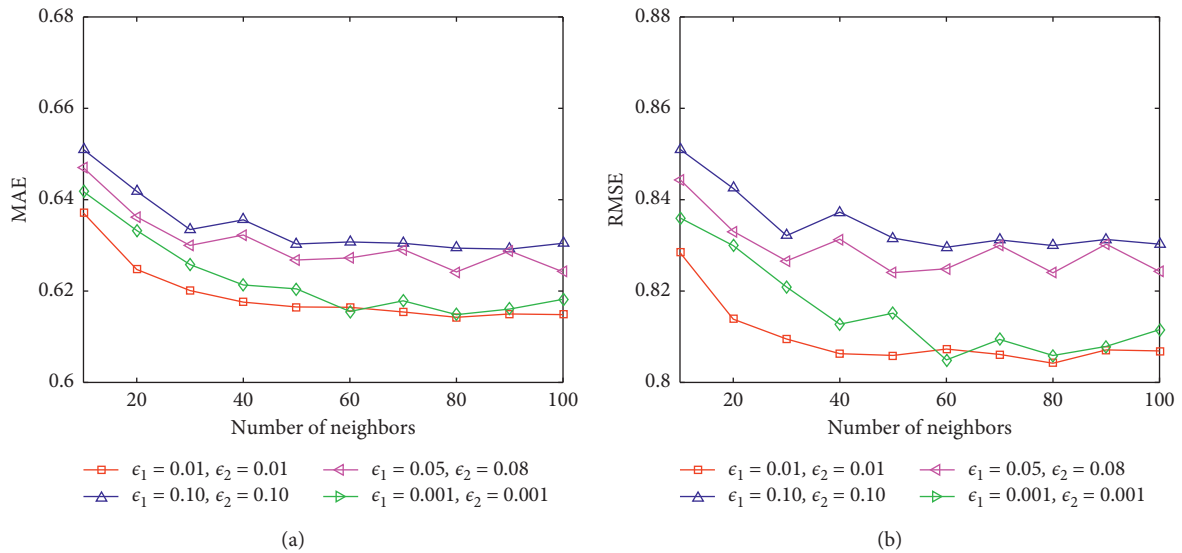


FIGURE 6: FilmTrust-1. (a) MAE results between different values of  $\epsilon_1$  and  $\epsilon_2$ . (b) RMSE results between different values of  $\epsilon_1$  and  $\epsilon_2$ .

As shown in Table 3, as the number of recommendations changes, *Recall* and  $F_1$  values will increase, while the *Precision* will decrease. Compared with algorithm *DTNM*, *EHC-CF*, and *K-CF*, our *DN-VCA* has a slightly higher advantage in *Precision*, *Recall*, and  $F_1$  values.

**4.5.2. CiaoDVD-1.** To further test the performance of the algorithm, we test many experiments on the CiaoDVD-1 dataset. As shown in Figure 10, because the sparsity of this dataset is very large, the result of the algorithm has a strong oscillation. However, no matter how many neighbors we have, the MAE and RMSE values of our *DN-VCA* are lower than the other three compared algorithms, except that when the number of neighbors is 20, and the MAE and RMSE values are slightly higher than *EHC-CF*.

As shown in Table 4, compared with algorithm *DTNM* and *K-CF*, our *DN-VCA* has a significant advantage, and our *DN-VCA* is very close with *EHC-CF*. In general, the *DN-VCA* algorithm we proposed is higher than the other three algorithms in MAE, RMSE, *Precision*, *Recall*, and  $F_1$  values.

**4.5.3. FilmTrust-1.** Different from the previous three datasets, the scale of this dataset is 0.5, so in the final prediction results, the experiments in this paper do not conduct rounding processing on the predicted scores.

It can be seen from Figure 11 that the *DN-VCA* proposed in this paper is slightly lower than *EHC-CF* when the number of neighbors is 10, and the MAE of *EHC-CF* is 0.6365, while the MAE of *DN-VCA* is 0.6371. When the number of neighbors is taken as other values, the accuracy of

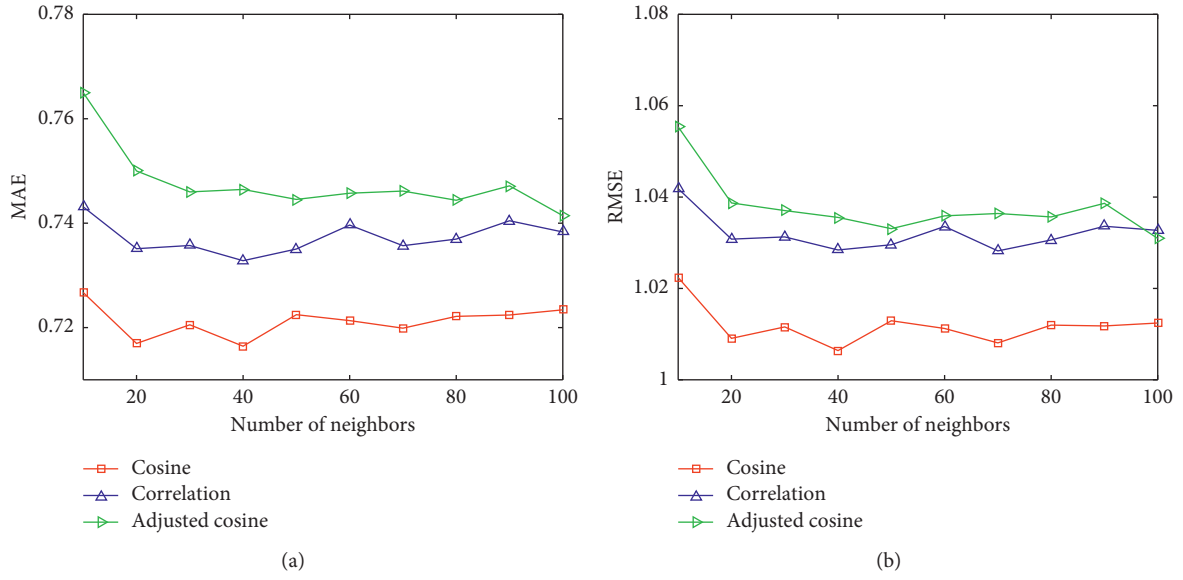


FIGURE 7: Movielens. (a) MAE results between three different similarities. (b) RMSE results between three different similarities.

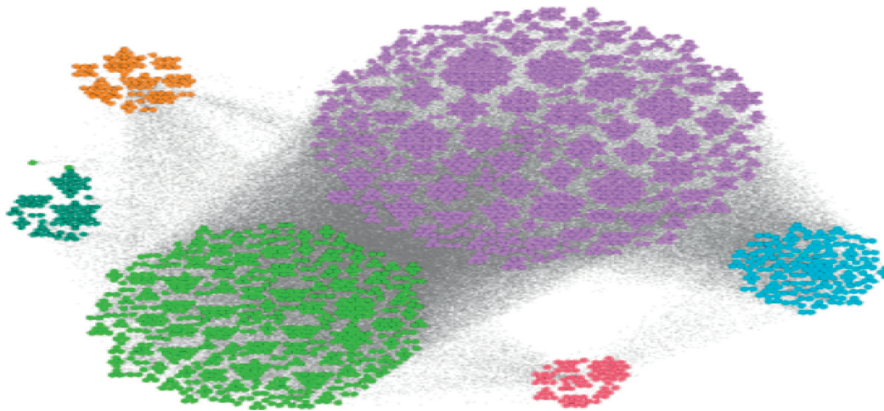


FIGURE 8: Clustering results based on *Movielens-100k*.

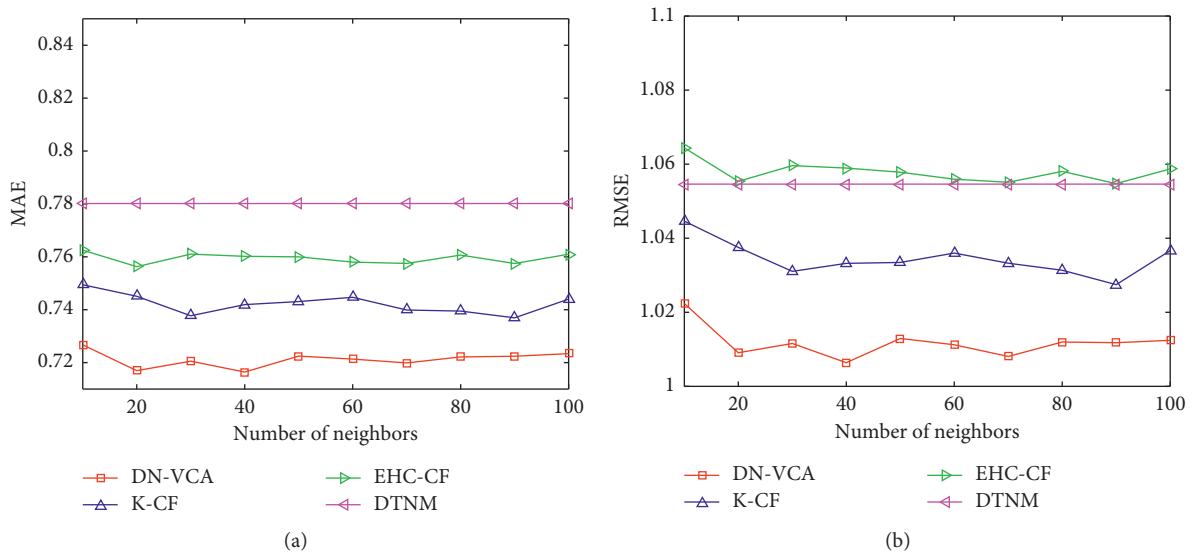


FIGURE 9: Movielens. (a) Comparison results of MAE of four algorithms. (b) Comparison results of RMSE of four algorithms.

TABLE 3: Precision, recall, and  $F_1$  of EHC-CF, DTNM, K-CF, and DN-VCA based on Movielens dataset.

Algorithm	Metrics (%)	Recommended number									
		2	4	6	8	10	12	14	16	18	20
EHC-CF	Precision	90.3805	89.7659	89.3578	88.7287	88.3573	88.2147	87.8782	87.7129	87.6330	87.4568
	Recall	7.8480	15.1348	21.7199	27.5310	32.6933	37.3786	41.5252	45.3709	48.9157	52.1453
	$F_1$	14.3758	25.7226	34.6698	41.6885	47.3606	52.1235	56.0057	59.4090	62.3873	64.9418
DTNM	Precision	91.2053	89.7964	89.1047	88.8487	88.4917	88.2884	88.2264	87.9556	87.7681	87.6785
	Recall	7.9072	15.1289	21.6744	27.5701	32.7656	37.4209	41.6886	45.4873	48.9955	52.2710
	$F_1$	14.4857	25.7146	34.5931	41.7475	47.4577	52.1785	56.2263	59.5654	62.4880	65.1010
K-CF	Precision	91.2274	90.1739	89.5158	89.1396	88.8328	88.4051	88.2116	88.0297	87.9305	87.7079
	Recall	7.9072	15.2000	21.7640	27.6597	32.8672	37.4738	41.6812	45.5312	49.0895	52.2939
	$F_1$	14.4860	25.8335	34.7383	41.8829	47.6137	52.2509	56.2164	59.6208	62.6062	65.1270
DN-VCA	Precision	<b>91.7247</b>	<b>90.4707</b>	<b>89.7541</b>	<b>89.4433</b>	<b>88.9443</b>	<b>88.5499</b>	<b>88.3380</b>	<b>88.0890</b>	<b>87.9783</b>	<b>87.8778</b>
	Recall	<b>7.9570</b>	<b>15.2437</b>	<b>21.8078</b>	<b>27.7400</b>	<b>32.9021</b>	<b>37.5253</b>	<b>41.7313</b>	<b>45.5466</b>	<b>49.1077</b>	<b>52.3803</b>
	$F_1$	<b>14.5765</b>	<b>25.9100</b>	<b>34.8122</b>	<b>42.0086</b>	<b>47.6662</b>	<b>52.3267</b>	<b>56.2874</b>	<b>59.6466</b>	<b>62.6328</b>	<b>65.2411</b>

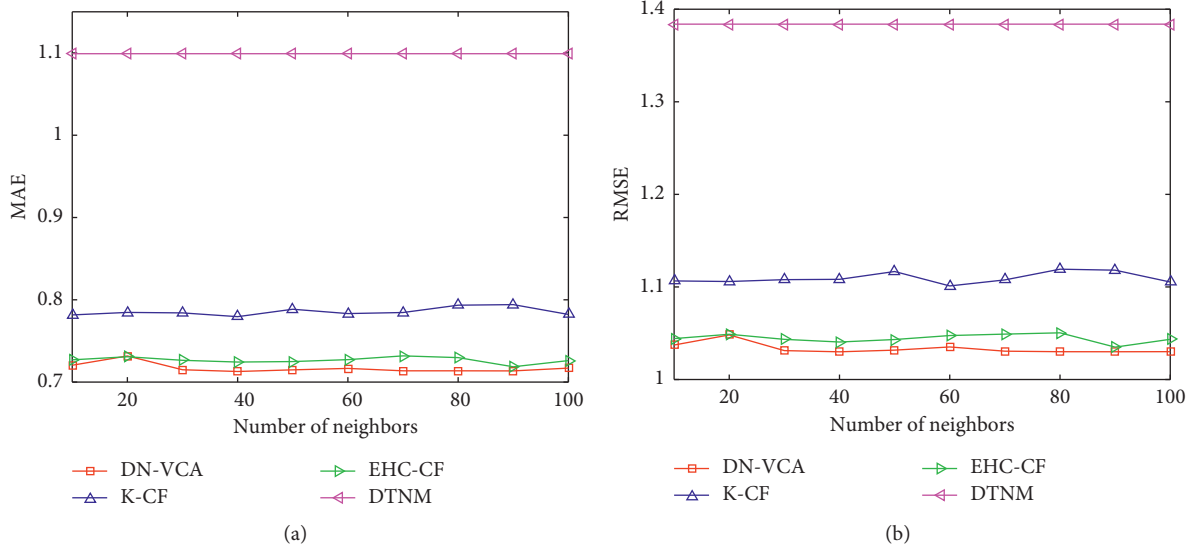


FIGURE 10: CiaoDVD-1. (a) Comparison results of MAE of four algorithms. (b) Comparison results of RMSE of four algorithms.

TABLE 4: Precision, recall, and  $F_1$  of EHC-CF, DTNM, K-CF, and DN-VCA based on CiaoDVD-1 dataset.

Algorithm	Metrics (%)	Recommended number									
		2	4	6	8	10	12	14	16	18	20
EHC-CF	Precision	24.5900	13.6557	9.6985	7.6019	6.3958	5.5571	4.9320	4.4214	4.0643	3.7721
	Recall	11.3877	11.8937	12.3989	12.8134	13.3843	13.8904	14.3312	14.6425	15.1090	15.5502
	$F_1$	15.5662	12.7134	10.8831	9.5420	8.6551	7.9379	7.3381	6.7916	6.4052	6.0711
DTNM	Precision	13.9274	8.4802	5.9297	5.1109	4.4484	3.8422	3.6072	3.4007	3.1722	3.0071
	Recall	6.4437	7.3774	7.5714	8.6098	9.2974	9.5947	10.4794	11.2553	11.7857	12.3821
	$F_1$	8.8107	7.8901	6.6504	6.4139	6.0174	5.4869	5.3667	5.2231	4.9987	4.8388
K-CF	Precision	20.6154	12.1088	8.2584	6.7321	5.5848	4.9203	4.3153	4.0033	3.7157	3.3792
	Recall	9.5448	10.5441	10.5573	11.3497	11.6845	12.2939	12.5426	13.2556	13.8116	13.9290
	$F_1$	13.0480	11.2720	9.2669	8.4509	7.5571	7.0276	6.4210	6.1492	5.8557	5.4387
DN-VCA	Precision	<b>24.8136</b>	<b>13.7930</b>	<b>9.8938</b>	<b>7.6697</b>	<b>6.5324</b>	<b>5.6663</b>	<b>4.9717</b>	<b>4.4229</b>	<b>4.1280</b>	<b>3.8314</b>
	Recall	<b>11.4897</b>	<b>12.0078</b>	<b>12.6432</b>	<b>12.9301</b>	<b>13.6684</b>	<b>14.1613</b>	<b>14.4467</b>	<b>14.6415</b>	<b>15.3412</b>	<b>15.7958</b>
	$F_1$	<b>15.7062</b>	<b>12.8381</b>	<b>11.1002</b>	<b>9.6277</b>	<b>8.8396</b>	<b>8.0936</b>	<b>7.3972</b>	<b>6.7932</b>	<b>6.5052</b>	<b>6.1667</b>

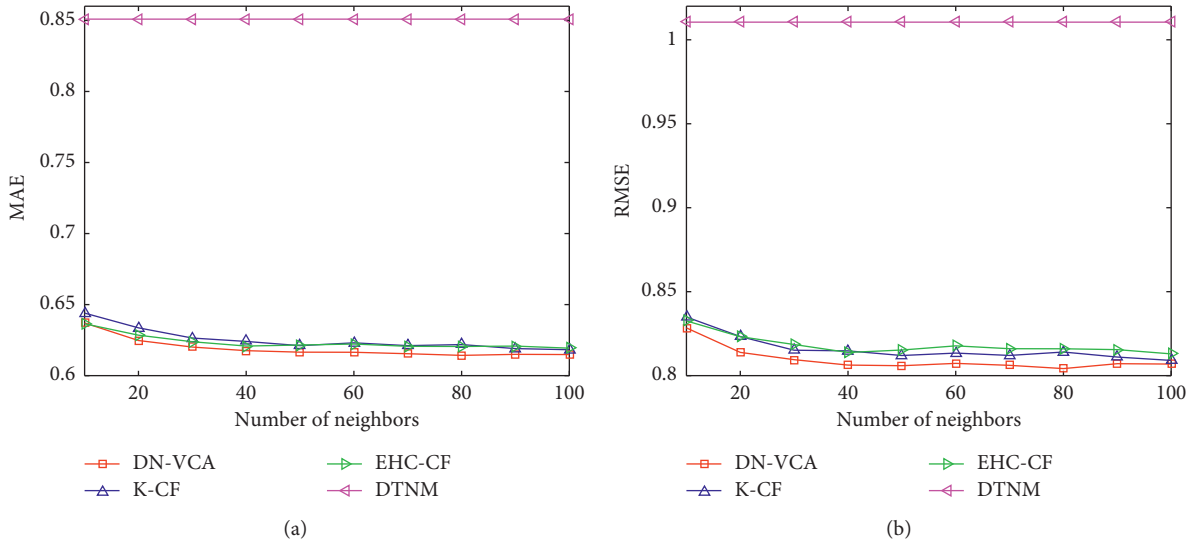


FIGURE 11: FilmTrust-1. (a) Comparison results of MAE of four algorithms. (b) Comparison results of RMSE of four algorithms.

TABLE 5: Precision, recall, and  $F_1$  of EHC – CF, DTNM, K – CF, and DN – VCA based on FilmTrust-1 dataset.

Algorithm	Metrics (%)	Recommended number									
		2	4	6	8	10	12	14	16	18	20
EHC-CF	Precision	82.2555	83.8450	82.8444	82.2364	81.5371	81.2952	81.1014	81.0292	80.9482	80.9327
	Recall	34.3482	57.5925	73.4343	84.5687	91.3565	95.2834	97.2184	98.1894	98.6618	99.0253
	$F_1$	48.9674	68.2819	77.8552	83.3853	86.1669	87.7341	88.4304	88.7868	88.9306	89.0687
DTNM	Precision	85.2386	84.0522	83.0014	82.1382	81.6513	81.2537	81.0887	80.9642	80.9636	80.9419
	Recall	34.3415	57.7353	73.5729	84.4672	91.4840	95.2343	97.2034	98.1105	98.6805	99.0364
	$F_1$	48.9578	68.4510	78.0024	83.2855	86.2874	87.6891	88.4167	88.7155	88.9475	89.0788
K-CF	Precision	85.1992	84.1723	83.0559	82.2075	81.7250	81.3210	81.1107	81.0446	<b>81.0157</b>	80.9571
	Recall	34.3256	57.8176	73.6217	84.5388	91.5668	95.3132	97.2297	98.2083	<b>98.7445</b>	99.0555
	$F_1$	48.9351	68.5486	78.0539	83.3559	86.3654	87.7617	88.4407	88.8037	<b>89.0050</b>	89.0957
DN-VCA	Precision	<b>85.8327</b>	<b>84.3690</b>	<b>83.3180</b>	<b>82.4043</b>	<b>81.7817</b>	<b>81.3914</b>	<b>81.1733</b>	<b>81.0943</b>	81.0097	<b>80.9695</b>
	Recall	<b>34.5804</b>	<b>57.9529</b>	<b>73.8548</b>	<b>84.7414</b>	<b>91.6301</b>	<b>95.3956</b>	<b>97.3045</b>	<b>98.2682</b>	98.7368	<b>99.0703</b>
	$F_1$	<b>49.2986</b>	<b>68.7090</b>	<b>78.3006</b>	<b>83.5556</b>	<b>86.4252</b>	<b>87.8376</b>	<b>88.5088</b>	<b>88.8580</b>	88.9982	<b>89.1092</b>

our *DN-VCA* has significant advantages with *DTNM*. Compared with *EHC-CF* and *K-CF*, the *MAE* and *RMSE* of *DN-VCA* are also the lowest. The results show that the proposed *DN-VCA* achieves good performance on FilmTrust-1 dataset when predicting scores.

As shown in Table 5, each of the four algorithms gives similar results in *Precision*, *Recall*, and  $F_1$  values. However, *DN-VCA* that we proposed has the best results, which are optimal in *Precision*, *Recall*, and  $F_1$  values except that when the number of recommendations is 18, the  $F_1$  value of *DN-VCA* is lower than 0.0068 of *K-CF*, the *Precision* of *DN-VCA* is lower than 0.006 of *K-CF*, and the *Recall* of *DN-VCA* is lower than 0.0077 of *K-CF*. *DN-VCA* obtained the best recommendation result when the number of recommendations is 20.

Experimental results on three datasets show that *MAE* and *RMSE* of our proposed *DN-VCA* are lower than the other four comparison algorithms in predicting scores, so our proposed algorithm is effective in predicting scores. In addition, in terms of recommendation, *DN-VCA* also achieves better results than other algorithms in *Precision*, *Recall*, and  $F_1$  values. The *K-CF* and *EHC-CF* of the three

comparison methods are recommendation algorithms based on clustering algorithms. Our results show that our vector dynamic evolution clustering algorithm outperforms these other clustering algorithms, hence suggesting that our proposed method can also be effective for generating recommendations. So, it proves that the algorithm we proposed is effective in the recommend system. That means our double-layer network construction is meaningful and the dynamic clustering in the double-layer network can gather the similar interest users together to the same community. Neighbor users with high similarity give valuable suggestions in the collaborative filter recommendation process.

## 5. Conclusion

In this paper, a novel vector dynamic evolutionary clustering recommendation algorithm *DN-VCA* based on double-layer network and attention mechanism is proposed. Our algorithm firstly constructs a double-layer network model through node similarity and an attention mechanism. Then, the improved vector dynamic evolution clustering equation

is used in the double-layer network to cluster the nodes into the most suitable community. Finally, the similarity between nodes is calculated within the community for enabling collaborative filtering recommendation. We not only verify the validity of the DN-VCA, but also prove the theoretical results of the algorithm. Additionally, we solve the shortcomings of the existing methods. For example, previous algorithms are only based on the single-layer network, and the state of nodes is only a scalar when clustering is dynamically evolving. With the development of big data, the recommendation system faces more and more users and items. It is impossible for us to carry out similar comparison with all users for collaborative filtering based on users, and the vector dynamic evolution clustering proposed in this paper is a good community detection method to solve this problem. Please note that our algorithm has a number of degrees of freedom that are highly nontrivial to optimize. Actually, our results do not use optimal parameters values; if those parameters could be further optimized then the performance of our algorithm would further improve. Finding efficient optimization methods for these parameters constitute an interesting field of future research.

## Appendix

### A. Proof of Theorem 1

In order to evolve the double-layer network model, the following matrices are firstly defined:

$$\begin{aligned}
P^+ &= \left\{ A_{U_{ij}} \mid A_{U_{ij}} \geq h_1, \quad i = 1, \dots, m, j = 1, \dots, n \right\}, \\
P^- &= \left\{ A_{U_{ij}} \mid A_{U_{ij}} < h_1, \quad i = 1, \dots, m, j = 1, \dots, n \right\}, \\
Q^+ &= \left\{ A_{I_{ij}} \mid A_{I_{ij}} \geq h_2, \quad i = 1, \dots, n, j = 1, \dots, m \right\}, \\
Q^- &= \left\{ A_{I_{ij}} \mid A_{I_{ij}} < h_2, \quad i = 1, \dots, n, j = 1, \dots, m \right\}, \\
U^+ &= \left\{ S_{U_{ij}} \mid S_{U_{ij}} \geq \text{Ave}(S_{U_i}), \quad i = 1, \dots, m, j = 1, \dots, m \right\}, \\
U^- &= \left\{ S_{U_{ij}} \mid S_{U_{ij}} < \text{Ave}(S_{U_i}), \quad i = 1, \dots, m, j = 1, \dots, m \right\}, \\
I^+ &= \left\{ S_{I_{ij}} \mid S_{I_{ij}} \geq \text{Ave}(S_{I_i}), \quad i = 1, \dots, n, j = 1, \dots, n \right\}, \\
I^- &= \left\{ S_{I_{ij}} \mid S_{I_{ij}} < \text{Ave}(S_{I_i}), \quad i = 1, \dots, n, j = 1, \dots, n \right\}.
\end{aligned} \tag{A.1}$$

Matrix  $P^+$  keeps the elements of  $\text{Att}_{U_{ij}} \geq h_1$ , and the rest are all 0. Matrix  $Q^+$  keeps the elements of  $\text{Att}_{I_{ij}} \geq h_2$ , and the rest elements in  $Q^+$  are all 0. And similarly, matrix  $P^-$  keeps the elements of  $\text{Att}_{U_{ij}} < h_1$ , and the rests are all 0. Matrix  $Q^-$  keeps the elements of  $\text{Att}_{I_{ij}} < h_2$ , and the rest elements in  $Q^-$  are all 0.  $U^+$  and  $U^-$  contain elements that satisfy  $S_{U_{ij}} \geq \text{Ave}(S_{U_i})$  and  $S_{U_{ij}} < \text{Ave}(S_{U_i})$ , and the other elements are all 0. Matrix  $I^+$  contains all elements that are greater than or equal to the mean similarity of each item. Matrix  $I^-$  contains all elements less than the mean similarity of each item, and all other elements are equal to 0.

**Theorem A.1.** *Vector dynamic evolution process equations (13) and (15) can be converted into the following forms:*

$$\begin{aligned}
\mathbf{x}(t+1) &\leq \mathbf{x}(t) + \mathbf{S}\mathbf{x}(t) + \mathbf{T}\mathbf{y}(t), \\
\mathbf{y}(t+1) &\leq \mathbf{y}(t) + \mathbf{G}\mathbf{y}(t) + \mathbf{H}\mathbf{x}(t).
\end{aligned} \tag{A.2}$$

*Proof*

(i) The user layer vector dynamic evolution process equation (13) is

$$\begin{aligned}
\mathbf{x}_i(t+1) &= \mathbf{x}_i(t) + K_1 \sum_{j \in DU_i} S_{U_{ij}} \cdot \sin(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \\
&\quad + K_2 \sum_{j \notin DU_i} S_{U_{ij}} \cdot \sin(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \\
&\quad + K_3 \sum_{j \in AU_i} A_{U_{ij}} \cdot \sin(M \cdot \mathbf{y}_j(t) - \mathbf{x}_i(t)) \\
&\quad + K_4 \sum_{j \notin AU_i} A_{U_{ij}} \cdot \sin(M \cdot \mathbf{y}_j(t) - \mathbf{x}_i(t)).
\end{aligned} \tag{A.3}$$

The initial values are all selected from the range of  $[0, \pi/2]$ .  $S_{U_{ij}} \in [0, 1]$  and  $A_{U_{ij}} \in [0, 1]$ . We can obtain  $-(\pi/2) \leq \mathbf{x}_j(t) - \mathbf{x}_i(t) \leq (A_{U_{ij}} \in [0, 1])$  and  $-(\pi/2) \leq \mathbf{y}_j(t) - \mathbf{y}_i(t) \leq (\pi/2)$ . And then can obtain  $\sin(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \leq \mu_{ij}(\mathbf{x}_j(t) - \mathbf{x}_i(t))$  and  $\sin(\mathbf{y}_j(t) - \mathbf{y}_i(t)) \leq \lambda_{ij}(\mathbf{y}_j(t) - \mathbf{y}_i(t))$ , where  $\mu_{ij} > 0$  and  $\lambda_{ij} > 0$  and they meet the following conditions:

$$\begin{aligned}
\mu_{ij} &= \begin{cases} \frac{1}{2}, & -\frac{\pi}{2} \leq \mathbf{x}_j(t) - \mathbf{x}_i(t) < 0, \\ 1, & 0 \leq \mathbf{x}_j(t) - \mathbf{x}_i(t) \leq \frac{\pi}{2}, \end{cases} \\
\lambda_{ij} &= \begin{cases} \frac{1}{2}, & -\frac{\pi}{2} \leq \mathbf{y}_j(t) - \mathbf{y}_i(t) < 0, \\ 1, & 0 \leq \mathbf{y}_j(t) - \mathbf{y}_i(t) \leq \frac{\pi}{2}. \end{cases}
\end{aligned} \tag{A.4}$$

Then, equation (13) can be converted to

$$\begin{aligned}
\mathbf{x}_i(t+1) &= \mathbf{x}_i(t) + K_1 \sum_{j=1}^m U_{ij}^+ \cdot \mu_{ij}(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \\
&\quad + K_2 \sum_{j=1}^m U_{ij}^- \cdot m_{ij}(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \\
&\quad + K_3 \sum_{j=1}^n P_{ij}^+ \cdot \lambda_{ij}(M\mathbf{y}_j(t) - \mathbf{x}_i(t)) \\
&\quad + K_4 \sum_{j=1}^n P_{ij}^- \cdot \lambda_{ij}(M\mathbf{y}_j(t) - \mathbf{x}_i(t)).
\end{aligned} \tag{A.5}$$



Then,

$$\begin{aligned}
\mathbf{x}_i(t+1) &= \mathbf{x}_i(t) + K_1 \sum_{j=1}^m (U_{ij}^+ \mu_{ij} \mathbf{x}_j(t)) - K_1 \left( \sum_{j=1}^m U_{ij}^+ \mu_{ij} \right) \mathbf{x}_i(t) + K_2 \sum_{j=1}^m (U_{ij}^- \mu_{ij} \mathbf{x}_j(t)) - K_2 \left( \sum_{j=1}^m U_{ij}^- \mu_{ij} \right) \mathbf{x}_i(t) \\
&\quad + K_3 \sum_{j=1}^n (P_{ij}^+ \lambda_{ij} M \mathbf{y}_j(t)) - K_3 \left( \sum_{j=1}^n P_{ij}^+ \lambda_{ij} \right) \mathbf{x}_i(t) + K_4 \sum_{j=1}^n (P_{ij}^- \lambda_{ij} M \mathbf{y}_j(t)) - K_4 \left( \sum_{j=1}^n P_{ij}^- \lambda_{ij} \right) \mathbf{x}_i(t) \\
&\leq \mathbf{x}_i(t) + K_1 \sum_{j=1}^m U_{ij}^+ \mathbf{x}_j(t) - \frac{K_1}{2} \left( \sum_{j=1}^m U_{ij}^+ \right) \mathbf{x}_i(t) + K_2 \sum_{j=1}^m U_{ij}^- \mathbf{x}_j(t) - \frac{K_2}{2} \left( \sum_{j=1}^m U_{ij}^- \right) \mathbf{x}_i(t) \\
&\quad + K_3 \sum_{j=1}^n P_{ij}^+ M \mathbf{y}_j(t) - \frac{K_3}{2} \left( \sum_{j=1}^n P_{ij}^+ \right) \mathbf{x}_i(t) + K_4 \sum_{j=1}^n P_{ij}^- M \mathbf{y}_j(t) - \frac{K_4}{2} \left( \sum_{j=1}^n P_{ij}^- \right) \mathbf{x}_i(t).
\end{aligned} \tag{A.6}$$

Suppose the user state vector  $\mathbf{x}_i(t)$  is an  $f$ -dimensional column vector, and the item state vector  $\mathbf{y}_i(t)$  is a  $g$ -dimensional column vector. Denote  $\mathbf{x}(t) = (\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_m(t))^T$  and  $\mathbf{y}(t) = (\mathbf{y}_1(t), \mathbf{y}_2(t), \dots, \mathbf{y}_n(t))^T$ . So,  $\mathbf{x}(t)$  is a column vector in  $mf$  dimension and  $\mathbf{y}(t)$  is a column vector in  $ng$  dimension.

Define a special matrix  $A^+$  as follows:

$$\begin{bmatrix}
U_{11}^+ & 0 & \cdots & 0 & \cdots & U_{1m}^+ & 0 & \cdots & 0 \\
0 & U_{11}^+ & \cdots & 0 & \cdots & 0 & U_{1m}^+ & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & U_{11}^+ & \cdots & 0 & 0 & \cdots & U_{1m}^+ \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
U_{m1}^+ & 0 & \cdots & 0 & \cdots & U_{mm}^+ & 0 & \cdots & 0 \\
0 & U_{m1}^+ & \cdots & 0 & \cdots & 0 & U_{mm}^+ & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & U_{m1}^+ & \cdots & 0 & 0 & \cdots & U_{mm}^+
\end{bmatrix}. \tag{A.7}$$

Matrix  $A^+$  is a large matrix of  $mf \times mf$  dimension, composed of  $m^2$  small diagonal matrices with  $f$  dimension. Similarly, the elements of  $A^-$  are made up of the elements of  $U^-$ .

Define the diagonal elements of matrix  $B^+$  as follows:

$$\begin{aligned}
B_{ii}^+ &= \sum_{j=1}^m U_{1j}^+, \quad i = 1, \dots, f, \\
B_{ii}^+ &= \sum_{j=1}^m U_{2j}^+, \quad i = f+1, \dots, 2f, \\
&\dots \\
B_{ii}^+ &= \sum_{j=1}^m U_{mj}^+, \quad i = (m-1)f+1, \dots, mf.
\end{aligned} \tag{A.8}$$

$B^+$  matrix is the diagonal matrix of  $mf \times mf$ , the first  $f$  diagonal elements are the same, the  $f+1$  to  $2f$  elements are the same, and so on, and the last  $f$  elements are the same. Similarly, we can define  $B^-$ , and the elements of  $B^-$  are made up of the elements of the matrix  $P^-$ .

Then, define the matrix  $C^+$  as follows:

$$\begin{bmatrix}
P_{11}^+ & 0 & \cdots & 0 & \cdots & P_{1n}^+ & 0 & \cdots & 0 \\
0 & P_{11}^+ & \cdots & 0 & \cdots & 0 & P_{1n}^+ & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & P_{11}^+ & \cdots & 0 & 0 & \cdots & P_{1n}^+ \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
P_{m1}^+ & 0 & \cdots & 0 & \cdots & P_{mn}^+ & 0 & \cdots & 0 \\
0 & P_{m1}^+ & \cdots & 0 & \cdots & 0 & P_{mn}^+ & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & P_{m1}^+ & \cdots & 0 & 0 & \cdots & P_{mn}^+
\end{bmatrix}. \tag{A.9}$$

Matrix  $C^+$  is a large matrix of  $mf \times nf$  dimension. When the elements of  $C^+$  become are substituted by the elements of  $P^-$ , the above matrix becomes  $C^-$ .

Define the matrix  $E$  as follows:

$$E = \begin{bmatrix}
M & & & \\
& M & & \\
& & \ddots & \\
& & & M
\end{bmatrix}. \tag{A.10}$$

The matrix  $E$  is made up of multiple  $M$  matrices and it is diagonal block matrix. Matrix  $M$  is the  $f \times g$  dimension matrix defined by equation (14). The dimension of matrix  $E$  is  $nf \times ng$ , which means that matrix  $E$  is made up of  $n$  matrices  $M$ .

And finally, we define the matrix  $D^+$  as follows:

$$\begin{aligned} D_{ii}^+ &= \sum_{j=1}^n P_{1j}^+, \quad i = 1, \dots, f, \\ D_{ii}^+ &= \sum_{j=1}^n P_{2j}^+, \quad i = f + 1, \dots, 2f, \\ &\dots \\ D_{ii}^+ &= \sum_{j=1}^n P_{mj}^+, \quad i = (m-1)f + 1, \dots, mf. \end{aligned} \quad (\text{A.11})$$

The composition of the matrix  $D^-$  is the same as the composition of the matrix  $D^+$ , and the element source is the matrix  $P^-$ .

Based on the above definitions,  $\mathbf{x}(t+1)$  can be further transformed:

$$\begin{aligned} \mathbf{x}(t+1) &\leq \mathbf{x}(t) + K_1 A^+ \mathbf{x}(t) - \frac{K_1}{2} B^+ \mathbf{x}(t) + K_2 A^- \mathbf{x}(t) - \frac{K_2}{2} B^- \mathbf{x}(t) \\ &\quad + K_3 C^+ E \mathbf{y}(t) - \frac{K_3}{2} D^+ \mathbf{x}(t) + K_4 C^- E \mathbf{y}(t) - \frac{K_4}{2} D^- \mathbf{x}(t) \\ &\leq \mathbf{x}(t) + \left( K_1 A^+ - \frac{K_1}{2} B^+ + K_2 A^- - \frac{K_2}{2} B^- - \frac{K_3}{2} D^+ - \frac{K_4}{2} D^- \right) \mathbf{x}(t) \\ &\quad + (K_3 C^+ E + K_4 C^- E) \mathbf{y}(t). \end{aligned} \quad (\text{A.12})$$

Denote

$$\begin{aligned} S &= K_1 A^+ - \frac{K_1}{2} B^+ + K_2 A^- - \frac{K_2}{2} B^- - \frac{K_3}{2} D^+ - \frac{K_4}{2} D^-, \\ T &= K_3 C^+ E + K_4 C^- E. \end{aligned} \quad (\text{A.13})$$

Then, we have

$$\mathbf{x}(t+1) \leq \mathbf{x}(t) + S \mathbf{x}(t) + T \mathbf{y}(t). \quad (\text{A.14})$$

(ii) Similarly, the item layer vector dynamic evolution equation 15 can also be in a similar form.

$J^+$  is defined as follows:

$$\begin{aligned} &\begin{bmatrix} I_{11}^+ & 0 & \dots & 0 & \dots & I_{1n}^+ & 0 & \dots & 0 \\ 0 & I_{11}^+ & \dots & 0 & \dots & 0 & I_{1n}^+ & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{11}^+ & \dots & 0 & 0 & \dots & I_{1n}^+ \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ I_{n1}^+ & 0 & \dots & 0 & \dots & I_{nm}^+ & 0 & \dots & 0 \\ 0 & I_{n1}^+ & \dots & 0 & \dots & 0 & I_{nm}^+ & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{n1}^+ & \dots & 0 & 0 & \dots & I_{nm}^+ \end{bmatrix}, \end{aligned} \quad (\text{A.15})$$

$$K_{ii}^+ = \sum_{j=1}^n I_{1j}^+, \quad i = 1, \dots, g,$$

$$K_{ii}^+ = \sum_{j=1}^n I_{2j}^+, \quad i = g + 1, \dots, 2g,$$

...

$$K_{ii}^+ = \sum_{j=1}^n I_{nj}^+, \quad i = (n-1)g + 1, \dots, ng.$$

$L^+$  is defined as follows:

$$\begin{bmatrix} Q_{11}^+ & 0 & \dots & 0 & \dots & Q_{1m}^+ & 0 & \dots & 0 \\ 0 & Q_{11}^+ & \dots & 0 & \dots & 0 & Q_{1m}^+ & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_{11}^+ & \dots & 0 & 0 & \dots & Q_{1m}^+ \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Q_{n1}^+ & 0 & \dots & 0 & \dots & Q_{nm}^+ & 0 & \dots & 0 \\ 0 & Q_{n1}^+ & \dots & 0 & \dots & 0 & Q_{nm}^+ & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_{n1}^+ & \dots & 0 & 0 & \dots & Q_{nm}^+ \end{bmatrix}, \quad (\text{A.16})$$

$$F = \begin{bmatrix} M^T & & & \\ & M^T & & \\ & & \ddots & \\ & & & M^T \end{bmatrix},$$

$$N_{ii}^+ = \sum_{j=1}^m Q_{1j}^+, \quad i = 1, \dots, g,$$

$$N_{ii}^+ = \sum_{j=1}^m Q_{2j}^+, \quad i = g + 1, \dots, 2g,$$

...

$$N_{ii}^+ = \sum_{j=1}^m Q_{nj}^+, \quad i = (n-1)g + 1, \dots, ng.$$

Similarly, the elements of  $J^-$  are made up of the elements of  $I^-$ . The elements of  $K^-$  are made up of the elements of the matrix  $Q^-$ . When the elements of the  $L^+$  become elements of  $Q^-$ , this matrix becomes the  $L^-$ . The matrix  $F$  is made up of

multiple  $M^T$  matrices. The composition of the matrix  $N^-$  is the same as the composition of the matrix  $N^+$ , and the element source is the matrix  $Q^-$ .

Then, we have

$$\mathbf{y}(t+1) \leq \mathbf{y}(t) + G\mathbf{y}(t) + H\mathbf{x}(t). \quad (\text{A.17})$$

Here,

$$G = K_5 J^+ - \frac{K_5}{2} K^+ + K_6 J^- - \frac{K_6}{2} K^- - \frac{K_7}{2} N^+ - \frac{K_8}{2} N^-,$$

$$H = K_7 L^+ F + K_8 L^- F.$$

(A.18)  $\square$

## B. Proof of Theorem 2

**Theorem B.1.** *If appropriate parameters  $K_1, K_2, K_3, K_4, K_5, K_6, K_7$ , and  $K_8$  make the  $\theta < 0$  and  $\omega < 0$ , then the fixed points in equations (13) and (15) are uniformly stable.*

*Proof.* Since any isolated equilibrium state can be moved to the origin of the state space by coordinate transformation, we only discuss the stability of the equilibrium state at the origin of the coordinates.

Lyapunov function is defined as follows:

$$\begin{aligned} V(t) &= \mathbf{x}(t)^T \mathbf{x}(t), \\ W(t) &= \mathbf{y}(t)^T \mathbf{y}(t). \end{aligned} \quad (\text{B.1})$$

So, the following transformations are conducted:

$$\begin{aligned} \Delta V &= V(t+1) - V(t) \\ &= \mathbf{x}(t+1)^T \mathbf{x}(t+1) - \mathbf{x}(t)^T \mathbf{x}(t) \\ &\leq [\mathbf{x}(t) + S\mathbf{x}(t) + T\mathbf{y}(t)]^T [\mathbf{x}(t) + S\mathbf{x}(t) + T\mathbf{y}(t)] - \mathbf{x}(t)^T \mathbf{x}(t) \\ &\leq [\mathbf{x}(t)^T + \mathbf{x}(t)^T S^T + \mathbf{y}(t)^T T^T] [\mathbf{x}(t) + S\mathbf{x}(t) + T\mathbf{y}(t)] - \mathbf{x}(t)^T \mathbf{x}(t) \\ &\leq \mathbf{x}(t)^T (S + S^T + S^T S) \mathbf{x}(t) + \mathbf{x}(t)^T (T + S^T T) \mathbf{y}(t) + \mathbf{y}(t)^T (T^T + T^T S) \mathbf{x}(t) + \mathbf{y}(t)^T T^T T \mathbf{y}(t). \end{aligned} \quad (\text{B.2})$$

Similarly,

$$\begin{aligned} \Delta W &= W(t+1) - W(t) \\ &= \mathbf{y}(t+1)^T \mathbf{y}(t+1) - \mathbf{y}(t)^T \mathbf{y}(t) \\ &\leq [\mathbf{y}(t) + G\mathbf{y}(t) + H\mathbf{x}(t)]^T [\mathbf{y}(t) + G\mathbf{y}(t) + H\mathbf{x}(t)] - \mathbf{y}(t)^T \mathbf{y}(t) \\ &\leq [\mathbf{y}(t)^T + \mathbf{y}(t)^T G^T + \mathbf{x}(t)^T H^T] [\mathbf{y}(t) + G\mathbf{y}(t) + H\mathbf{x}(t)] - \mathbf{y}(t)^T \mathbf{y}(t) \\ &\leq \mathbf{y}(t)^T (G + G^T + G^T G) \mathbf{y}(t) + \mathbf{y}(t)^T (H + G^T H) \mathbf{x}(t) + \mathbf{x}(t)^T (H^T + H^T G) \mathbf{y}(t) + \mathbf{x}(t)^T H^T H \mathbf{x}(t). \end{aligned} \quad (\text{B.3})$$

If appropriate parameters  $K_1, K_2, K_3, K_4, K_5, K_6, K_7$ , and  $K_8$  make

$$\begin{aligned} \theta &= \mathbf{x}(t)^T (S + S^T + S^T S) \mathbf{x}(t) + \mathbf{x}(t)^T (T + S^T T) \mathbf{y}(t) + \mathbf{y}(t)^T (T^T + T^T S) \mathbf{x}(t) + \mathbf{y}(t)^T T^T T \mathbf{y}(t) < 0, \\ \omega &= \mathbf{y}(t)^T (G + G^T + G^T G) \mathbf{y}(t) + \mathbf{y}(t)^T (H + G^T H) \mathbf{x}(t) + \mathbf{x}(t)^T (H^T + H^T G) \mathbf{y}(t) + \mathbf{x}(t)^T H^T H \mathbf{x}(t) < 0, \end{aligned} \quad (\text{B.4})$$

according to the stability theory of Lyapunov [34], the fixed points in equations (13) and (15) are uniformly stable. In fact, given the proper values of  $K_1, K_2, K_3, K_4, K_5, K_6, K_7$ , and  $K_8$ , the stability condition can be achieved.

The above theoretical analysis shows that the evolution process of the vector dynamic evaluation equations is consistent and stable. As time goes on, the state vectors of the nodes are stable and the clustering results are discovered.

In fact, the condition of  $\theta < 0$  and  $\omega < 0$  can be easily satisfied and we can find clustering parameters  $K_1, \dots, K_8$  with a wide value range.  $\square$

## Data Availability

Firstly, the data used to support the findings of this study are all open online and we have listed in the article on page 8. (<https://grouplens.org/datasets/movielens/100k/>, <http://www.public.asu.edu/jtang20/datasetcode/truststudy.htm>, and <https://www.librec.net/datasets/filmtrust.zip>) Secondly, the Matlab code 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 research was supported by the National Natural Science Foundation of China (nos. 71561020, 61503203, 61702317, and 61771297) and Fundamental Research Funds for the Central Universities (no. GK201802013).

## References

- [1] M. Deshpande and G. Karypis, "Item-based top-N recommendation algorithms," *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 143–177, 2004.
- [2] P. Zhou, Y. Zhou, D. Wu, and H. Jin, "Differentially private online learning for cloud-based video recommendation with multimedia big data in social networks," *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1217–1229, 2016.
- [3] J.-H. Su, H.-H. Yeh, P. S. Yu, and V. S. Tseng, "Music recommendation using content and context information mining," *IEEE Intelligent Systems*, vol. 25, no. 1, pp. 16–26, 2010.
- [4] L. Yao, Q. Z. Sheng, A. H. H. Ngu, J. Yu, and A. Segev, "Unified collaborative and content-based web service recommendation," *IEEE Transactions on Services Computing*, vol. 8, no. 3, pp. 453–466, 2015.
- [5] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, 2003.
- [6] S. H. Choi, Y.-S. Jeong, and M. K. Jeong, "A hybrid recommendation method with reduced data for large-scale application," *IEEE Transactions on Systems*, vol. 40, no. 5, pp. 557–566, 2010.
- [7] K. Yu, A. Schwaighofer, V. Tresp, X. Xu, and H.-P. Kriegel, "Probabilistic memory-based collaborative filtering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 1, pp. 56–69, 2004.
- [8] A. Hernando, J. Bobadilla, and F. Ortega, "A non negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model," *Knowledge-Based Systems*, vol. 97, pp. 188–202, 2016.
- [9] X. Li, X. Cheng, S. Su, S. Li, and J. Yang, "A hybrid collaborative filtering model for social influence prediction in event-based social networks," *Neurocomputing*, vol. 230, pp. 197–209, 2017.
- [10] D. Billsus and M. J. Pazzani, "Learning collaborative information filters," in *Proceedings of the International Conference on Machine Learning*, pp. 46–54, Madison, WI, USA, 1998.
- [11] X. Ma, D. Dong, and Q. Wang, "Community detection in multi-layer networks using joint nonnegative matrix factorization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 2, pp. 273–286, 2019.
- [12] X. Ma, P. Sun, and Y. Wang, "Graph regularized nonnegative matrix factorization for temporal link prediction in dynamic networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 496, pp. 121–136, 2018.
- [13] X. Ling, D. Guo, F. Cai, and H. Chen, "User-based clustering with top-N recommendation on cold-start problem," in *Proceedings of the 2013 Third International Conference on Intelligent System Design and Engineering Applications*, pp. 1585–1589, Hong Kong, China, January 2013.
- [14] J. D. West, I. Wesley-Smith, and C. T. Bergstrom, "A recommendation system based on hierarchical clustering of an article-level citation network," *IEEE Transactions on Big Data*, vol. 2, no. 2, pp. 113–123, 2016.
- [15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.
- [16] S. Zahra, M. A. Ghazanfar, A. Khalid, M. A. Azam, U. Naeem, and A. Prugel-Bennett, "Novel centroid selection approaches for K means-clustering based recommender systems," *Information Sciences*, vol. 320, pp. 156–189, 2015.
- [17] J. Wu, Y. Hou, Y. Jiao, Y. Li, X. Li, and L. Jiao, "Density shrinking algorithm for community detection with path based similarity," *Physica A: Statistical Mechanics and Its Applications*, vol. 433, pp. 218–228, 2015.
- [18] L. Bai, J. Liang, H. Du, and Y. Guo, "A novel community detection algorithm based on simplification of complex networks," *Knowledge-Based Systems*, vol. 143, pp. 58–64, 2018.
- [19] J. Wu and Y. Jiao, "Clustering dynamics of complex discrete-time networks and its application in community detection," *Chaos*, vol. 24, Article ID 033104, 2014.
- [20] J. Wu, F. Wang, and P. Xiang, "Automatic network clustering via density-constrained optimization with grouping operator," *Applied Soft Computing*, vol. 38, pp. 606–616, 2016.
- [21] J. Chen, H. Wang, L. Wang, and W. Liu, "A dynamic evolutionary clustering perspective: community detection in signed networks by reconstructing neighbor sets," *Physica A: Statistical Mechanics and Its Applications*, vol. 447, pp. 482–492, 2016.
- [22] J. Chen, I. Zhang, W. Liu, and Z. Yan, "Community detection in signed networks based on discrete-time model," *Chinese Physics B*, vol. 26, no. 1, Article ID 018901, 2017.
- [23] Z. Bu, H.-J. Li, J. Cao, Z. Wang, and G. Gao, "Dynamic cluster formation game for attributed graph clustering," *IEEE Transactions on Cybernetics*, vol. 49, no. 1, pp. 328–341, 2019.
- [24] P. Moradi, S. Ahmadian, and F. Akhlaghian, "An effective trust-based recommendation method using a novel graph clustering algorithm," *Physica A: Statistical Mechanics and Its Applications*, vol. 436, pp. 462–481, 2015.
- [25] J. Ren, J. Long, and Z. Xu, "Financial news recommendation based on graph embeddings," *Decision Support Systems*, vol. 125, Article ID 113115, 2019.
- [26] T. Liang, L. Zheng, L. Chen, Y. Wan, P. S. Yu, and J. Wu, "Multi-view factorization machines for mobile app recommendation based on hierarchical attention," *Knowledge-Based Systems*, vol. 187, Article ID 104821, 2020.
- [27] C. Feng, Z. Liu, S. Lin, and T. Q. S. Quek, "Attention-based graph convolutional network for recommendation system," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 7560–7564, Brighton, UK, April 2019.
- [28] W. Shang, S. Shang, S. Feng, and M. Shi, "An improved video recommendations based on the hyperlink-graph model," in *Proceedings of the 2016 4th International Conference on Applied Computing and Information Technology/3rd International Conference on Computational Science/Intelligence and Applied Informatics/1st International Conference on Big Data, Cloud Computing, Data Science and Engineering*, pp. 379–383, Las Vegas, NV, USA, December 2016.
- [29] Y. Yasami, "A new knowledge-based link recommendation approach using a non-parametric multilayer model of dynamic complex networks," *Knowledge-Based Systems*, vol. 143, pp. 81–92, 2018.
- [30] H. Ebbinghaus, H. A. Ruger, and C. E. Bussenius, "Memory: a contribution to experimental psychology," *Journal of Annals of Neurosciences*, vol. 20, no. 4, p. 155, 2013.
- [31] J. Shao, X. He, Q. Yang, C. Bohm, and C. Plant, "Synchronization inspired partitioning and hierarchical clustering,"

- IEEE Transactions on Knowledge and Data Engineering*, vol. 25, pp. 893–905, 2013.
- [32] J. Chen, H. Uliji, and C. Zhao, “Collaborative rating prediction based on dynamic evolutionary heterogeneous clustering,” *Bio-Inspired Computing—Theories and Applications*, vol. 682, pp. 394–399, 2017.
- [33] J. Wu, L. Jiao, J. Chao, F. Liu, M. Gong, and R. Shang, “Overlapping community detection via network dynamics,” *Physical Review E*, vol. 85, no. 1, Article ID 016115, 2012.
- [34] G. Teschl, “Ordinary differential equations and dynamical systems,” *Atlantis Studies in Differential Equations*, vol. 140, no. 3, pp. 189–194, 2004.
- [35] U. Liji, Y. Chai, and J. Chen, “Improved personalized recommendation based on user attributes clustering and score matrix filling,” *Computer Standards and Interfaces*, vol. 57, pp. 59–67, 2018.
- [36] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.
- [37] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu, “A new user similarity model to improve the accuracy of collaborative filtering,” *Knowledge-Based Systems*, vol. 56, pp. 156–166, 2014.
- [38] F. Shang, Y. Liu, J. Cheng, and D. Yan, “Fuzzy double trace norm minimization for recommendation systems,” *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 4, pp. 2039–2049, 2018.

## Research Article

# Ranking Influential Nodes in Complex Networks with Information Entropy Method

Nan Zhao <sup>1</sup>, Jingjing Bao <sup>1,2</sup> and Nan Chen <sup>1</sup>

<sup>1</sup>State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China

<sup>2</sup>College of Computer Science and Technology, Inner Mongolia Normal University, Hohhot 010022, China

Correspondence should be addressed to Nan Zhao; zhaonan@xidian.edu.cn

Received 23 March 2020; Accepted 22 April 2020; Published 8 June 2020

Guest Editor: Hongshu Chen

Copyright © 2020 Nan Zhao 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 ranking of influential nodes in networks is of great significance. Influential nodes play an enormous role during the evolution process of information dissemination, viral marketing, and public opinion control. The sorting method of multiple attributes is an effective way to identify the influential nodes. However, these methods offer a limited improvement in algorithm performance because diversity between different attributes is not properly considered. On the basis of the k-shell method, we propose an improved multiattribute k-shell method by using the iterative information in the decomposition process. Our work combines sigmoid function and iteration information to obtain the position index. The position attribute is obtained by combining the shell value and the location index. The local information of the node is adopted to obtain the neighbor property. Finally, the position attribute and neighbor attribute are weighted by the method of information entropy weighting. The experimental simulations in six real networks combined with the SIR model and other evaluation measure fully verify the correctness and effectiveness of the proposed method.

## 1. Introduction

Multidimensional information flows rapidly on the network, while different nodes have different effects on information transmission [1], viral marketing [2], public opinion guidance [3], and social recommendation [4, 5] due to their different influences. From the perspective of information transmission, different social networks have different modes of information transmission because of the diversity of functional focuses and user structures. From the perspective of marketing, by providing influential user rankings for different hobbies and groups, it can help new users quickly and effectively obtain relevant information sources of interest so as to achieve a smooth cold start. From the perspective of public opinion guidance and control, the event evolution process of hot public opinions often includes the forwarding and comments of users with different influences on different platforms. These simple operations often lead to an enormous development of public opinions in different directions.

The influence of nodes is evaluated from global structure information, such as betweenness centrality [6], closeness centrality [7], and Katz centrality [8]. These methods show good performance in node sorting. However, because of  $O(n^2)$  or even higher computational complexity, these methods are not suitable for large-scale networks. The influence of nodes is quantified by local information, such as degree centrality [9], semilocal centrality [10], hybrid degree centrality [11], average shortest path centrality [12], and  $h$  index [13]. Local measures are less efficient because they only consider local neighborhood information. There are many heuristic algorithms [14, 15] combined with local neighborhood information. Research based on random walk evaluates the influence of nodes through multiple iterative operations with high-computational complexity such as feature vector centrality [16], PageRank [17], LeaderRank [18], and Hits [19].

Kitsak et al. [20] argues that the most influential nodes are those located at the core of the network. Each node is assigned a fixed shell value after k-shell decomposition.

However, k-shell decomposition tends to assign the same shell value to many nodes so that the influence of these nodes with same shell value cannot be further distinguished. On this basis, plenty of methods have been proposed to further improve the performance of k-shell method. Zeng and Zhang [21] propose a mixed degree decomposition method, which combines the residual degree and depletion degree to update the nodes. In each step of decomposition, the nodes are removed and decomposed based on the mixed degree. However, the  $\lambda$  parameter is difficult to optimize. Liu et al. [22] proposes an improved ranking method to generate a more differentiated ranking list. This method is realized by calculating the shortest distance between the target node and the core node of the network. The core nodes of the network are in a node set with the highest shell value in k-shell decomposition. The computational cost of this method is relatively expensive by calculating the shortest distances to the core nodes. Bae and Kim [23] propose a new measurement of neighborhood coreness centrality, which calculates the diffusion influence of nodes in the network by summing all neighborhood shell values. The influence of nodes with the same  $ks$  value can be further distinguished by using the iterative information of removal nodes to identify the position difference of nodes in the network. The degree decomposition method based on iteration factor [24] is to improve the performance of the traditional method by using the iteration information and node degree in the decomposition. In addition, some other node-sorting algorithms were introduced to improve the sorting performance.

On the basis of the k-shell method, our work makes full use of the iterative information in the decomposition process and proposes an improved multiattribute k-shell method. First, the iteration information is processed by sigmod function to obtain the position index. Then, the position attribute is captured by combining the shell value and the position index. The local information of the node is adapted to obtain the neighbor property. Finally, the position attribute and neighbor attribute are weighted by the method of information entropy weighting. In the experiment, the SIR model, Kendall coefficient, and imprecision function are used to, respectively, evaluate the propagation capability of different probabilities of propagation, the imprecision of ranking and the correlation coefficient of different probability of propagation. Furthermore, we evaluate ranking results of the proposed method by selecting seeds in influence maximization problem and measuring the ranking uniqueness and distribution. The experimental results prove that the proposed method can effectively distinguish the differences of different attributes and significantly promote the performance of identifying the influence of nodes.

The following parts are organized as follows. We briefly review the definition of related algorithms used for comparison in Section 2. In Section 3, our improved multiattribute k-shell method is proposed and a meaningful example is illustrated to show how the proposed measure works. In Section 4, we present the details of the data, the spreading model, and the evaluation measure that are used to evaluate the performance of our measure. The

experimental results are presented in Section 5. Finally, we expose the conclusion of the work in Section 6.

## 2. Related Work

Kitsak et al. proposed the k-shell method to determine the influence of nodes in the network. This method considers that the closer the node is to the center of the network, the higher the influence of the node will be. This method uses node degree to rank the importance of nodes. The following details show the decomposition principle of the k-shell method.

First, we need to remove the nodes and edges with a medium degree as 1 in the network. At this time, nodes with a degree as 1 still exist in the remaining network. We should continue to remove them until no nodes with a degree as 1 exist in the network. At this point, the removed nodes form a layer and its  $ks$  value is assigned to 1. According to the abovementioned method, continue to remove nodes with a degree value as 2 in the network and repeat this operation until there are no nodes in the network. As can be seen in Figure 1, the  $ks$  value allocated for nodes 8, 9, 10, 11, 12, 13, 14, 15, and 16 is 1. The value allocated for nodes 5, 6, and 7 is 2. The  $ks$  value allocated for nodes 1, 2, 3, and 4 is 3.

K-shell decomposition method is suitable for large networks because of low-computational complexity. Of course, the disadvantages of this method are obvious. First, most nodes are assigned the same  $ks$  value so that the importance of these nodes cannot be further distinguished. For example, from the perspective of degree, the degree of node 8 is 3 and the degree of node 11 is 1. The influence of node 8 is obviously greater than 11, but they have the same  $ks$  value. Second, in the process of removing nodes, the edges that have been removed are not considered and the influence of residue is only concerned about. In this way, it is considered that nodes with the same  $ks$  have the same number of edges in the outer layer, which is obviously not consistent with common sense. For example, node 1 and node 2 have abundant first-order and second-order neighbors in the outer layer, while node 1 has no neighbors in the outer layer. The same  $ks$  value assigned to them does not reflect the difference between them. Third, in the regular network, the  $ks$  value of most nodes is 1, which is obviously not suitable for this background.

The traditional K-shell decomposition method only updates the nodes according to the residual degree of the remaining nodes and completely ignores the depletion degree of the removed nodes. In the mixed degree decomposition method, the decomposition process is based on the residual degree of the remaining nodes and the depletion degree of the removed nodes. For node  $i$ , its residual degree and depletion degree are represented by  $k_i^r$  and  $k_i^e$ , respectively. In each step of the mixing degree decomposition method, the node removal is determined by the mixing degree  $k_i^m$ :

$$k_i^m = k_i^r + \lambda k_i^e, \quad (1)$$

where  $\lambda$  is the adjustable parameter between 0 and 1. When  $\lambda$  is 0, the mixing degree decomposition method is consistent

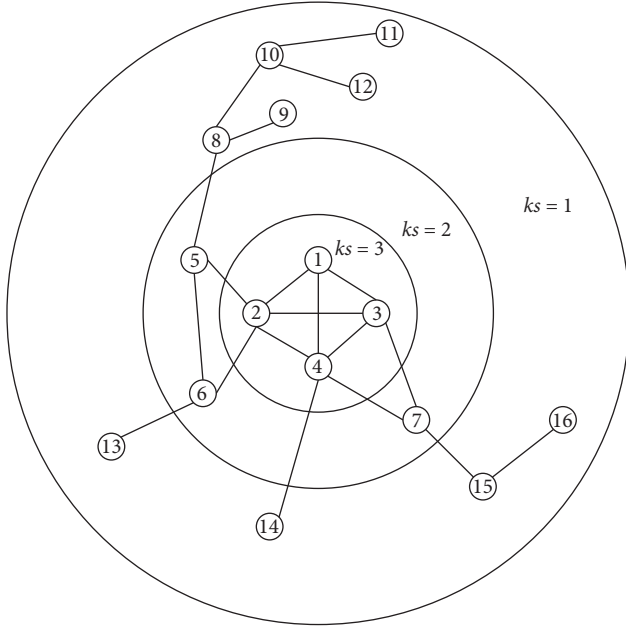


FIGURE 1: The schematic diagram of basic principles of the k-shell method.

as the K-shell decomposition method. When  $\lambda$  is 1, the mixed degree decomposition method is equivalent to the degree centrality method. Different from the traditional K-shell decomposition method, in the mixed degree decomposition method, the value of  $ks$  of all nodes can be decimal. The parameter  $\lambda$  is usually set to 0.7.

The traditional K-shell decomposition method is improved by using the shortest distance from the source node to the core node of the network. The propagation capability of nodes with the same  $ks$  value can be further distinguished by the following methods:

$$\theta(i) = (ks_{\max} - ks_i + 1) \sum_{j \in S_c} d_{ij}, \quad (2)$$

where  $ks_{\max}$  is the maximum  $ks$  value in K-shell decomposition,  $ks_i$  is the  $ks$  value of node  $i$ , and  $S_c$  is the set of nodes whose  $ks$  value is maximum. Although this nonparametric method can identify nodes with the same  $ks$  value, the computational complexity is high because of calculation of the shortest path distance to the core. More seriously, if the network is not fully connected, the shortest path distance between partial node pairs cannot be obtained.

If the iterative information is used to identify the position difference of nodes in the network, the propagation ability of nodes with the same  $ks$  value can be further distinguished. Degree decomposition method based on the iteration factor is proposed to improve the performance of the traditional method by using iteration information and node degree in decomposition. It is worth noting that the degree is a local variable and the iteration factor is a global variable. This method fully combines local and global factors to identify influential nodes more effectively. The iterative factor  $\delta_i$  of node  $i$  is

$$\delta_i = ks_i * \left(1 + \frac{\text{iter}(i)}{m}\right), \quad (3)$$

where  $m$  is the total number of iterations in K-shell decomposition and  $\text{iter}(i)$  is the number of times that node  $i$  has been removed from the decomposition. The influence of node  $i$  in the degree decomposition method based on iterative factor is

$$IC_i = \delta_i * d_i + \sum_{j \in \Gamma(i)} \delta_j * d_j. \quad (4)$$

The influence of node will be great if a node has many neighbors at the core of the network. Based on this assumption, the neighbor core of the node is

$$C_{nc}(i) = \sum_{j \in \Gamma(i)} ks(j), \quad (5)$$

where  $\Gamma(i)$  is the neighbor of node  $i$ . Recursively, the extended neighbor kernel is defined as

$$C_{nc+}(i) = \sum_{j \in \Gamma(i)} C_{nc}(j). \quad (6)$$

### 3. Materials and Methods

Practice has proved that the combination of multiple attributes can further improve the sorting effect. In recent years, researchers have combined different attributes and strategies to mine the influence of nodes. The performance of these methods proves that considering multiple attributes is an effective strategy to evaluate the impact capability of nodes. At present, there are many attribute weighting methods, such as least squares weighting method and principal component analysis method. Among the many attributes of nodes, location attributes play a significant role in the sorting process of nodes. At the same time, the influencing ability of nodes depends largely on the neighbor attributes. Combining these two attributes, it is an effective strategy to use the attribute weighting method to further identify the influence of nodes.

In the K-shell decomposition process, the number of iterations reveals very important location information, and it can further distinguish the location differences of the removed nodes. A node with a higher number of iterations is closer to the core of the network, or it is closer to the edge of the network. The number of iterations here refers to the number of global iterations decomposed from K-shell to the end. In this paper, sigmod function is used to further process the number of iterations when the node is deleted, so as to define the node position index  $p(i)$ :

$$p(i) = \frac{3}{4} \frac{1}{1 + e^{-\sqrt{\text{iter}(i)}}}. \quad (7)$$

The relationship between the position index and the number of iterations is shown in Figure 2. As the number of iterations increases, the position index increases in a downward slope with a critical value of 0.75.



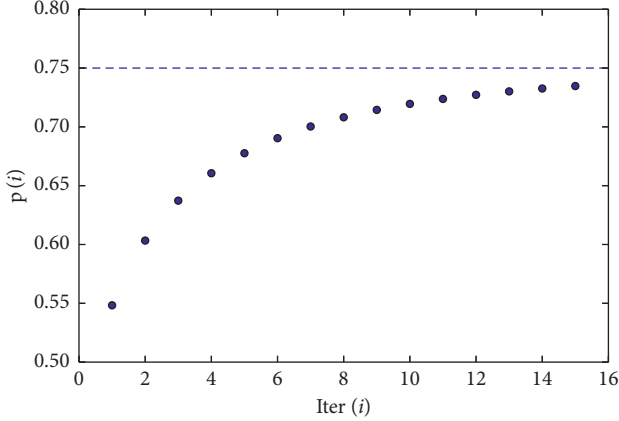


FIGURE 2: The diagram of the relationship between iteration number and position index.

The position attribute of the node is represented by  $PN_p(i)$ , which is composed of the  $ks$  value of the node and the sum of the location index of the neighbor of the node:

$$PN_p(i) = ks_i + \sum_{j \in \Gamma(i)} p(j), \quad (8)$$

where  $ks_i$  is the  $ks$  value of the node in the K-shell decomposition.

The position attribute of a node cannot effectively distinguish the influence of a node in the same position. For example, all nodes on the edge of a network have the same position information, so their influence should be the same. In fact, due to the difference of local structure, the influence of edge nodes in the same position will vary greatly. The local attributes of nodes should be further used to distinguish the influence differences of nodes with the same location attributes. If a node has more neighbors, it can have a greater impact on the network. Furthermore, the influence of a node's neighbor is also impacted by its neighbors. Consider the second-order neighbor can improve the capability of measuring the influence of a node.

The neighbor attribute of a node is represented by  $PN_n(i)$ , and it is represented by the second-order neighbor number of the node:

$$PN_n(i) = \sum_{j \in \Gamma(i)} \sum_{l \in \Gamma(j)} k_l, \quad (9)$$

where  $k_l$  represents the degree of node  $l$ .

Both position attribute and local attribute play a significant role in identifying the influence of nodes. Combining these two key attributes, the influence of nodes can be accurately calculated and the performance of node influence ranking can be further improved. Many traditional multi-attribute sorting methods treat all attribute weights to be consistent. At the same time, there are many weighting methods, such as analytic hierarchy process, multiobjective programming, principal component analysis, and weighted least square method. Information entropy weighting method is an excellent weighting method, which has been verified by

many examples. Our method adapts the method of information entropy weighting to avoid the defects of the traditional weighting method:

$$PN(i) = w_1 * PN_p(i) + w_2 * PN_n(i), \quad (10)$$

where  $w_1$  represents the weight of the location attribute and  $w_2$  represents the weight of the neighbor attribute.

The calculation process of information entropy weighting method is as follows. First, the entropy value of each attribute is calculated:

$$H_i = -\frac{1}{\ln n} \sum_{j=1}^n r_{ij} \ln r_{ij}, \quad i = 1, 2, \quad (11)$$

where  $H_i$  represents the entropy of the  $i$ th attribute, and  $r_{ij}$  represents the normalized value of the  $i$ th attribute of the  $j$ th node. Because this method has only location and neighbor attributes,  $i$  is set as 1 and 2:

$$r_{1j} = \frac{PN_p(j)}{\sum_{j=1}^n PN_p(j)}, r_{2j} = \frac{PN_n(j)}{\sum_{j=1}^n PN_n(j)}. \quad (12)$$

Then, the information entropy is combined to calculate the weight of the two attributes:

$$w_i = \frac{1 - H_i}{2 - \sum_i H_i}, \quad i = 1, 2. \quad (13)$$

Whether the multiattribute-improved K-shell algorithm proposed in this section is feasible, the feasibility can be explained with the help of diagrams. The graph is an undirected graph with 16 nodes and 20 edges. The PN value of each node is calculated according to the algorithm. The values in the calculation process are shown in Table 1. As we can see from the table, the importance of all nodes can be sorted in the descending order according to the PN value. The PN values of node 11 and node 12 are the same and their importance cannot be distinguished. The PN values of nodes 1, 2, 3, and 4 are in the first gradient, and the PN value of node 2 is the largest. Its importance can be seen in Figure 1 in the network. PN values of nodes 5, 6, 7, and 8 are located in the second gradient. These nodes are not outer edge nodes.

The PN value of node 14 is larger than that of edge nodes 9, 11, 12, 13, and 16. It can be seen from the figure that since it is directly connected to core node 4, its influence has been enhanced. From the calculation results shown in Table 1, it can be preliminarily concluded that the improved K-shell method based on multiple attributes is feasible to some extent.

## 4. Experimental Setup

**4.1. Data Description.** We conduct several experiments on six different real networks to evaluate the performance of our proposed centrality measure. The real networks are drawn from disparate fields. CA-HepTh [25] is a collaboration network of Arxiv High Energy Physics Theory. Netscience [25] is the network of co-authorship of scientists in network theory and experiments. Cond-Mat [25] is from the e-print arXiv and covers scientific collaborations

TABLE 1: Example verification results of K-shell-improved algorithm based on multiple or inner core nodes.

Node	$ks$	Iter ( $i$ )	PN $p$ ( $i$ )	PN $n$ ( $i$ )	PN ( $i$ )
1	3	5	5.032744485944649	50	36.79944673550679
2	3	5	6.353940102911473	66	48.490348105799676
3	3	5	5.693342294428061	59	43.35133818263464
4	3	5	6.241636228400565	64	47.04449874200889
5	2	4	3.9754352186156736	34	25.186011640937103
6	2	4	3.8864732377707982	32	23.747014505664833
7	2	4	3.958485252509983	36	26.593917345765504
8	1	3	2.812214004336133	19	14.247925877904956
9	1	1	1.6372559148173789	7	5.425716932665104
10	1	2	2.7338437827623867	13	9.986275003487945
11	1	1	1.6033222618802176	5	4.002873871402858
12	1	1	1.6033222618802176	5	4.002873871402858
13	1	1	1.6605978084834119	9	6.8454506851464405
14	1	1	1.6775814953148829	16	11.795521737181323
15	1	2	2.208891742455916	13	9.832170482539274
16	1	1	1.6033222618802176	4	3.2964331094214456

between authors papers submitted to Condense Matter category. DNC Email [26] is the network of emails in the 2016 Democratic National Committee email leak. Nodes in the network correspond to persons in the dataset and the edge is the mail exchange between users. Ego-Twitter [27] contains Twitter user-user following information. A node represents a user and an edge indicates that the user follows the other user. Route Views [28] is the network of autonomous systems of the Internet connected with each other. Nodes are autonomous systems (AS), and edges denote communication. A brief overview of the networks is shown in Table 2.

**4.2. Spreading Model.** To evaluate the lists ranked by all the centrality measures, we need to know the list ranked by the real spreading process of the nodes. In the spreading process, the probability of accepting a message from another user depends on the user’s influence [31]. So, the spreading efficiency of nodes is used to measure the ranking result of influential nodes. There are many information diffusion models, such as SIR model, Independent Cascade model, and Linear Threshold model, and some information diffusion models independent of network topology [32, 33]. In this paper, we employ the standard SIR model [34] to simulate the spreading process on networks and record the spreading efficiency for every node. In the SIR model, every node belongs to one of the susceptible states, the infected state or the recovered state. In detail, we set one node as an infected node and the other nodes are susceptible nodes. At each step, for every infected node, it can infect its susceptible neighbors with infection probability  $\beta$  and then can be removed with probability  $\lambda$ . Generally, we set  $\lambda = 1.0$ . The appropriate propagation probabilities are needed to be chosen, in case too small or too large propagation probability makes the propagation effect not ideal and leads to the failure to distinguish the influence of nodes. According to the heterogenous mean-field method, the epidemic threshold of network is  $\beta_{th} = \langle k \rangle / (\langle k^2 \rangle - \langle k \rangle)$ .  $k$  and  $k^2$  are degree and second-order degree of node. The

TABLE 2: The basic topological features of four real network datasets.

Network	$N$	$M$	$K$	$K_{max}$	$C$	$r$	$\beta_{th}$	$\beta$
CA-HepTh	9877	51971	5.26	65	0.47	0.268	0.087	0.12
Netscience	379	914	4.82	43	0.74	-0.082	0.142	0.25
Cond-Mat	16264	47595	5.85	107	0.62	0.185	0.084	0.14
DNC Email	2029	5598	4.72	404	8.9	-0.307	0.014	0.08
Ego-Twitter	23370	33101	2.83	239	2.15	-0.478	0.027	0.08
Route Views	6474	13895	4.3	1459	0.96	-0.182	0.007	0.06

$N$  and  $M$  are the numbers of nodes and edges, respectively.  $K$  and  $K_{max}$  denote the average degree and the maximum degree.  $C$  and  $r$  are the clustering coefficient [29] and assortative coefficient [30].  $\beta_{th}$  and  $\beta$  are the epidemic threshold of network and the infection probability used in our experiment.

propagation probabilities are set just larger than the epidemic threshold. In the experiment, this dynamical process of infection and recovering will repeat until there are no infected nodes. The sum of infected and recovered nodes at time  $t$ , denoted by  $F(t)$ , can be considered as an indicator to evaluate the influence of the initially infected node at time  $t$ . Obviously,  $F(t)$  increases with the increasing of  $t$  and will reach stable state at time  $t_c$ , denoted by  $F(t_c)$ , where  $t_c$  represents the final time and  $F(t_c)$  represents the eventual influence of the initially infected node. To guarantee the reliability of the results, all of them are averaged over a large number of realizations.

**4.3. Evaluation.** In order to evaluate the performance of the centrality measures, we use Kendall’s coefficient  $\tau$  [35] to measure the correlation between one topology-based ranking list and the one generated by the SIR model, which is approached by a large number of simulations. Let  $(x_i, y_i)$  and  $(x_j, y_j)$  be a randomly selected pair of joint from two ranking list,  $X$  and  $Y$ , respectively. If both  $(x_i > x_j)$  and

$(y_i > y_j)$  or if both  $(x_i < x_j)$  and  $(y_i < y_j)$ , they are said to be concordant. If  $(x_i > x_j)$  and  $(y_i < y_j)$  or  $(x_i < x_j)$  and  $(y_i > y_j)$ , they are said to be discordant. If  $(x_i = x_j)$  or  $(y_i = y_j)$ , the pair is neither concordant nor discordant. Kendall's coefficient  $\tau$  is defined as

$$\tau = \frac{n_c - n_d}{0.5n(n-1)}, \quad (14)$$

where  $n_c$  and  $n_d$  denote the number of concordant and discordant pairs, respectively. The value  $\tau$  lies between +1 and -1. The higher the  $\tau$  value indicates, the more accurate ranked list a centrality measure could generate. The most ideal case is  $\tau = 1$ , where the ranked list generated by the centrality measure is exactly the same as the ranked list generated by the real spreading process.

To measure the imprecision [17] of methods in ranking influential nodes, we compare propagation capability of influential nodes obtained by ranking result with nodes which have largest propagation capability and the propagation capability of nodes generated from the SIR model. Kendall's correlation coefficient considers the correlation between the ranking order of all nodes in the network and the order of propagation capability of all nodes. However, the imprecision function is used to evaluate the cumulative propagation capability of top-ranked nodes in different proportions. The imprecision function  $\varepsilon(p)$  is defined as

$$\varepsilon(p) = 1 - \frac{\sum_{i \in \phi_m(p)} F_i(t_c)}{\sum_{j \in \phi_s(p)} F_j(t_c)}, \quad (15)$$

where  $p$  is the proportion of top-ranked nodes and  $\phi_m(p)$  and  $\phi_s(p)$  denote the set of nodes at the top when proportion is  $p$ .  $F_i(t_c)$  is propagation capability of node  $i$ . The value of  $\varepsilon$  lies between 0 and 1. The lower  $\varepsilon$  value indicates, the more precise the centrality measure is in ranking propagation capability of node.

## 5. Results and Discussion

In this chapter, SIR model, Kendall coefficient, and imprecision function are used to verify the performance of the proposed method. Degree, Ks (K-shell), MDD (Mixed Degree Decomposition), ksIF (K-shell iteration factor method), and Cnc+ (Extended Neighborhood coreness centrality measure) were compared with the proposed multiattribute PN method.

**5.1. Evaluate the Spreading Capability of Nodes.** This section verifies that different propagation probabilities are selected under the fixed proportion of transmission sources to calculate the propagation capability of the influence node set. By comparing the propagation capability of different algorithms under the same propagation probability, the performance of different methods can be compared. The simulation set the propagation source ratio as 0.1. The max time step of infection process is set as 100, but the infection will stop when there is no user in the infection state. The results were based on an average of 500 independent experiments. The simulation results are shown in Figure 3. The

abscissa is the propagation probability and the ordinate is the propagation capability, which is expressed as a percentage.

Firstly, with the increasement of the propagation probability, the node propagation ability of the six methods is improved. The performance of the six methods is relatively close when the propagation probability is small and obviously different when the propagation probability is large. Specifically, in the CA-HepTh, DNC Email, and Cond-Mat, the performance of PN maintained the best under various propagation probabilities. In the Ego-Twitter dataset, the performance of PN and Cnc+ is significantly higher than that of the other four methods. Moreover, when the transmission probability is 0.03, the node transmission ability of Cnc+ is higher, and in other cases, the node transmission ability of PN method is higher. In the Netscience and Route Views dataset, the performance of PN is relatively better than the other method, but Cnc+ outperforms the PN method when the transmission probability is less than 0.1 in the Netscience. PN is not the best in several points in Route Views, but it had highest propagation ability in most cases, as shown in Figure 3(f). In general, the proposed PN in this chapter has better performance. The difference in performance is related to the specific network structure.

**5.2. Evaluate the Imprecision of Ranking.** This section verifies that different proportions of top-ranked nodes are selected under the fixed propagation probability to calculate the imprecision of ranking the influence of nodes. By comparing the imprecision of different algorithms under the same proportion of top-ranked, the performance of different methods in ranking the influence of nodes can be distinguished. The simulation set the propagation probability of node, as shown in Table 2, and the proportion top-ranked nodes varies from 0.01 to 0.2. The max time step of infection process set as 100 and the results were based on an average of 500 independent experiments. Simulation results in six networks are shown in Figure 4. The abscissa is the proportion of top-ranked nodes that are considered and the ordinate is the imprecision of ranking.

In Figure 4(a), the method PN, KsIF, and Cnc+ have low imprecision when  $p$  is higher than 0.02 and they have poor performance in identifying the 2% top-ranked nodes. However, the imprecision of PN is much lower than the other two in CA-HepTh dataset. The imprecision of Degree to rank top 1% nodes is low about 0.07 than the method PN. However, Degree has highest  $\varepsilon$  when  $p$  is greater than 0.04. In Netscience and Cond-Mat, Degree, K-shell, and MDD have higher value, and PN is lowest in most  $p$  except that  $p$  is in the range of 0.02 to 0.04, KsIF outperforms PN in Netscience. When  $p$  is larger than 0.04 and smaller than 0.06, Cnc+ outperforms PN in Cond-Mat. In the dataset, DNC Email, K-shell, and PN are better compared with other method, and K-shell has a slight edge when  $p$  between 0.04 and 0.1 or 0.13 and 0.15. However, K-shell performs poor in 3% top-ranked nodes. In the dataset, Ego-Twitter and Route Views, it is obvious that the method we proposed performs

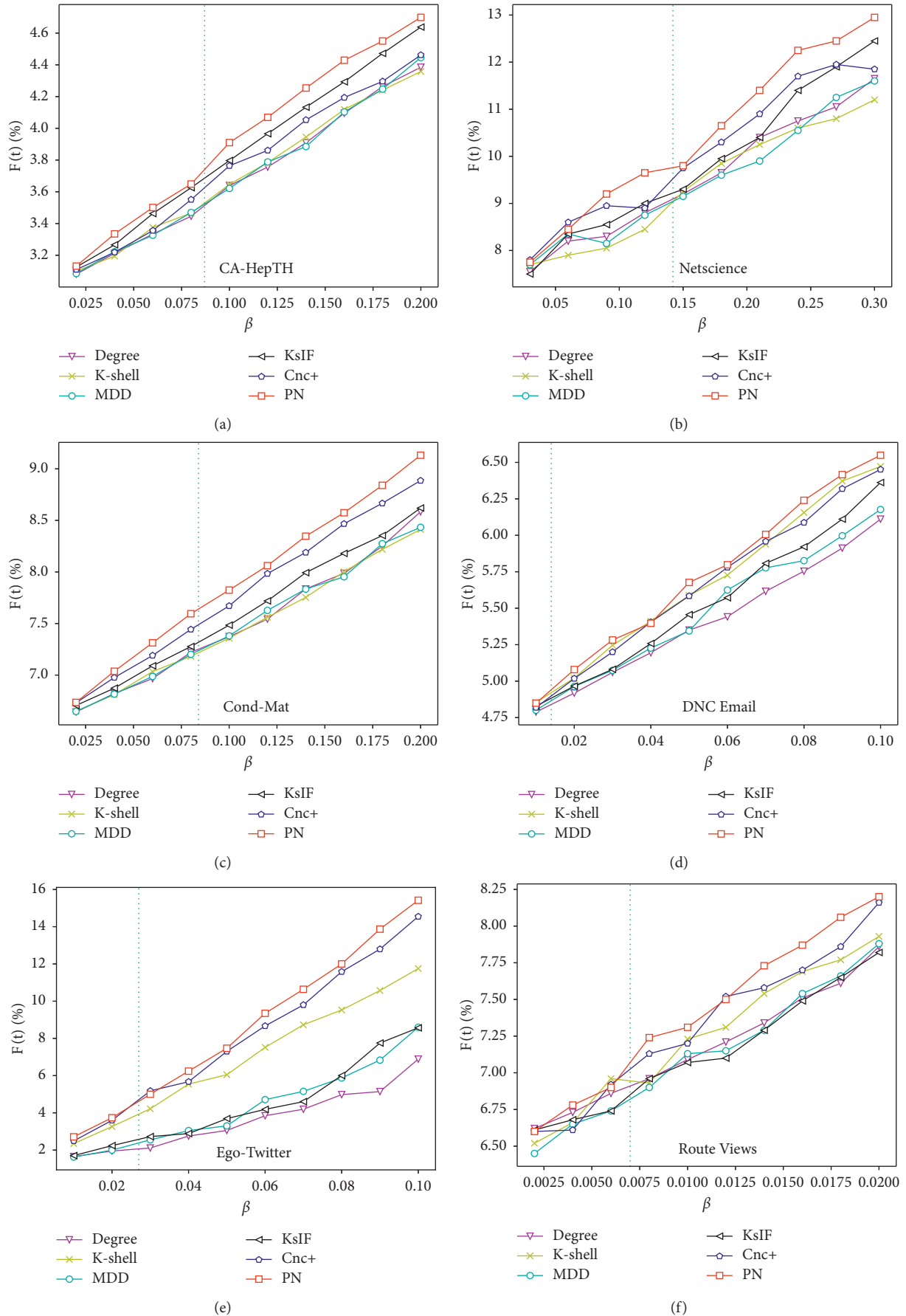


FIGURE 3: The propagation capability graph of the five methods under different propagation probabilities. The experiments are simulated on six different datasets: CA-HepTh (a), Netscience (b), Cond-Mat (c), DNC Email (d), Ego-Twitter (e), and Route Views (f). The green vertical line is the epidemic threshold of the corresponding network.

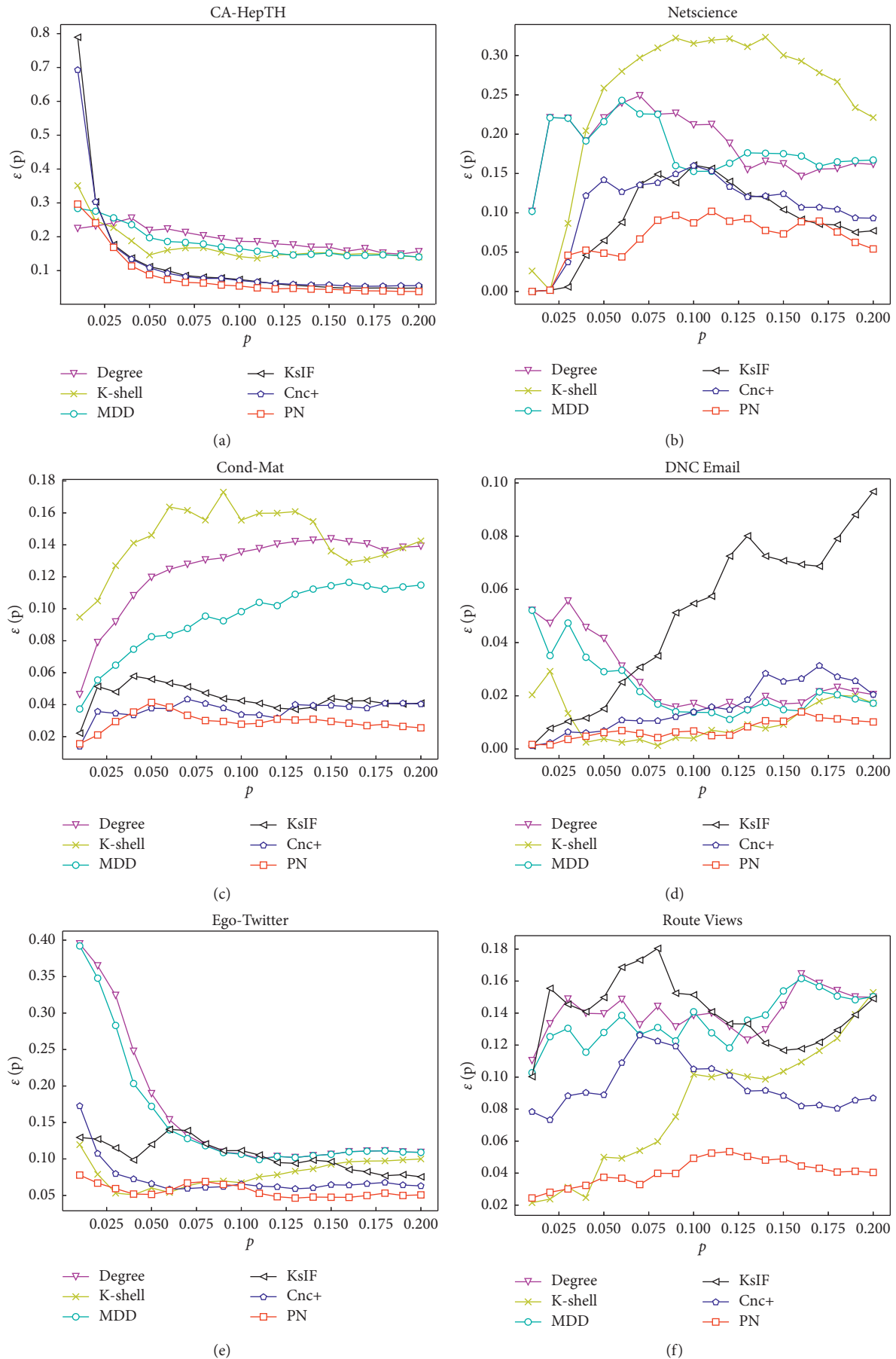


FIGURE 4: The imprecision of six ranking methods with different proportion of top-ranked nodes. The experiments are simulated on six different datasets: CA-HepTh (a), Netscience (b), Cond-Mat (c), DNC Email (d), Ego-Twitter (e), and Route Views (f).

well and is more stable in all cases. From the simulation results of the abovementioned six datasets, it can be seen that the method PN can not only precisely identify most of the top 1% to 4% nodes but also rank the following nodes by their influence stably.

*5.3. Evaluate the Correlation Coefficients of Method.* This section verifies the correlation between the influence value and propagation ability of nodes under different propagation probabilities by Kendall's coefficient  $\tau$ . The value range of propagation probability in this simulation is 0.02 to 0.2 in the datasets CA-HepTh, Cond-Mat, DNC Email, Ego-Twitter, and Route Views, and extended to 0.3 in the dataset Netscience because of higher epidemic threshold of it. In the DNC Email and Route Views, we also put the simulation result when the value of propagation probability is equal to their epidemic threshold. The max time step of infection process is 100. The results were based on an average of 500 independent experiments. The simulation results are shown in Figure 5. The abscissa is the propagation probability and the ordinate is the correlation coefficient. The epidemic threshold of each network is also drawn into the figure as a vertical line.

The correlation of ranking result of different methods and real propagation abilities obtained by the SIR model has different manifestations in different datasets, but the general trend is the same. When the propagation probability is small, the ranking correlation of Degree, K-shell, and MDD with real propagation ability is higher than the other three methods and the thresholds are roughly the same as the epidemic threshold  $\beta_{th}$  of network. It can be seen in Figure 5 that the Degree has high Kendall's coefficient  $\tau$  when  $\beta$  is less than 0.06 in the CA-HepTh and Cond-Mat and less than 0.07 in the Netscience, and it is not obvious in other three networks because of the small epidemic threshold. However, in the cases where  $\beta$  is higher than  $\beta_{th}$ , these methods that have considered both degree and other aspects such as the influence of neighbors to the spreading ability of nodes perform better than the method that mainly take degree into account. This is because the very small propagation probability will make the infection behaviour in a small range near the initial infected node, unable to spread over a large area, so the degree is the decisive factor. As the probability of propagation increases, the act of infection becomes easier, so the extent of infection depends not only on the number of neighbors of a node but also on the ability of its neighbors to propagate, or even the neighbors' neighbors. On the contrary, when the propagation probability increases to a point much larger than  $\beta_{th}$  of network, infection becomes too easy so that the nodes in the core of the network or with high degree had largest scope of infection. It is clearly shown in the networks with small epidemic threshold such as DNC Email and Route Views in the figure. When  $\beta > \beta_{th}$ , the method PN has the highest Kendall's coefficient  $t$  in CA-HepTh, Netscience, and Cond-Mat. Our proposed method also performs best when  $\beta$  is between 0.06 to 0.14 in DNC Email and between 0.06 to 0.18 in Route Views. In short, compared with the other five methods, the PN method

proposed in this chapter has good performance under the appropriate propagation probability value range.

*5.4. Evaluate the Performance of Selecting Seeds in Influence Maximization Problem.* This section verifies the reliability of PN when it is applied in the influence maximization problem. Maximization of influence is widely used in real life. For example, viral marketing is a typical application that can promote new products or ideas for merchants or publicity departments. It aims to select a group of nodes in the network called seed node set as initial propagation nodes and spread in the network as widely as possible according to a certain diffusion mode. There are many related studies. We use the selection of top from the ranking to specify seed nodes to measure the propagation range of seed nodes selected by several ranking methods. We examine the spreading efficiency of different seed node set with six real networks: CA-HepTh, Netscience, Cond-Mat, DNC Email, Ego-Twitter, and Route Views.

In the experiment, we set  $P$  as the proportion of seed nodes in the whole network, ranging from 0.01 to 0.05. We also used the SIR Model to simulate the propagation process and calculated the influence range by using the number of users who were finally infected. The propagation probability in the simulation is determined according to the epidemic threshold and is shown in Table 2. The max time step of infection process is set as 100 and the results were based on an average of 500 independent experiments. The simulation results are shown in Figure 6. The abscissa is the ratio of the seed set in all users, and the ordinate is the propagation scope of initial seed nodes and expressed as a percentage.

The results manifest that the PN method performs best in the datasets Cond-Mat, DNC Email, Ego-Twitter, and Route Views. In the CA-HepTh dataset, the seed set selected by the method PN when  $P$  is greater than 0.02 can infect wilder than others. When  $P$  is 0.02 and below, degree is most effective and the PN is better than KsIF and Cnc+. In the Netscience network, KsIF is best when  $P$  is 0.03 and 0.04, and the infected rate of seeds selected by PN is almost same as KsIF in other cases.

*5.5. Measure the Ranking Uniqueness and Distribution.* This section verifies the monotonicity of the new method on the sorting by using Bae and Kim's ranking monotonicity method [23]. Since the K-shell method will calculate the K-shell value of many nodes into the same value, it is difficult to distinguish their differences in influence. In this respect, the method we proposed can do better. According to the definition of Bae and Kim, the monotonicity of the ranking result is expressed as follows:

$$M(R) = \left[ 1 - \frac{\sum_{r \in R} n_r (n_r - 1)}{n(n-1)} \right]^2. \quad (16)$$

In (16),  $R$  represents the ranking list and  $n$  is the size of it, every element of the ranking list is a set of nodes that have the same ranking value, and  $n_r$  is the number of nodes in  $R$  that have the same ranking position  $r$ . The

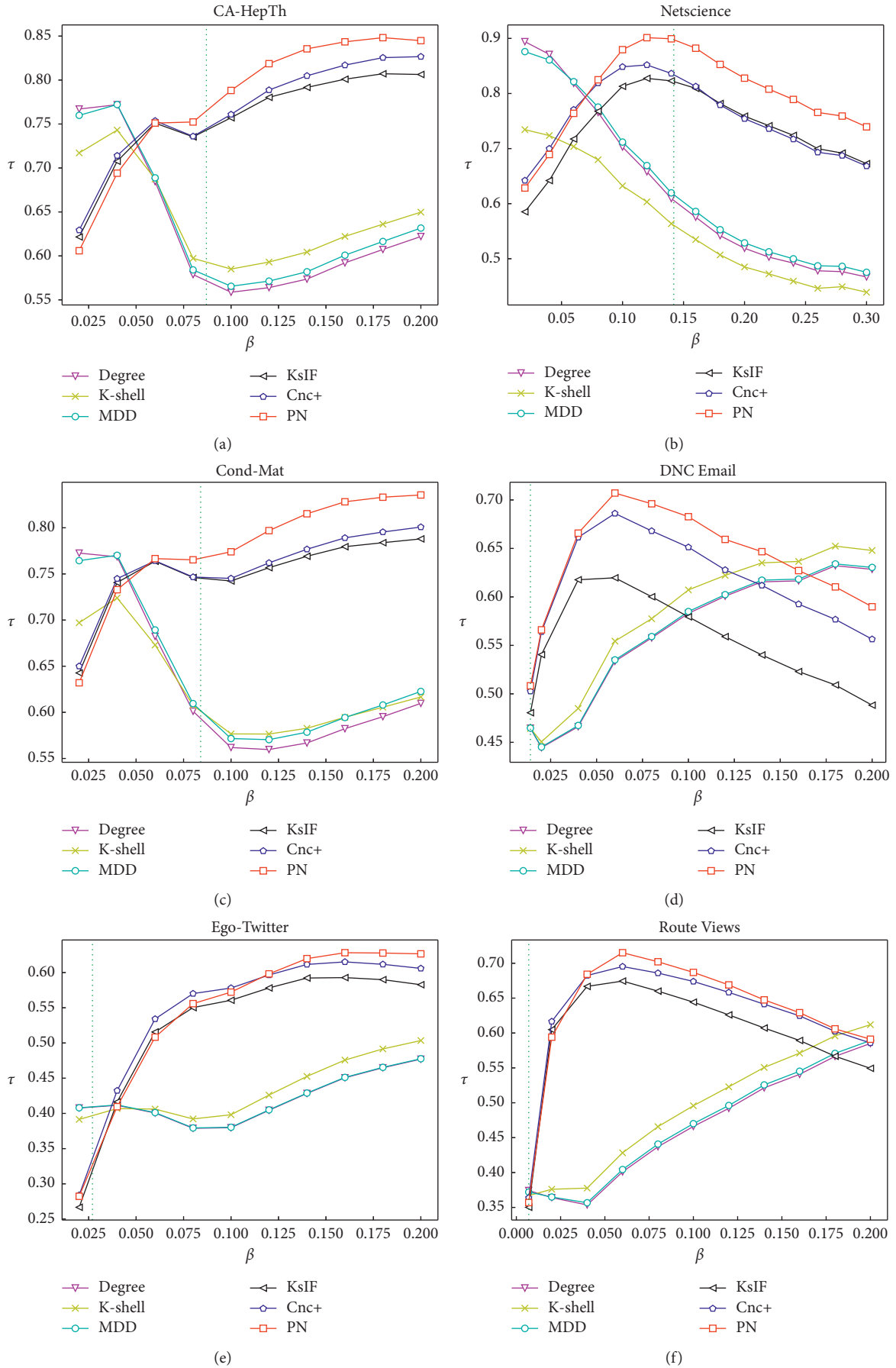


FIGURE 5: The correlation coefficient variation diagram of the six methods under different propagation probabilities. The experiments are simulated on six different datasets: CA-HepTh (a), Netscience (b), Cond-Mat (c), DNC Email (d), Ego-Twitter (e), and Route Views (f). The green vertical line is the epidemic threshold of the corresponding network.

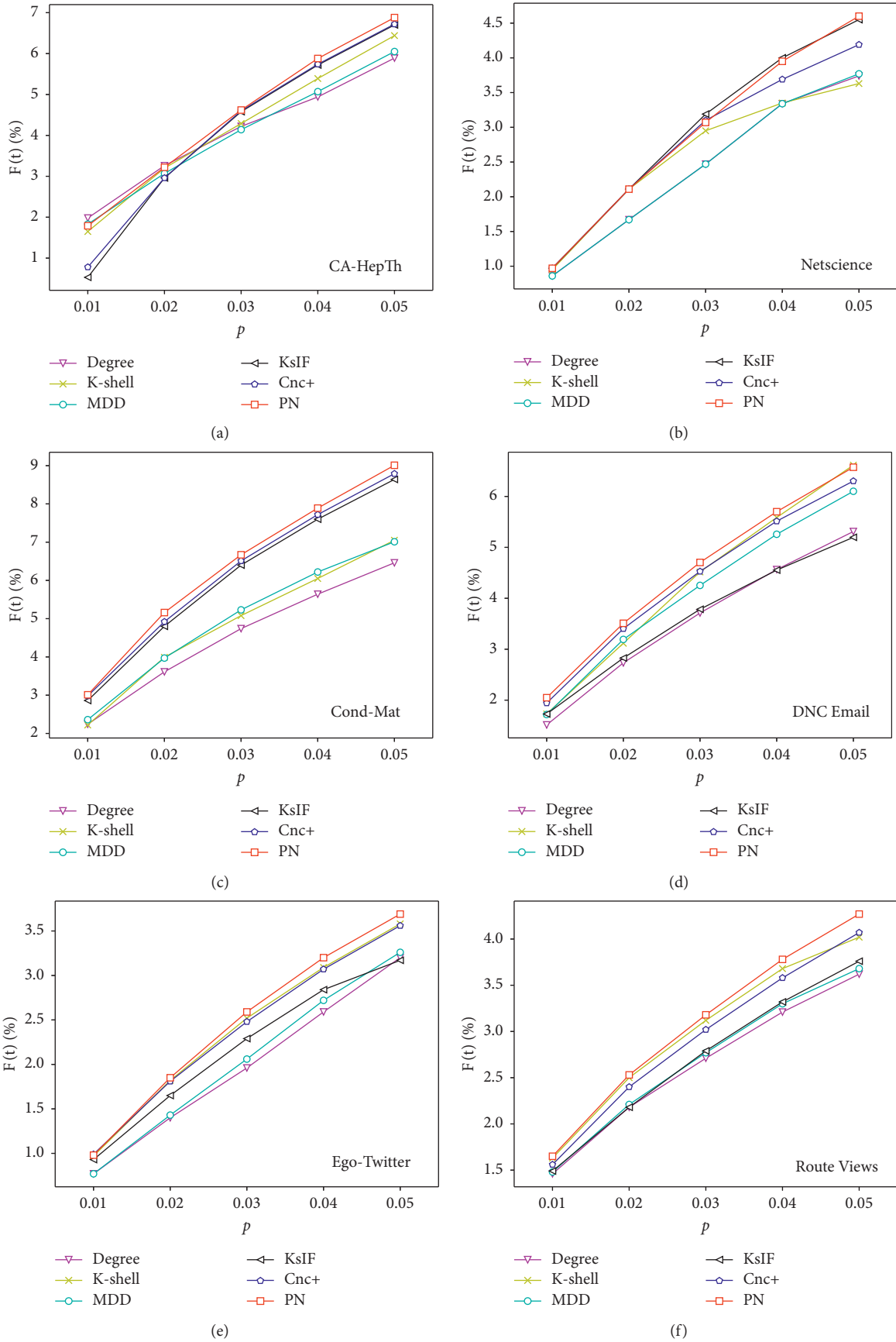


FIGURE 6: The performance of selecting seeds of the six methods in influence maximization problem. The experiments are simulated on six different datasets: CA-HepTh (a), Netscience (b), Cond-Mat (c), DNC Email (d), Ego-Twitter (e), and Route Views (f).



TABLE 3: The monotonicity  $M(R)$  of six real networks,  $R$  is representing the ranking vector of different methods.

Network	$M$ (Degree)	$M$ (Ks)	$M$ ( $MDD$ )	$M$ (KsIF)	$M$ (Cnc+)	$M$ (PN)
CA-HepTh	0.762683	0.674237	0.81057	0.992416	0.988876	0.993566
Netscience	0.764206	0.642083	0.821527	0.994702	0.989307	0.995036
Cond-Mat	0.809054	0.740879	0.86048	0.994138	0.991407	0.994788
DNC Email	0.339762	0.32344	0.344565	0.936082	0.936028	0.936465
Ego-Twitter	0.131463	0.122307	0.131656	0.995574	0.991322	0.996321
Route Views	0.587174	0.549622	0.608671	0.993285	0.992548	0.993893

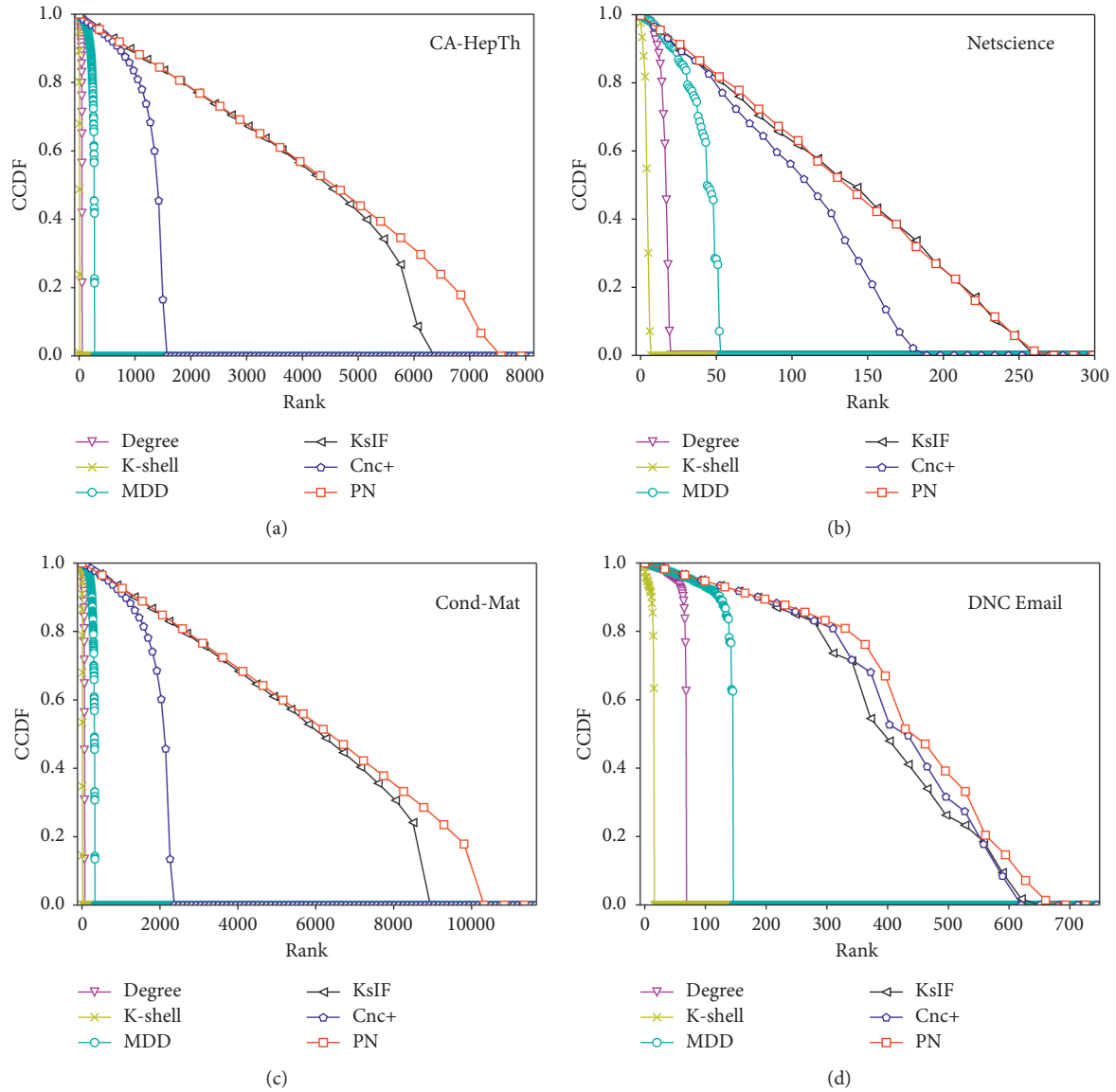


FIGURE 7: Continued.

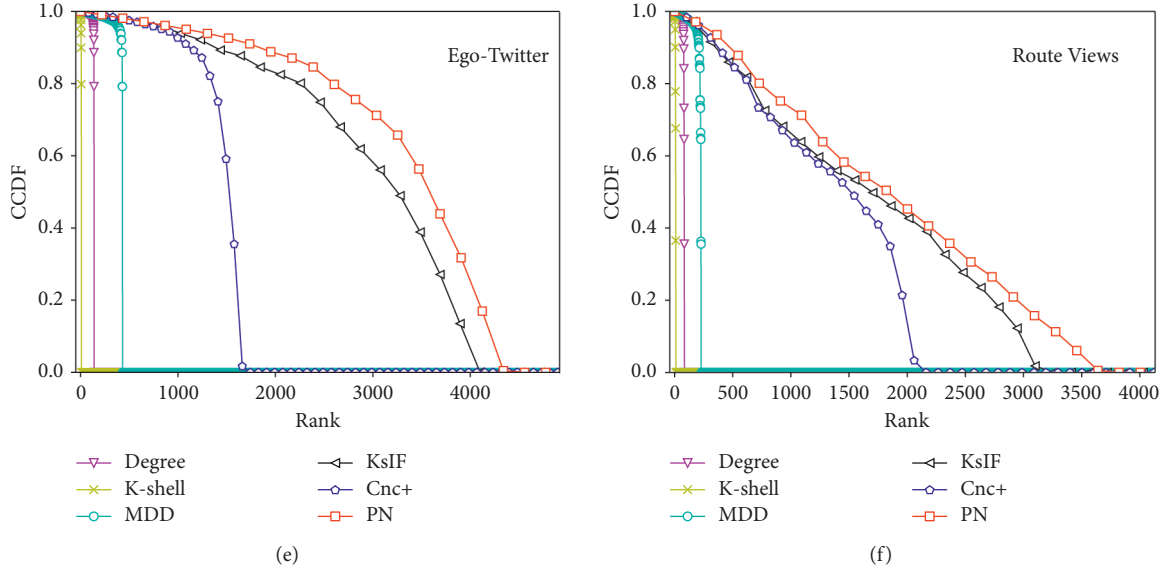


FIGURE 7: The complementary cumulative distribution (CCDF) of the six methods in six different datasets: CA-HepTh (a), Netscience (b), Cond-Mat (c), DNC email (d), Ego-Twitter (e), and Route Views (f).

value of  $M(R)$  fluctuates between  $[0, 1]$  and the higher the value, the stronger the uniqueness. In extreme cases, 1 means that each node is assigned a different sort value, whereas 0 is the opposite and all nodes are in the same rank.

We examine the monotonicity of different methods with the same six datasets as above. The calculation results are shown in Table 3. It can be seen from the table that the monotonicity result of the PN is apparently higher than the degree,  $ks$  and MDD, and approximate to KsIF and Cnc+.

In order to clarify the ranking distribution of the different measures more clearly, a complementary cumulative distribution function (CCDF) is plotted. According to the CCDF principle, if many nodes are in the same rank, the CCDF plot will decrease rapidly; otherwise, the CCDF plot will slow down. Figure 7 shows the ranking distribution in six networks.

The line representing Degree,  $ks$ , or MDD drop sharply, as can be seen on the left side of each graph. This is especially true when the number of nodes in the dataset is large. For the method PN, the ranking distribution is slightly improved compared with Cnc+ and KsIF in datasets DNC Email and Route Views. The curves of KsIF and PN in the dataset Netscience are basically overlapping and drop off more slowly than that of Cnc+. In the dataset CA-HepTh and Cond-Mat, the method Cnc+ also does not perform well compared to the method KsIF and PN. The KsIF and PN are equally good at identifying the influential nodes, while in the latter part of the ranking, the downward trend of KsIF curve is more obvious than that of PN curve. It is indicated that the ability of method PN to distinguish the nodes' spreading capability is better than Cnc+. It can be seen in the Ego-Twitter that the performance of PN is better than the method Cnc+. So, we can say that PN performs well in most networks.

## 6. Conclusions

On the basis of the K-shell method, we proposed a new multiattribute ranking method based on node position and neighborhood. We made full use of the iterative information in the decomposition process. First, the iteration information is processed by sigmod function to obtain the position index. The position attribute is obtained by combining the shell value and the position index. Then, the local information of the node is adopted to obtain the neighbor property. Furthermore, the position attribute and neighbor attribute are weighted by the method of information entropy weighting. Finally, we evaluated the propagation capability of different propagation probabilities, the imprecision of different proportions, the correlation coefficient of different propagation probabilities, and the propagation capability of selected seed nodes in influence maximization problem. At the same time, we also verified the good performance of our method in distinguishing influence of nodes. Compared with other K-shell decomposition and its improved algorithms, the method proposed in this paper had better performance. Through simulation experiments, it is found that the PN method can make full use of the iterative information in the decomposition process and the influence of neighbors to further distinguish the difference of nodes with the same  $ks$  value. Experiments with SIR model, Kendall's coefficient, and imprecision function fully verified the correctness and effectiveness of the proposed method. In a word, the effectiveness of the proposed method in the identification of influence nodes was verified by various forms of experiments.

## Data Availability

Previously reported network datasets were used to support this study and are available at <http://networkrepository.com>

and <http://konect.uni-koblenz.de>. These datasets are cited at relevant places within the text as references [22–25].

## 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 Key Industry Projects in Shaanxi Province (Grant no. 2019ZDLGY09-03), Natural Science Foundation of Shaanxi Province (Grant nos. 2018JM6053 and 2018JZ6006), National Natural Science Foundation of China (Grant nos. 61372076, 61301171, and 61771296), and 111 Project (Grant no. B08038).

## References

- [1] L. Gao, W. Wang, P. Shu, H. Gao, and L. A. Braunstein, "Promoting information spreading by using contact memory," *EPL (Europhysics Letters)*, vol. 118, no. 1, p. 18001, 2017.
- [2] V. P. T. Menta and P. K. Singh, "Efficient selection of influential nodes for viral marketing in social networks," in *Proceedings of the IEEE International Conference on International Conference on Current Trends in Advanced Computing*, Bangalore, India, March 2017.
- [3] Y. Cho, J. Hwang, and D. Lee, "Identification of effective opinion leaders in the diffusion of technological innovation: a social network approach," *Technological Forecasting and Social Change*, vol. 79, no. 1, pp. 97–106, 2012.
- [4] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [5] F. Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2018.
- [6] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
- [7] G. Sabidussi, "The centrality index of a graph," *Psychometrika*, vol. 31, no. 4, pp. 581–603, 1966.
- [8] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.
- [9] L. C. Freeman, "Centrality in social networks' conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1979.
- [10] B. Tian, J. Hu, and Y. Deng, "Identifying influential nodes in complex networks based on AHP," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 4, pp. 1777–1787, 2017.
- [11] Q. Ma and J. Ma, "Identifying and ranking influential spreaders in complex networks with consideration of spreading probability," *Physica A: Statistical Mechanics and Its Applications*, vol. 465, pp. 312–330, 2017.
- [12] Z. Lv, N. Zhao, F. Xiong, and N. Chen, "A novel measure of identifying influential nodes in complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 523, pp. 488–497, 2019.
- [13] L. Lü, T. Zhou, Q. Zhang et al., "The H-index of a network node and its relation to degree and coreness," *Nature Communications*, vol. 7, no. 1, p. 10168, 2016.
- [14] A. Sheikhhahmadi and M. A. Nematbakhsh, "Identification of multi-spreader users in social networks for viral marketing," *Journal of Information Science*, vol. 43, no. 3, pp. 412–423, 2017.
- [15] X. Wang, X. Zhang, C. Zhao et al., "Maximizing the spread of influence via generalized degree discount," *Plos One*, vol. 11, no. 10, Article ID e0164393, 2016.
- [16] P. Bonacich, "Some unique properties of eigenvector centrality," *Social Networks*, vol. 29, no. 4, pp. 555–564, 2007.
- [17] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1–7, pp. 107–117, 1998.
- [18] L. Lu, Y. C. Zhang, C. H. Yeung et al., "Leaders in social networks, the delicious case," *PLoS One*, vol. 6, no. 6, Article ID e21202, 2011.
- [19] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM)*, vol. 46, no. 5, pp. 604–632, 1999.
- [20] M. Kitsak, L. K. Gallos, S. Havlin et al., "Identification of influential spreaders in complex networks," *Nature Physics*, vol. 6, no. 11, pp. 888–893, 2010.
- [21] A. Zeng and C.-J. Zhang, "Ranking spreaders by decomposing complex networks," *Physics Letters A*, vol. 377, no. 14, pp. 1031–1035, 2013.
- [22] J.-G. Liu, Z.-M. Ren, and Q. Guo, "Ranking the spreading influence in complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 392, no. 18, pp. 4154–4159, 2013.
- [23] J. Bae and S. Kim, "Identifying and ranking influential spreaders in complex networks by neighborhood coreness," *Physica A: Statistical Mechanics and Its Applications*, vol. 395, no. 4, pp. 549–559, 2014.
- [24] Z. Wang, Y. Zhao, J. Xi, and C. Du, "Fast ranking influential nodes in complex networks using a k-shell iteration factor," *Physica A: Statistical Mechanics and Its Applications*, vol. 461, pp. 171–181, 2016.
- [25] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, TX, USA, January 2015.
- [26] *DNC Emails Network Dataset*, KONECT, Landau Germany, 2017.
- [27] J. Leskovec and J. McAuley, "Learning to discover social circles in ego networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, pp. 548–556, Lake Tahoe, NE, USA, December 2012.
- [28] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph Evolution: densification and shrinking diameters," *ACM Transactions on Knowledge Discovery from Data*, vol. 1, no. 1, 2007.
- [29] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [30] M. E. J. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, Article ID 208701, 2002.
- [31] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Transactions on Cybernetics*, pp. 1–14, 2019.
- [32] S. Fang, N. Zhao, N. Chen, F. Xiong, and Y. Yi, "Analyzing and predicting network public opinion evolution based on group persuasion force of populism," *Physica A: Statistical Mechanics and Its Applications*, vol. 525, pp. 809–824, 2019.
- [33] N. Zhao, S. Fang, N. Chen, and C. Pei, "Modeling and analyzing the influence of multi-information coexistence on attention," *IEEE Access*, vol. 7, pp. 117152–117164, 2019.

- [34] L. Guan, D. Li, K. Wang, and K. Zhao, "On a class of nonlocal SIR models," *Journal of Mathematical Biology*, vol. 78, no. 6, pp. 1581–1604, 2019.
- [35] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1-2, pp. 81–93, 1938.

## Research Article

# Effective Data Transmission and Control Based on Social Communication in Social Opportunistic Complex Networks

**Weiyu Yang, Jia Wu , and Jingwen Luo**

*School of Computer Science and Engineering, Central South University, Changsha 410083, China*

Correspondence should be addressed to Jia Wu; [jjawu5110@163.com](mailto:jjawu5110@163.com)

Received 7 March 2020; Revised 26 April 2020; Accepted 6 May 2020; Published 8 June 2020

Guest Editor: Fei Xiong

Copyright © 2020 Weiyu 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.

In opportunistic complex networks, information transmission between nodes is inevitable through broadcast. The purpose of broadcasting is to distribute data from source nodes to all nodes in the network. In opportunistic complex networks, it is mainly used for routing discovery and releasing important notifications. However, when a large number of nodes in the opportunistic complex networks are transmitting information at the same time, signal interference will inevitably occur. Therefore, we propose a low-latency broadcast algorithm for opportunistic complex networks based on successive interference cancellation techniques to improve propagation delay. With this kind of algorithm, when the social network is broadcasting, this algorithm analyzes whether the conditions for successive interference cancellation are satisfied between the broadcast links on the assigned transmission time slice. If the conditions are met, they are scheduled at the same time slice, and interference avoidance scheduling is performed when conditions are not met. Through comparison experiments with other classic algorithms of opportunistic complex networks, this method has outstanding performance in reducing energy consumption and improving information transmission efficiency.

## 1. Introduction

With the development of mobile communications in recent years, opportunistic complex networks appear in people's vision. The difference between opportunistic complex networks and the traditional mesh network mainly lying in their nodes deployment is not uniform, the network size and initial location of the nodes are not set in advance, and there is no need for a complete path between the source node and the target node [1–4]. As nodes move, two nodes can enter the communication range and exchange data [5]. Opportunistic complex networks can handle problems that are difficult to solve with existing wireless technology networks, such as network splits and delays, and can meet the network communication that requires low-cost and terrible conditions [6, 7]. In fact, part of the concept of opportunistic complex networks is derived from the early research on delay tolerant network (DTN) [8]. At present, the main application fields of opportunistic complex networks include wildlife tracking, handheld device networking, in-vehicle network, and remote area network transmission [9, 10].

Nowadays, with the advent of the 5G era, mobile devices such as smart phones, Bluetooth, and tablet computers have significantly increased and are widely distributed in various regions. Because people always move with these mobile devices, these devices have random mobility [11, 12]. In this way, they can be regarded as a social node, and the community can be established through the relationship between the nodes [13, 14]. However, when data are transmitted from the source node to other nodes in the network through broadcast, in many cases, it is desirable that the broadcast delay can be as low as possible [7]. Therefore, research on a low-latency broadcast algorithm is of great theoretical significance and practical application value for the study of information transmission in opportunistic complex networks [15, 16].

There are many nodes in the opportunistic complex networks. In this way, when large-scale communication occurs in a social network, many nodes in the network will simultaneously spread information [17]. So, due to the broadcast characteristics of wireless signal transmission, the signal transmission of a node will affect the data acceptance

of other nodes within its interference range, and signal interference is an important factor that affects the node broadcast delay [18]. The above situation will result in low transmission rate and high latency for social networks. However, it has been proved [19] that, due to the influence of signal interference, the minimum delay broadcasting problem in wireless sensor networks is NP (nonpolynomial) difficult problem, so it is difficult to design a polynomial time optimization algorithm. In order to solve the problem of low-latency broadcasting in opportunistic complex networks, researchers have proposed many approximation algorithms. The broadcast delay is continuously optimized by approximating optimization to approximate the performance of the optimal algorithm [20–22]. Through these studies, the approximate ratio is continuously reduced and the performance is continuously improved. The link scheduling strategy designed under the physical interference model can effectively improve the actual performance of the broadcast algorithm [23, 24]. However, it is necessary to consider all data links transmitting simultaneously, which makes the problem more complicated and difficult, and the research challenges are greater [25].

However, the abovementioned research work has designed the broadcast algorithm through interference avoidance scheduling technology [26]. That is, when there is a signal interference between broadcast links, the transmission of information between these nodes is distributed to different time slices to avoid mutual interference. Although the interference avoidance scheduling technology can effectively reduce the interference between signals, it reduces the number of broadcast links that can be transmitted concurrently, which is not conducive to reducing the broadcast delay. In this paper, a greedy broadcast algorithm GreedyA (greedy algorithm) is first proposed. This algorithm mainly uses the breadth first search tree with the source node as the root node to stratify the nodes in the social network [27, 28]. Then GreedyA algorithm according to the rules that the node covering more number of nodes have prioritize to become the parent node to construct the broadcast tree. Finally, the broadcast link scheduling is carried out around the broadcast tree using the idea of layer-by-layer scheduling and interference avoidance scheduling. So as to better enhance the transmission performance, this paper proposes a broadcast algorithm (EDTC) based on the GreedyA algorithm, which increases the number of broadcast links that can be transmitted concurrently through successive interference cancellation techniques [29–31].

The main contributions of this research are as follows:

- (1) A greedy broadcast algorithm, GreedyA. This algorithm uses the method of layer-by-layer scheduling and interference avoidance scheduling to allocate the transmission time slices of the broadcast link, which effectively solves the problem of signal interference. But the number of broadcast links that can be transmitted concurrently is limited.
- (2) Combining with successive interference cancellation techniques, another low-latency broadcast algorithm

EDTC is proposed. This algorithm is based on the greedy broadcast algorithm GreedyA and makes full use of the benefits of successive interference cancellation to perform broadcast link scheduling. The EDTC algorithm improves the performance of information transmission in opportunistic complex networks by relaxing the link interference limit and increasing the number of broadcast links that can be transmitted concurrently.

- (3) Experiments show that using the EDTC algorithm for information transmission in opportunistic complex networks has excellent performance in reducing energy consumption and improving data transmission efficiency.

This paper is divided into five parts. The first part introduces our research. In the second part, we briefly explained the related work. The third part introduces the proposed algorithm model. In the fourth part, a simulation experiment is performed using the proposed algorithm, and the experimental results are analyzed. Section five summarizes this study.

## 2. Related Work

Over the years, research on routing algorithms has always been a hot issue in opportunistic complex networks. So far, many routing algorithms have been proposed. Among them, there are many algorithms applied to opportunistic complex networks. Several routing algorithms are described below.

Vahdat and Becker [32] proposed the epidemic routing algorithm, whose core idea is using several meeting nodes to transmit information. Lenando and Alrfaay [33] studied epidemic with social features to improve routing performance in opportunistic social networks. The core idea of this algorithm is to utilize the social activities of the nodes. Compared with the epidemic protocol, it can increase the transmission rate and reduce the transmission overhead, average delay, and average hops. Mundur et al. [34] proposed an improvement based on epidemic routing protocols, and its core idea is to use the information that the messages have been transmitted in the anti-interference list to prevent future exchange of these messages. With this technology, there will be better buffers and network utilization, which can increase the percentage of messages delivered with a lower latency. To enhance the epidemic routing in delay-tolerant networks from an energy perspective, Rango et al. [35] proposed a new strategy to dynamically adjust the  $n$ -parameter. This strategy considers the energy consumption and node degree of mobile nodes to increase or decrease the amount of data distributed in the network. With this strategy, when the remaining energy of the node is low, the scalability of the popular strategy is greatly improved, and the  $n$ -parameter is also improved. Conversely, when the mobile node has a good energy budget, more transmissions can be allowed and the  $n$ -parameter can be reduced to increase the transmission probability.

Spyropoulos et al. [36] proposed a simple solution called Spray and Wait, which managed to overcome the shortcomings of epidemic routing and other flooding-based schemes. The algorithm can avoid the performance dilemma inherent in

utility-based schemes. In order to avoid the Spray and Wait algorithm making random and blind forwarding decisions in delay tolerant networks, Xue et al. [37] proposed a Spray and Wait algorithm based on average transfer probability in delay tolerant networks. The core idea of the algorithm is using pass prediction to forward messages. Huang et al. [38] proposed a Spray and Wait routing based on location prediction in social networks. The main idea of the algorithm is that, in the waiting phase, each relay node uses polynomial interpolation to predict the future position. A copy of the message can be forwarded to another relay node closer to the target without waiting for the target node to be encountered. This solution makes full use of mobility information so that messages can be delivered to their destination faster. Jain et al. [39] proposed enhanced fuzzy logic-based Spray and Wait routing protocol for delay tolerant networks. The core idea of this algorithm is to achieve a high transfer rate by appropriately aggregating multiple message parameters. Experiments prove that compared with other Spray and Wait routing protocol variants, the proposed buffer management scheme successfully achieves the goal of increasing the delivery ratio and the overhead ratio.

For flooding routing strategies, multiple copies generated by the original nodes are used for forwarding. The network nodes have large information redundancy and high dependence on network resources. For the purpose of reducing network resources consumption to a greater extent, a routing strategy based on prediction is proposed. Dhurandher et al. [40] proposed a history-based routing prediction in opportunistic complex networks. The core idea of this algorithm is to use movement history to model the behavior of nodes. Markov predictors were used to make predictions and choose the best next node. Yu et al. [41] proposed a probabilistic routing algorithm based on contact time and message redundancy. This algorithm estimates the transit probability of a node based on the history of encounter information and contact time. And, by using a controlled replicating scheme, messages can be transmitted in parallel on multiple paths.

Based on the context-aware routing strategy, choosing the best transmission path through context-aware parameters obtained by intermediate nodes can greatly improve network performance. Wong [42] proposed the social relation opportunistic routing (SROR) algorithm. It is mainly based on social relations, social profiles, and social mobility patterns. The optimal relay node for routing data is calculated to maximize the delivery ratio. And they proved that the proposed algorithm can achieve the highest data transmission rate with the highest routing efficiency in the social environment. Xu et al. [43] proposed an intelligent distributed routing algorithm based on social similarity. This algorithm can use the social environment information in the network to predict the mobile attributes of network nodes through the BP neural network. This routing decision fully considers the time and space attributes of the mobile node. Through simulation experiments, in the comparison of other existing well-known algorithms, they find that their algorithm can improve the network's ability to adapt to topology changes.

However, in social networks, due to the broadcast characteristics of wireless signals, there will be interference

between wireless signals, which will make receiving nodes to not receive the message correctly. The related research of low-latency broadcasting algorithms is introduced below with respect to the physical interference model.

Yu et al. [24] studied the basic communication primitives in unstructured wireless networks under the physical interference model and the method for distributing broadcast messages from multiple nodes to the entire network with minimum delay. They proved that the proposed random distributed algorithm can be completed in  $O((D + nb)\log n + \log^2 n)$  time slices with a higher probability, where  $D$  is the network diameter,  $n_b$  is the number of nodes that need to send broadcast messages, and  $n$  is the network scale.

Tian et al. [44] proposed two global broadcast distributed determination algorithms based on the signal-to-interference plus noise ratio model. In these two algorithms, any node can become the source node, and the remaining nodes are divided into different layers according to their distance from the source node. Broadcast messages are transmitted layer by layer from the source node to all other nodes. For the first algorithm, by carefully selecting multiple subsets of the largest independent set of each layer, most concurrent transmissions can be allowed. Its time complexity is  $O(D \log n)$ . For the second algorithm, the running time is improved by reducing the number of repeated broadcast messages in each layer, that is, eliminating redundant broadcasts in the same layer. Theoretical analysis shows that the time complexity of the second algorithm is  $O(D\Delta \log n)$ .

However, none of the abovementioned research works have used successive interference cancellation techniques to design low-latency broadcasting algorithms. In opportunistic complex networks, different network nodes have different signal interferences. How to choose the forwarding nodes reasonably and increase the number of links that can be transmitted concurrently is very challenging. Successive interference cancellation technology can effectively decode the required signals from the disturbing signals. As a result, network performance is improved. However, as far as we know, there is no research work to apply successive interference cancellation technology to the broadcast algorithm to reduce the broadcast delay in social opportunistic complex networks, so it is necessary to conduct in-depth research.

### 3. Materials and Methods

*3.1. Network Model.* This paper considers an opportunistic complex network with  $m$  nodes. Each sensor node uses a half-duplex omnidirectional antenna for wireless communication, and the maximum transmission distance is the same. Based on the characteristics of node wireless communication, the network is modeled as a unit circle graph  $G_0 = (V_0, E_0)$ . The set  $V_0$  contains all nodes in the network, and the set  $E_0$  includes all edges in the network. There is an edge between two nodes if and only if the distance between them is less than or equal to the maximum transmission distance.

Assuming that the time between the nodes is synchronized, the scheduling time is divided into several time slices of the same length. Each node can finish sending or receiving a piece of data in a time slice. This paper regards the physical

interference model as the signal interference model. That is to say, when a node's signal to noise ratio is not lower than a certain threshold, the node can correctly decode the required signal.

**3.2. Problem Definition.** This paper researches the broadcasting problem of opportunistic complex networks, among which the source node needs to transmit its data to all sensor nodes at time slice 1. When all sensor nodes receive data from the source node, the broadcast task is completed. Broadcast scheduling is used to allocate the transmission time slice of each node. The goal of this paper is to determine how to optimize the latency that all nodes receive source node data and ensure that the scheduled data transmission signals do not interfere with each other.

**Definition 1.** BDPIM (Broadcast Delay under Physical Interference Model) question. Given the wireless sensor network  $G_0 = (V_0, E_0)$  and a source node  $j$ , under the physical interference model, a broadcast algorithm is designed so that all nodes can receive data from the source node with the lowest broadcast delay.

Existing work has proved that the BDPIM problem is an NP-difficult problem [16], so it is impossible to design an optimization algorithm for polynomial time. In order to decrease the time complexity of the algorithm and to optimize the performance of the algorithm as much as possible, a low-latency broadcast algorithm with polynomial time needs to be designed.

In opportunistic complex networks, nodes use the "storage-carry-forward" routing mode to implement inter-node communication. When we analyze the nodes in the opportunistic complex networks, we must first know its characteristics. Therefore, we point out that all social complex network nodes meet the following conditions.

We can define that, at time  $s$ , the modularity of the community can be expressed as

$$U(s) = \frac{m_e}{M} - \frac{l_t^2}{4M^2}. \quad (1)$$

Among them,  $U$  is the modularity of the community.  $M$  is the total weight of the edge.  $m_e$  represents the total weight of edges in community  $e$ .  $l_t$  expresses the total level of node  $t$  in the community.

**Condition 1.** If the node is in an opportunistic complex network, then as the weight of edges formed with other nodes in the network increases, the total edge weight  $m_e$  of the community will increase. The above situation will increase the relevance degree of the community in the opportunistic complex networks.

**Proof.** Proof At time  $s$ , the modularity in the community is  $U(s)$ , and then, the change in modularity after time  $s + 1$  can be expressed as

$$U(s + 1) = \frac{m_e + \Delta m}{M + \Delta m} - \frac{(l_t + 2\Delta m)^2}{4(M + \Delta m)^2}. \quad (2)$$

Then,

$$\begin{aligned} U(s + 1) - U(s) &= \frac{m_e + \Delta m}{M + \Delta m} - \frac{(l_t + 2\Delta m)^2}{4(M + \Delta m)^2} - \left( \frac{m_e}{M} - \frac{l_t^2}{4M^2} \right) \\ &= \frac{(4M^3 m_e + 4M^2 \Delta m m_e + 4M^3 \Delta m + 4M^2 \Delta m^2) - (M^2 l_t^2 + 4M^2 \Delta m^2 + 4M^2 l_t \Delta m)}{4M^2 (M + \Delta m)^2} \\ &\quad - \left( \frac{(4M^3 m_e + 4M \Delta m^2 m_e + 8M^2 \Delta m m_e) - (l_t^2 M^2 + l_t^2 \Delta m^2 + 2l_t M \Delta m)}{4M^2 (M + \Delta m)^2} \right) \\ &\geq \frac{4M^3 \Delta m - 4M^2 l_t \Delta m - 4M^2 m_e \Delta m + 2M^2 l_t \Delta m}{4M^2 (M + \Delta m)^2} - \frac{4M l_t \Delta m^2 - (l_t \Delta m)^2}{2M^2 (M + \Delta m)^2} \\ &\geq \frac{4M^3 \Delta m - 6M^2 l_t \Delta m + 2M^2 l_t \Delta m - 2M^2 l_t \Delta m + (l_t \Delta m)^2}{4M^2 (M + \Delta m)^2} \\ &= \Delta m \frac{4M^3 \Delta m - 6M^2 l_t + 2M^2 l_t - 2M^2 l_t \Delta m + l_t^2 \Delta m}{4M^2 (M + \Delta m)^2} \\ &= \Delta m \frac{(2M^2 - 2M l_t - l_t \Delta m) \times (2M - l_t)}{4M^2 (M + \Delta m)^2}. \end{aligned} \quad (3)$$

We can get  $\Delta m > 0$ . So we just need to prove  $(2M^2 - 2M l_t - l_t \Delta m) \times (2M - l_t) > 0$ ; then, we can get  $U(s + 1) - U(s) > 0$ .

In other words,



$$\begin{cases} 2M^2 - 2Ml_t - l_t\Delta m > 0, \\ 2M - l_t > 0, \end{cases}$$

$$\begin{cases} 2M^2 - 2Ml_t - l_t\Delta m > 0 \\ 2M - l_t > 0 \\ \Delta m > 0 \end{cases} \implies \begin{cases} 0 < \Delta m < 2M\left(\frac{M}{l_t} - 1\right) \\ l_t < 2M \end{cases}$$

$$\implies \begin{cases} 0 < \Delta m < 2M\left(\frac{M}{l_t} - 1\right) \\ 2M\left(\frac{M}{l_t} - 1\right) > 0 \\ l_t < 2M \end{cases} \implies \begin{cases} 0 < \Delta m < 2M\left(\frac{M}{l_t} - 1\right), \\ l_t < M. \end{cases} \quad (4)$$

However, we also know that  $2U$  is the sum degree of the nodes in the network, and no community in the network has a saturation greater than  $2U$ . In summary, we can get that increasing the weight can increase the relevance to the community in the social opportunity network. That is to say, we can get that if a node belongs to this opportunistic complex networks, its weight will affect the relevance of the community in the social opportunity network.  $\square$

*Condition 2.* In an opportunistic complex network, if node  $N_i$  meets the condition  $l_i l_j / 2M < m_{ij} < \Delta m + (l_i l_j + l_t \Delta m + \Delta m^2 / 2(M + \Delta m))$ , then it will be separated from the community of node  $C_j$ .

*Proof.* We first assume that community  $E$  is divided into two subcommunities  $E_i$  and  $E_j$ , where nodes  $N_i$  and  $C_j$  are in different communities. The total weight of the community decreases. Then,

$$\begin{cases} M_i + M_j < M, \\ \frac{m_i}{M} - \frac{l_i^2}{4M^2} + \frac{m_j}{M} - \frac{l_j^2}{4M^2} < \frac{l_i + l_j + m_{ij}}{M} - \frac{(l_i + l_j)^2}{4M^2}, \\ m_{ij} > \frac{l_i l_j}{2M}. \end{cases} \quad (5)$$

When the total weight has decreased, we can express the formula as

$$M_i^* + M_j^* > M^*, \quad (6)$$

$$m_{ij} < \Delta m + \frac{l_i l_j + l_t \Delta m + \Delta m^2}{2(M + \Delta m)}. \quad (7)$$

As mentioned above, we can know that if two nodes in communities  $E_i$  and  $E_j$  satisfy the condition

$l_i l_j / 2M < m_{ij} < \Delta m + (l_i l_j + l_t \Delta m + \Delta m^2 / 2(M + \Delta m))$ , then the community has been divided.  $\square$

*Condition 3.* For node  $N_i$  in the opportunistic network, if its edge is connected to node  $C_j$  and this edge is the only edge of node  $C_j$ , then when the weight between nodes  $N_i$  and  $C_j$  drops, node  $C_j$  will still not be separated from the community.

*Proof.* If community  $E$  is divided, then it must meet the following three conditions:

$$\begin{cases} M_i + M_j < M, \\ \frac{u_i}{M} - \frac{l_i^2}{4M^2} + \frac{u_j}{M} - \frac{l_j^2}{4M^2} < \frac{l_i + l_j + m_{ij}}{M} - \frac{(l_i + l_j)^2}{4M^2}, \\ m_{ij} > \frac{l_i l_j}{2M}. \end{cases} \quad (8)$$

Along with the weight change, the formula can be expressed as

$$\begin{cases} M_i^* + M_j^* > M^*, \\ m_{ij} < \Delta m + \frac{l_i l_j + l_t \Delta m + \Delta m^2}{2(M + \Delta m)}. \end{cases} \quad (9)$$

It can be understood as

$$\frac{l_i l_j}{2M} < m_{ij} < \frac{l_i(l_j + \Delta m)}{2(M + \Delta m)} = \frac{l_i l_j + l_i \Delta m}{2(M + \Delta m)}. \quad (10)$$

Because

$$\frac{l_i l_j + l_i \Delta m}{2(M + \Delta m)} - \frac{l_i l_j}{2M} = \frac{l_i \Delta m (M - l_j)}{2M(M + \Delta m)} < 0, \quad \Delta m < 0, \quad (11)$$

then

$$\frac{l_i l_j + l_i \Delta m}{2(M + \Delta m)} < \frac{l_i l_j}{2M}. \quad (12)$$

So, in the end, we can get  $l_i l_j / 2M < m_{ij} < \Delta m + (l_i l_j + l_t \Delta m + \Delta m^2 / 2(M + \Delta m))$  is false.

By the above proof, we can get the conclusion that, for the nodes in opportunistic complex networks, if its edge is connected to another node and this edge is the only edge of the node, then when the weight between two nodes drops, the node will still not be separated from the community.

In summary, if a node is in an opportunistic complex network, then it should meet the above conditions.  $\square$

### 3.3. Basic Theory of Successive Interference Cancellation.

For the interference characteristics of wireless signals, traditional algorithms usually adopt the idea of interference avoidance scheduling for broadcast link scheduling. Different from the traditional interference avoidance method, this paper considers adopting successive interference cancellation technology to increase the number of broadcast

links that can be transmitted concurrently to improve the performance of information transmission. Successive interference cancellation is a multipacket receiving technology, which can decode the required data messages from the conflicting signals and thereby effectively reduce the signal interference in wireless networks. During the iterative detection of receiving nodes with successive interference cancellation, the strongest signals are decoded, while other signals are considered interference.

The condition that a signal meets the SIC at the receiving node is that its signal internet performance noise ratio (SINR) is not lower than a specific threshold. We will try to decode the signal with the strongest signal strength when it receives the conflict signal. At this time, other transmitted signals are regarded as noise. If the decoding is successful, the receiving node removes the signal. Then, the receiving node attempts to decode the strongest of the remaining signals. This process continues until all signals are extracted or decoded fails. Through this process, all the information carried in the conflict signal can be decoded gradually, and then the required information can be obtained. This process is called SIC's sequential detection feature. Obviously, the decoding of weak signals requires the successful decoding of all stronger signals as a prerequisite. In other words, in the conflict signal, the weak signal is dependent on the strong signal.

Firstly, this article shows the association between nodes in the communication domain, as shown in Figure 1. The source nodes directly transmit information to each other, and then the source nodes transmit the information to other nodes in the communication domain through broadcasting. In a communication domain, we define 1 source node and  $m$  sensor nodes.

In this paper, the noise power is defined as  $W_{N_0}$ , the specific threshold value meeting successive interference cancellation is  $\chi_{\text{SIC}}$ , and the distance between two nodes  $A_1$  and  $B_1$  is  $d_{A_1B_1}$ . When two data links  $I_{A_1B_1}$  and  $I_{A_2B_2}$  transmit simultaneously, according to the constraints of successive interference cancellation techniques, whether the node  $R_1$  can decode the signal of  $S_1$  has the following three cases:

- (1)  $I_{A_1B_1}$  is independent of  $I_{A_2B_2}$ .  $I_{A_1B_1}$  and  $I_{A_2B_2}$  are transmitted at the same time. Node  $B_1$  can still decode the signal of  $I_{A_1B_1}$  under the interference of  $I_{A_2B_2}$ . The following conditions must be met:

$$\frac{W_{B_1}(A_1)}{W_{B_1}(A_2) + W_{N_0}} \geq \chi_{\text{SIC}}. \quad (13)$$

Among them,  $W_{B_1}(A_1)$  and  $W_{B_1}(A_2)$  indicate the signal strength the node  $B_1$  received from two sending nodes  $A_1$  and  $A_2$ . Different signal fading models can obtain different received signal strengths. For the convenience of analysis, this paper uses the same signal fading model as in [24]; that is,  $W_{B_1}(A_1)$  and  $W_{B_1}(A_2)$  can be calculated according to the following formula:

$$W_{B_1}(A_1) = \frac{W_{A_1}}{d_{A_1B_1}^n}, \quad (14)$$

$$W_{B_1}(A_2) = \frac{W_{A_2}}{d_{A_2B_1}^n}. \quad (15)$$

Among them,  $W_{A_1}$  and  $W_{A_2}$  represent the signal transmission power of nodes  $A_1$  and  $A_2$ , respectively;  $n$  represents the signal attenuation index, and the value ranges from 2 to 6. By modifying equations (14) and (15), the proposed algorithm can also be extended to other signal fading models in this paper.

As shown in Figure 2, when the distance from node  $A_2$  to node  $B_1$  is farther than the distance from node  $A_2$  to node  $A_1$ , the above conditions may be satisfied.

- (2)  $I_{A_1B_1}$  depends on  $I_{A_2B_2}$ .  $I_{A_1B_1}$  and  $I_{A_2B_2}$  are transmitted at the same time. On node  $B_1$ , the signal of node  $A_2$  is stronger and satisfies the successive interference cancellation conditions. Therefore, the signal of node  $A_1$  can be used as the interference signal. Firstly, the signal of node  $A_2$  is decoded, and then, the signal is removed, thereby decoding the signal of node  $A_1$ . The following conditions need to be met:

$$\frac{W_{B_1}(A_2)}{(W_{B_1}(A_1) + W_{N_0})} \geq \chi_{\text{SIC}}. \quad (16)$$

As shown in Figure 3, when the distance from node  $A_2$  to node  $B_1$  is closer than the distance from node  $A_2$  to node  $A_1$ , the above conditions may be satisfied.

$I_{A_2B_2}$  interferes with  $I_{A_1B_1}$ .  $I_{A_1B_1}$  and  $I_{A_2B_2}$  are transmitted at the same time. Any signal as an interference signal does not satisfy the successive interference cancellation conditions. Node  $B_1$  cannot decode the signal of node  $A_1$ ; that is, it meets the following two conditions at the same time:

$$\frac{W_{B_1}(A_1)}{(W_{B_1}(A_2) + W_{N_0})} < \chi_{\text{SIC}}, \quad (17)$$

$$\frac{W_{B_1}(A_2)}{(W_{B_1}(A_1) + W_{N_0})} < \chi_{\text{SIC}}.$$

As shown in Figure 4, when the distance from node  $A_2$  to node  $B_1$  and the distance from node  $A_2$  to node  $A_1$  is similar, the above conditions may be satisfied.

For the first two cases, when two links are transmitting data at the same time, node  $B_1$  can still decode the signal of node  $A_1$ . While in the third case, node  $B_1$  cannot decode the signal of node  $A_1$ . Based on the above characteristics of successive interference cancellation techniques, this paper will specifically design different link scheduling strategies to maximize the number of data links that can be transmitted concurrently, thereby reducing broadcast delay.

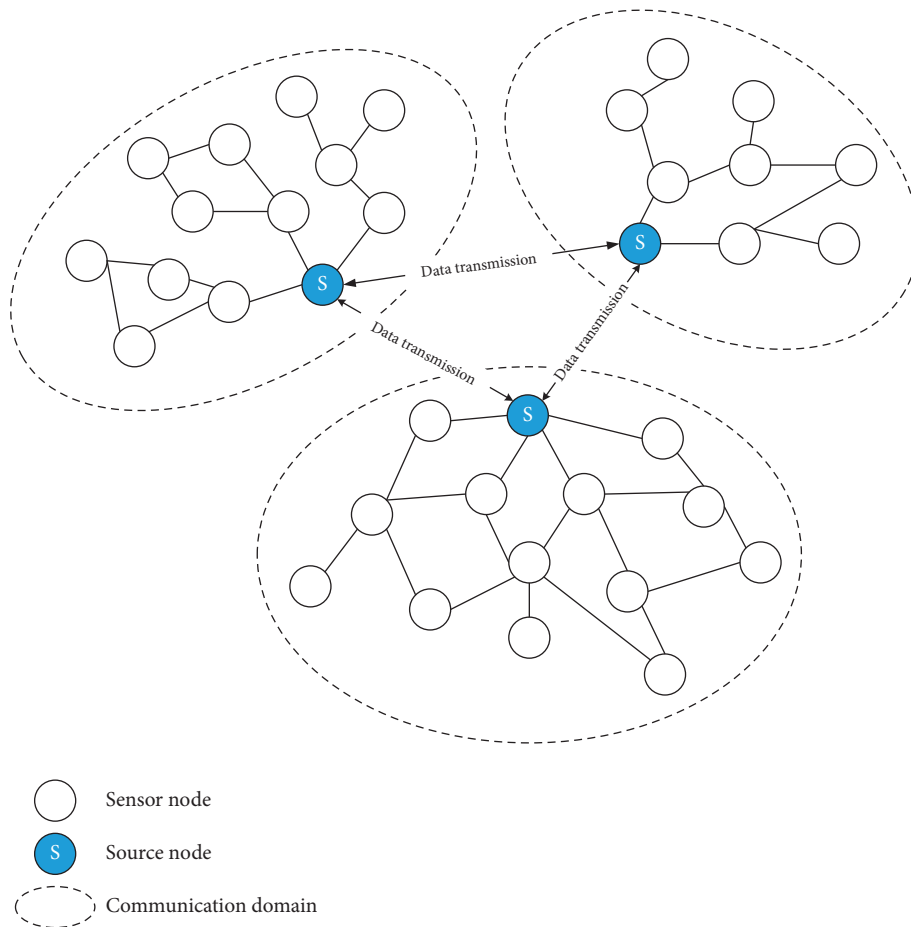


FIGURE 1: Data delivery on communication domain.

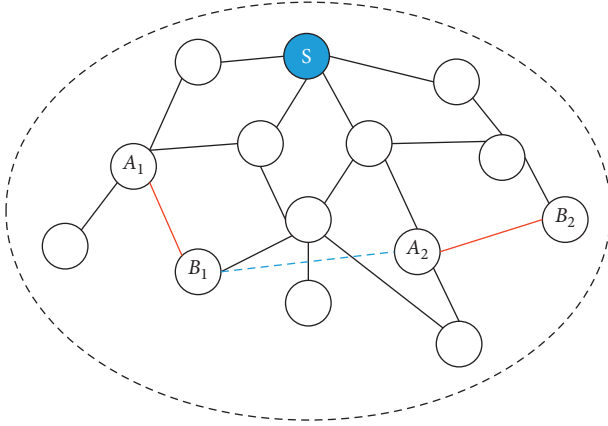
In [45], we can see a good application of successive interference cancellation technology in single- and multiple-antenna OFDM systems. Among them, the SIC-OF system has been applied to various famous network implementations, such as cellular, ad hoc, and infrastructure-based platforms. In [46], we can see the application of SIC in uplink massive MIMO systems. In the article, they research the energy efficiency when nonlinear successive-interference cancellation (SIC) receivers are employed at the BSs and provide an asymptotic analysis of the total transmit power with zero forcing SIC. As a result, as shown by the numerical results, the EE using the SIC receiver may be significantly higher than the EE using the linear receiver.

**3.4. Greedy Broadcast Algorithm.** Although the theoretical delay of the greedy broadcast algorithm is usually high in the worst case, it can often obtain a good average delay in the experiment. Therefore, this paper firstly considers designing a greedy broadcast algorithm (GreedyA) in this section. Then, on the basis of this algorithm, in the next part, we design another low-latency broadcast algorithm by combining successive interference cancellation techniques. The pseudocode of the GreedyA algorithm is shown in Algorithm 1.

The GreedyA algorithm uses a layer-by-layer scheduling method for broadcast scheduling, so the first step is to

construct a breadth first search tree with the source node as the root node and divide all nodes into different layers. With the purpose of effectively improving the performance of the broadcast algorithm and avoiding signal interference, the parent-child node relationship must be determined. Wireless signals have broadcast characteristics. In order to reduce the number of broadcast forwarding nodes, according to the rule that the node with the largest number of covered nodes takes precedence as the parent node, the algorithm constructs the broadcast tree  $T_s$  layer by layer starting from the top layer.

After the broadcast tree is constructed, the broadcast scheduling is performed layer by layer from the top layer. That is, after the broadcast data transmission of the nodes in the previous layer is completed, the next layer of nodes performs broadcast data transmission. In each layer, each node with child nodes is taken out in order, and then the transmission time slice of the node is allocated. The time slice allocation method is starting from the initial scheduling time slice  $t_j$  of this layer and analyzing whether the node has signal interference with the node that has been scheduled in the current time slice. The basis for judging the existence of signal interference is assuming that the node can perform broadcast data transmission with the node that has been scheduled in the current time slice at the same time and analyzing the receiving nodes of these nodes. If the SINR of a



— Data transmission  
 - - - Separation distance

FIGURE 2: The schematic diagram that  $I_{A_1B_1}$  is independent of  $I_{A_2B_2}$ .

receiving node is less than the specific threshold  $\chi_{\text{SIC}}$ , it is assumed that the assumption is not true; that is, there is signal interference between the node and the node that has been scheduled in the current time slice.

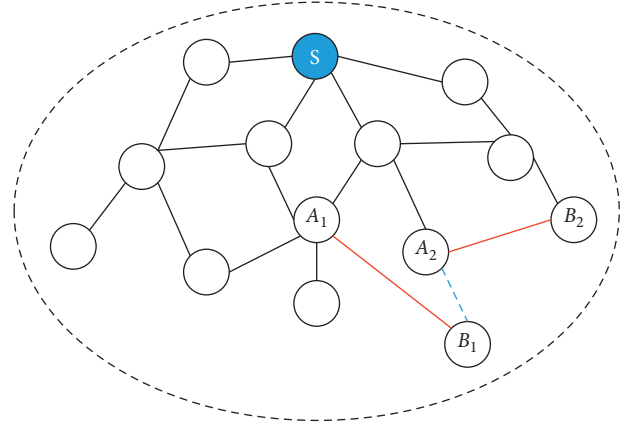
Because the GreedyA algorithm does not apply successive interference cancellation techniques, as long as the above situation occurs, the two links are considered to interfere with each other. The strategy adopted by the algorithm is performing interference avoidance scheduling, that is, dividing the interfering links into different time slices for data transmission.  $M(y)$  in Algorithm 1 represents the set of all nodes whose transmission time slices are allocated in  $y$  time slices, where  $y$  is a positive integer. However,  $|W|$  represents the number of nonempty elements  $M(y)$  contained in the set  $W$ , and it also represents the maximum transmission time slice of all scheduled nodes.

In the following, the main idea of the GreedyA algorithm is introduced through the communication domain shown in Figure 1. First, all nodes are divided into 4 layers according to the concept of breadth first search tree. As shown in Figure 5, we can get the number of layers for each node in the communication domain.

Based on the layer of each node, we can build a breadth first search tree, as shown in Figure 6.

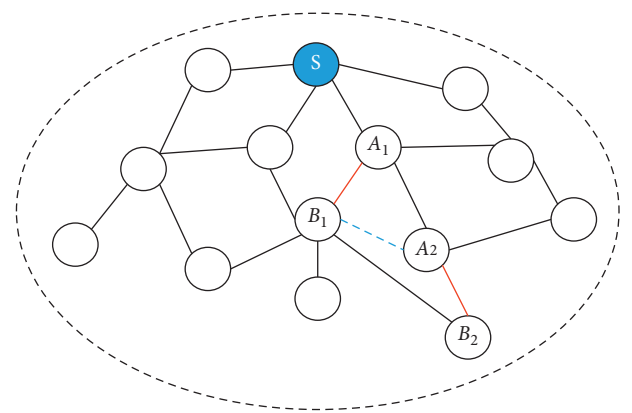
According to the rule that the node with the largest number of covered nodes takes precedence as the parent node, the algorithm constructs the broadcast tree  $T_s$  layer by layer starting from the top layer  $D_0$ . For example, the number of nodes covered by node  $n_3$  in layer  $D_1$  is the largest, so the node  $n_3$  is selected as the parent node of  $n_6, n_7$ , and  $n_8$  in layer  $D_2$ . As shown in Figure 7, we can get the final broadcast tree. Through the idea of the GreedyA algorithm mentioned above, we list three situations that may encounter signal interference during transmission in community 1, community 2, and community 3.

The next step of the GreedyA algorithm is to perform the broadcast chain scheduling. The method adopted is interference avoidance scheduling. Because there is only one



— Data transmission  
 - - - Separation distance

FIGURE 3: The schematic diagram that  $I_{A_1B_1}$  is dependent of  $I_{A_2B_2}$ .



— Data transmission  
 - - - Separation distance

FIGURE 4: The schematic diagram that  $I_{A_1B_1}$  interferes with  $I_{A_2B_2}$ .

sending node, scheduling in  $D_0$  and  $D_1$  layers is simple. In the  $D_1$  layer, after node  $n_3$  dispatches the transmission time slice, it needs to allocate the transmission time slice of node  $n_2$ . As shown in Figure 8, whether the signals of nodes  $n_3$  and  $n_2$  will affect their respective receiving nodes is analyzed; that is, whether the SINR of their receiving nodes will be below a specific threshold. If not affected, nodes  $n_3$  and  $n_2$  can be arranged for transmission at the same time slice; otherwise, node  $n_2$  will be arranged for transmission at the next time slice to avoid signal interference.

In the  $D_2$  layer, after node  $n_6$  dispatches the transmission time slice, it needs to allocate the transmission time slice of nodes  $n_5$  and  $n_7$ . As shown in Figure 9, first of all, it is analyzed whether the signals of nodes  $n_6$  and  $n_5$  will affect their respective receiving nodes, that is, whether the SINR of their receiving nodes will be below a specific threshold. If not affected, nodes  $n_6$  and  $n_5$  can be arranged for transmission at the same time slice; otherwise, node  $n_5$  will be arranged for transmission at the next time slice to avoid signal interference. It is analyzed whether the signals of nodes  $n_6$  and  $n_7$

**Input:** network size, source node, node position, maximum transmission distance, signal transmission power, SIC-specific threshold, noise power, signal attenuation index;  
**Output:** transmission time slice and broadcast delay of all nodes;  
 Begin  
**Let** breadth first search tree  $T$ ; //creating a breadth first search tree with the source node as the root node  $T$   
 $D_0, D_1, D_2, \dots, D_Y$ ; //dividing all nodes into different layers, including  $D_0, D_1, D_2, \dots, D_Y$ ;  
 $Y$ ; //represents the height of the breadth first search tree;  
 Broadcast tree  $T_S$ ; //according to the rule that the node with the largest number of covered nodes takes precedence as the parent node, construct the broadcast tree  $T_S$  layer by layer starting from the top layer  $D_0$ ;  
 $t_j \leftarrow 1$ ; //scheduling time slice  
 $M \leftarrow \emptyset$ ; //the set of nodes that have scheduled transmission time slices  
**FOR**  $m \leftarrow 0$  to  $Y - 1$  **DO**  
   **FOR**  $n \leftarrow 1$  to length ( $D_m$ ) **DO**  
     Take out the  $n$ -th node  $D_m(n)$  in  $D_m$ ;  
     **IF** the set of child node of node  $D_m(n)$  is not an empty set **THEN**  
        $y \leftarrow t_j$ ;  
       **WHILE** exist signal interference between nodes  $D_m(n)$  and  $M(y)$  **DO**  
          $y \leftarrow y + 1$ ;  
       **END**  
       The transmission time slice of the scheduling node  $D_m(n)$  is  $y$ ;  
        $M(y) \leftarrow M(y) \cup \{D_m(n)\}$   
     **END**  
**END**  
 $t_j \leftarrow |W| + 1$   
**END**  
 Broadcast delay  $\leftarrow t_j$ ;  
**RETURN** transmission time slice and broadcast delay of all nodes

ALGORITHM 1: GreedyA.

will affect their respective receiving nodes. If not affected, nodes  $n_6$  and  $n_7$  can be arranged for transmission at the same time slice; otherwise, node  $n_7$  will be arranged for transmission at the next time slice to avoid signal interference. The transit time relationship between nodes  $n_5$  and  $n_7$  is also analyzed as above.

**3.5. Broadcast Algorithm Based on Successive Interference Cancellation.** It is worth noting that the method adopted by the GreedyA algorithm in broadcasting chain scheduling is the idea of interference avoiding scheduling. However, combined with successive interference cancellation technology can increase the number of broadcast links that can be transmitted concurrently in a certain extent. In this section, designing another low-latency broadcast algorithm EDTC based on successive interference cancellation technology and GreedyA algorithm is to further improve transmission performance. The steps of the EDTC algorithm and the GreedyA algorithm are basically the same, but the main difference is that the basis for judging whether the broadcast link interferes; that is, the 10th line in the pseudocode is different. The two broadcast links,  $I_{A_1B_1}$  and  $I_{A_2B_2}$ , do not interfere with each other, and the judgment basis of the GreedyA algorithm is as follows:

$$\text{result} == (I_{A_1B_1} \text{ independent of } I_{A_2B_2}) \cap (I_{A_2B_2} \text{ independent of } I_{A_1B_1}). \quad (18)$$

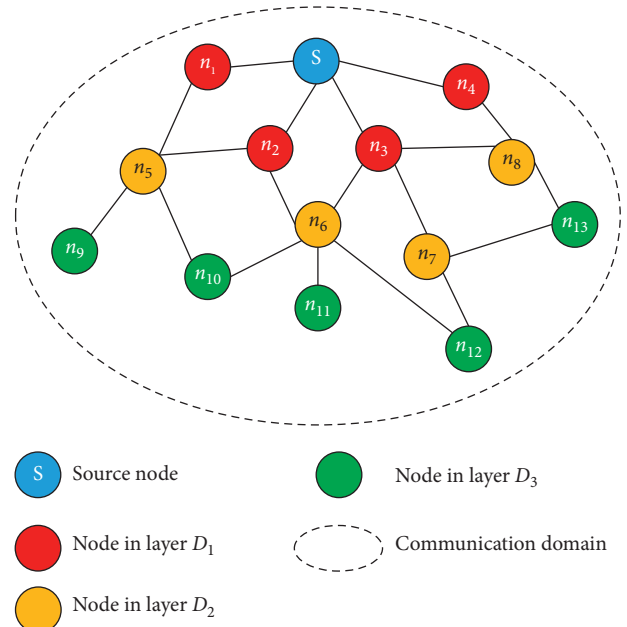


FIGURE 5: Node hierarchy graph.

Formula (7) requires that the SINR of the receiving nodes of the two links must be higher than the specific threshold. However, the judgment of the EDTC algorithm is

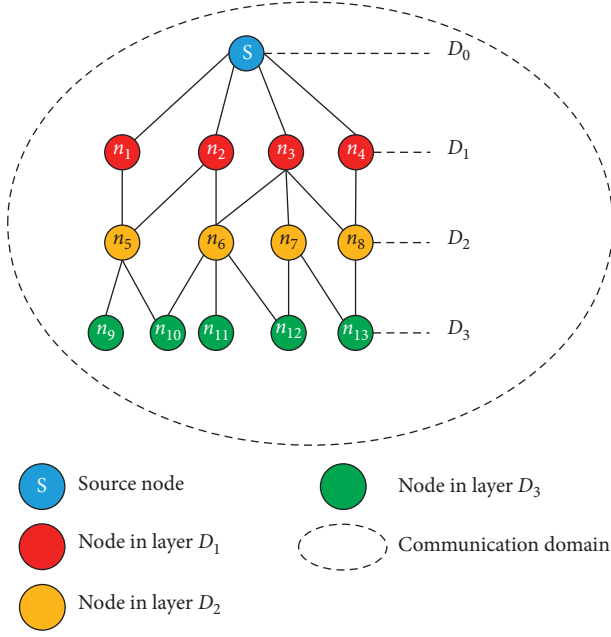


FIGURE 6: The layering of network node.

based on the interference between the two links under the SIC condition, which satisfies

$$\text{result} = (I_{A_1B_1} \text{ interfere } I_{A_2B_2}) \cap (I_{A_2B_2} \text{ interfere } I_{A_1B_1}). \quad (19)$$

Among them, result is the judgment consequent. It is different from GreedyA algorithm, when  $I_{A_1B_1}$  is independent of  $I_{A_2B_2}$  or  $I_{A_1B_1}$  depends on  $I_{A_2B_2}$ , EDTC algorithm considers that  $I_{A_1B_1}$  does not interfere with  $I_{A_2B_2}$ . Therefore, the interference restriction is relaxed, and the number of broadcast links that can be transmitted concurrently is increased, which is beneficial to improving the performance of information transmission.

In the example of the broadcast link scheduling shown in Figure 9, if the situation shown in Figure 10 exists, node  $n_6$  is closer to node  $n_{10}$ ; therefore, the GreedyA algorithm analyzes whether node  $n_6$  will interfere with node  $n_{10}$  to receive the signal from node  $n_5$ . As a result, two sending nodes cannot perform data transmission at the same time and need to perform interference avoidance scheduling. However, after combining with the successive interference cancellation technology, the EDTC algorithm analyzes the receiving node  $n_{10}$ , which can firstly decode the signal of the  $n_6$  node and then remove the signal. Thus, the signal of node  $n_5$  is obtained, so the simultaneous data transmission by the two sending nodes will not affect the normal reception of data by the receiving node.

Because in the practical application of the EDTC algorithm, there may be multiple broadcast links. Therefore, it is necessary to analyze the cumulative interference effects of multiple broadcast links; that is, the influence of multiple interference signals needs to be considered in the denominators of formulas (13) and (16).

### 3.6. Theoretical Analysis of Algorithm Performance

#### 3.6.1. Correctness Analysis

**Theorem 1.** *The GreedyA algorithm provides a correct broadcast scheduling scheme.*

*Certification.* Consider the entire opportunistic complex networks as a connected network. The GreedyA algorithm first stratifies all nodes, then constructs broadcast tree, and finally uses the method of layer-by-layer scheduling for information transmission. Interference avoidance scheduling allocates the transmission time slice of each sending node. Because the broadcast tree covers all nodes in the entire network, each node has the parent node. That is, broadcast data can be transmitted from the source node to all nodes in the network. Next, it is analyzed whether the forwarding node in the broadcast tree has scheduled its parent node to broadcast data to the node before sending the data. In addition, consider whether the scheduled broadcast links will affect the correct reception of data due to signal interference.

According to the construction rules of the broadcast tree, the parent node of a forwarding node is one level above the node. Because after the ending of scheduling at each layer, the scheduling time slice will be set as the maximum scheduled transmission time slice plus 1. Therefore, the transmission time slice of the forwarding node must be smaller than the transmission time slice of any node in the previous layer.

The idea of interference avoidance scheduling adopted by the GreedyA algorithm allocates the transmission time slice of the broadcast link. That is, when two broadcast links interfere with each other, the two broadcast links are allocated to different time slices for transmission. We can understand that when two broadcast links interfere with each other, the two broadcast links are allocated to different time slices for transmission. Therefore, the scheduled broadcast links will not affect the correct reception of data due to signal interference.

**Theorem 2.** *The EDTC algorithm provides a correct broadcast scheduling scheme.*

*Certification.* The steps of the EDTC algorithm are similar to the steps of the GreedyA algorithm, and the only difference is the basis for judging whether the link is interfered. According to the basic idea of successive interference cancellation, when the two links have a dependency relationship, the receiving node can correctly decode the required data signal. Therefore, the broadcasting chain scheduled by the EDTC algorithm can also realize the correct reception of data.

#### 3.6.2. The Analysis of Time Complexity

**Theorem 3.** *The time complexity of the GreedyA algorithm is  $O(Yn^3)$ .*

*Certification.* The first step of the EDTC algorithm is to construct a breadth first search tree to stratifies all nodes, and

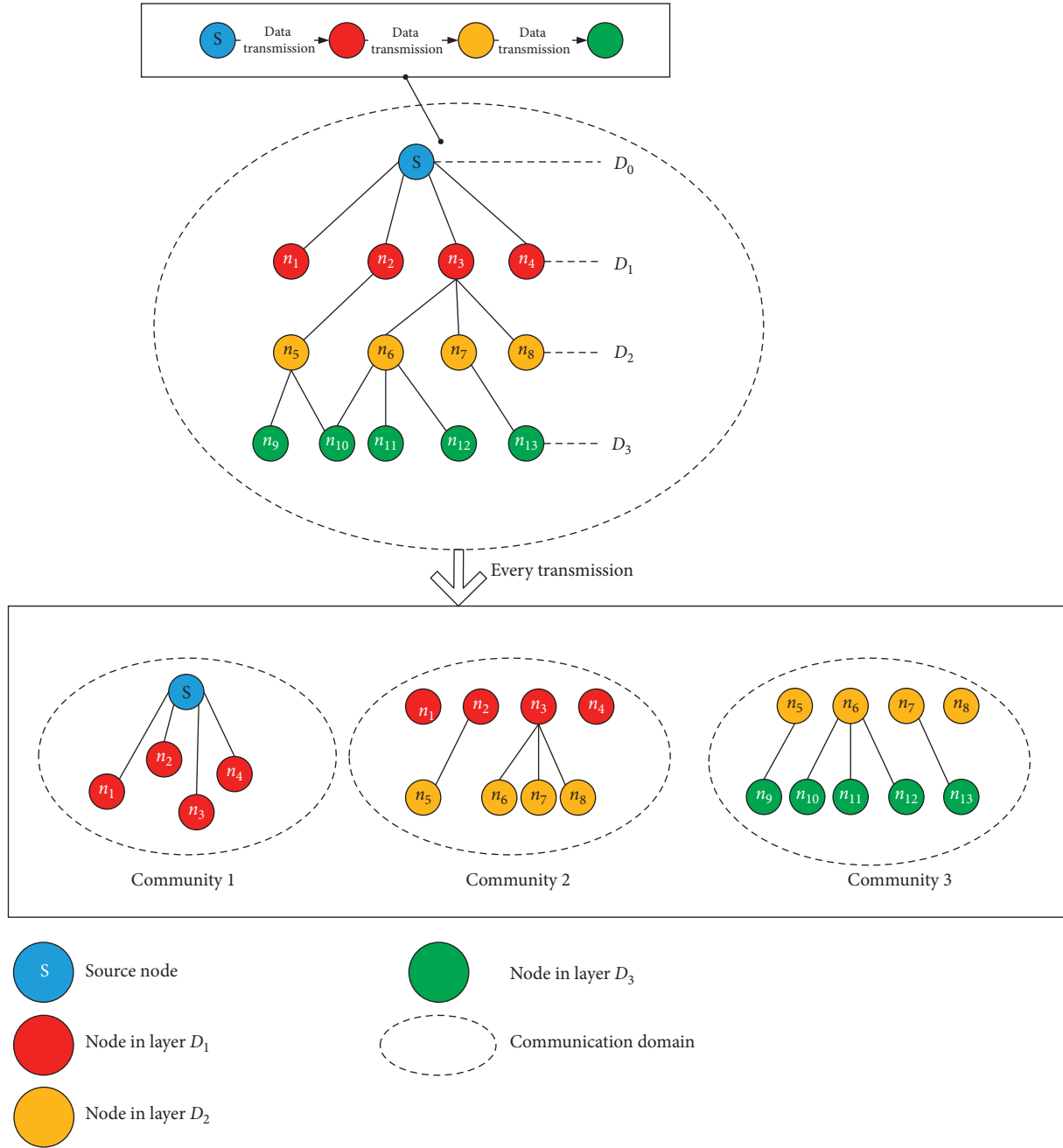


FIGURE 7: Broadcast tree and three communities.

it needs to traverse each node. So, the time overhead is  $O(n)$ . The second step is to construct a broadcast tree, which determines the parent-child relationship between nodes. The rule with the largest number of covered nodes needs to analyze the neighbors of each node, so the time overhead is  $O(n^2)$ ; The final step is to broadcast scheduling level by level. When performing broadcast scheduling at each layer, it is necessary to consider whether there is an interference relationship between different broadcast links; therefore, the time overhead of broadcast scheduling for each layer is  $O(n^3)$ . Because the total number of layers is  $Y$ , the last step

requires  $O(Yn^3)$  time. Therefore, the time overhead of all steps can be summed up to get this theorem.

**Theorem 4.** *The time complexity of the EDTC algorithm is  $O(Yn^3)$ .*

*Certification.* The steps of EDTC algorithm and GreedyA algorithm are basically the same. The main difference is the basis for judging interference. Therefore, by proof similar to

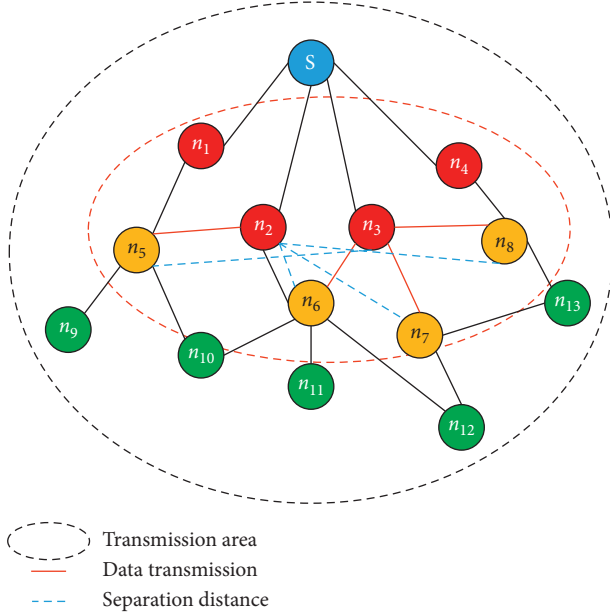


FIGURE 8: Broadcast link scheduling in community 2.

Theorem 4, it can be obtained that the time complexity of the EDTC algorithm is also  $O(Yn^3)$ .

### 3.6.3. The Analysis of Spatial Complexity

**Theorem 5.** *The spatial complexity of the GreedyA algorithm is  $O(n)$ .*

*Certification.* Analyze the amount of temporary storage space at each step of the GreedyA algorithm during the working process. The first step is to construct a breadth first search tree and stratify the nodes. The information of the neighbor node needs to be stored during the working process. Because the node has at most  $\Phi$  neighbor nodes, the space complexity of this step is  $O(\Phi)$ . The second step is to construct a broadcast tree. The information of the neighbor nodes is also required to be stored during the working process, so the spatial complexity is  $O(\Phi)$ . The last step is to perform broadcast scheduling according to the broadcast tree. During the operation, the scheduled broadcast link information needs to be stored. The spatial complexity is  $O(n)$ . After synthesizing the spatial complexity of all steps, it can be obtained that the spatial complexity of the GreedyA algorithm is  $O(n)$ .

**Theorem 6.** *The space complexity of the EDTC algorithm is  $O(n)$ .*

*Certification.* The steps of EDTC algorithm and GreedyA algorithm are basically the same. The main difference is the basis for judging interference. No extra storage space is required in the judgment process. Therefore, this theorem can be established through a certificate similar to Theorem 5.

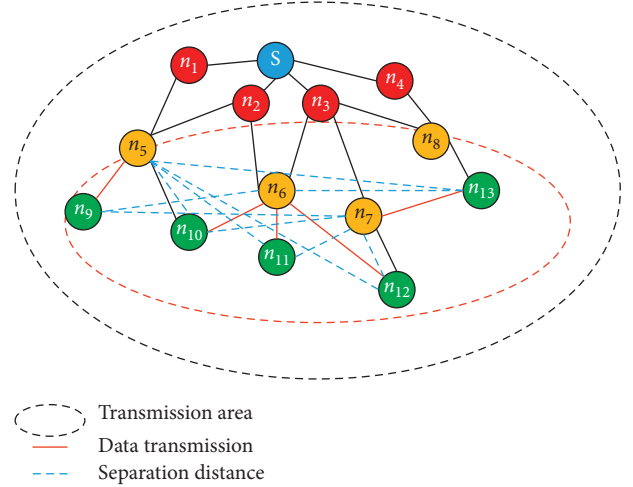


FIGURE 9: Broadcast link scheduling in community 3.

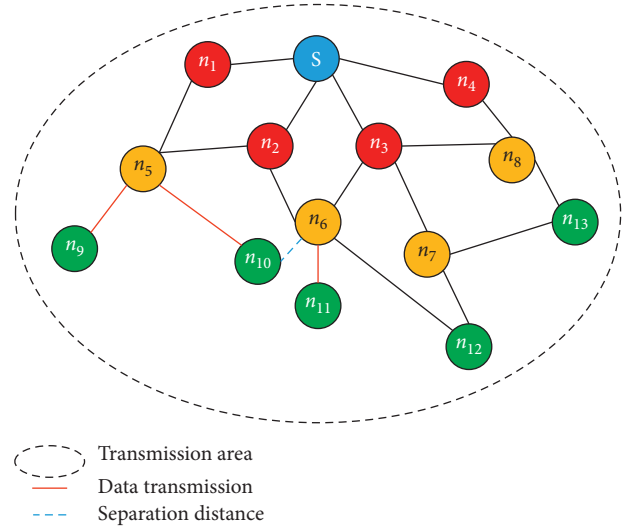


FIGURE 10: Broadcast link scheduling based on EDTC.

## 4. Results and Discussion

For the evaluation of the experimental performance of the EDTC algorithm, we test it with the opportunistic complex networks environment (ONE). In addition, in order to better judge its performance, EDTC will be compared with four other algorithms: ICMT (information cache management and data transmission algorithm) [1], SECM (status estimation and cache management algorithm) [47], Spray and Wait routing algorithm [36], and GreedyA algorithm. The following is an introduction to the principles of these algorithms:

- (1) ICMT: this algorithm is an information cache management and transmission algorithm based on node data information cache. For achieving the purpose of adjusting the cache, the algorithm evaluates the transmission probability between nodes by identifying the neighbor nodes and then adjusts the cache data



distribution to ensure that nodes with a higher transmission probability have priority to access information. Simultaneously, neighbor nodes share the cache tasks of the nodes and effectively distribute data [1].

- (2) SECM: to evaluate the probability of nodes in the project by establishing a method to identify surrounding neighbors, ensuring that nodes with high project probability obtain information first and achieve the purpose of adjusting the cache [47].
- (3) Spray and Wait: this algorithm successfully overcomes the shortcomings of epidemic routing and other flood-based schemes, avoiding the performance dilemma inherent in utility-based schemes [36].
- (4) In addition, we compare the proposed EDTC algorithm with the GreedyA algorithm. The algorithm we proposed is based on the GreedyA algorithm of the greedy broadcast algorithm and fully utilizes the benefits of successive interference cancellation to schedule the broadcast link in the social network, relaxing the limit of link interference and increasing the number of broadcast links that can be transmitted concurrently. By comparing the simulation experimental data between the two algorithms, we can see more clearly the impact of successive interference cancellation technology on the information transmission in the social network.

In simulation experiments, we mainly evaluate the performance of the algorithm based on the following parameters:

- (1) Delivery ratio: this parameter represents the possibility of selecting a relay node during transmission. The delivery ratio is an important indicator in the performance evaluation of opportunistic social networks, which directly reflects the performance of the data distribution mechanism.
- (2) Overhead on average: the network overhead of successfully transmitting information between a pair of nodes can be expressed as this parameter, referring to the ratio of the difference between the total number of forwarded message copies in the network and the total number of messages successfully delivered to the destination node.
- (3) Energy consumption: this parameter represents the energy consumption during transmission.
- (4) End-to-end delay on average: this parameter represents the average delay in transmitting information between two nodes. End-to-end network delay is the time from when a message is generated to when it is successfully delivered to the destination node.

During the simulation experiment, we set its parameters as follows: the maximum transmission distance is 35 m, the signal transmission power is 1300 W, the noise power is 1.5 W, and the signal attenuation index is 2. Other parameters for environment configuration are shown in Table 1.

In the EDTC algorithm, the size of the specific threshold will directly affect whether a receiving node can decode the

TABLE 1: Experimental parameter settings.

Parameter	Value
Communication area	3800 m × 3000 m
Total number of nodes	400
Node movement speed	0.5 ~ 1.5 m/s
Node cache storage	3 MB
Simulation time	5 hours
Movement mode	Random site moving model
Maximum transmission domain	15 m <sup>2</sup>
The interval between each two times	20–25 s
Initial energy	100 Joules
Each energy depletion	0.5 Joules

obtained signal and remove it from the mixed signal, thereby receiving the required signal. Therefore, different specific thresholds will cause the EDTC algorithm to exhibit different performances during the transmission of information. In order to better explain the relationship between the specific threshold value and transmission performance, this paper uses subjective judgment to set the specific threshold value to 1.5, 2.5, 3.5, 4.5, and 5.5 for experiments.

First, through Figures 11(a)–11(d), we can find that no matter what the value  $\chi_{SIC}$  is, their delivery ratio, overhead on average, and end-to-end delay on average all tend to stabilize over time. For energy consumption, it gradually increases with time. By comparing the experimental data curves obtained with different thresholds, it can be found that when the value of threshold is too high or too low, the performance of information transmission in social networks is not very ideal. For example, when  $\chi_{SIC} = 1.5$ , the message delivery ratio is only maintained at about 0.56, and when  $\chi_{SIC} = 5.5$ , the energy consumption is the most. This is because the standard of whether the receiving node can decode the signal is determined by the specific threshold value. When the specific threshold is continuously increased, the conditions for decoding interference signals become more and more severe so that the propagation performance of the EDTC algorithm shows a certain upward trend. However, as the specific threshold continues to increase, the interference signal cannot be effectively decoded, thereby reducing the performance of information transmission. But when the specific threshold value is in the interval [2.5, 3.5], a better transmission result can be obtained. In order to better verify the performance of the EDTC algorithm in subsequent experiments, we set the specific threshold to 3.5.

Through the simulation experiments, we found that the relationship between the four parameters to be evaluated and time is shown in Figures 12–15.

Firstly, in Figure 12, we show the relationship between each algorithm's delivery ratio in simulation experiments and simulation experiment time. We can clearly see from the figure that the lowest delivery ratio are Spray and Wait routing algorithms (copy = 30) and SECM. Their values are 0.31–0.36 and 0.35–0.39 separately. We can know that when the Spray and Wait routing algorithm (copy = 30) and SECM algorithm are used for information transmission, because they use the flooding method to transmit information to the nodes in the community, a large amount of information is

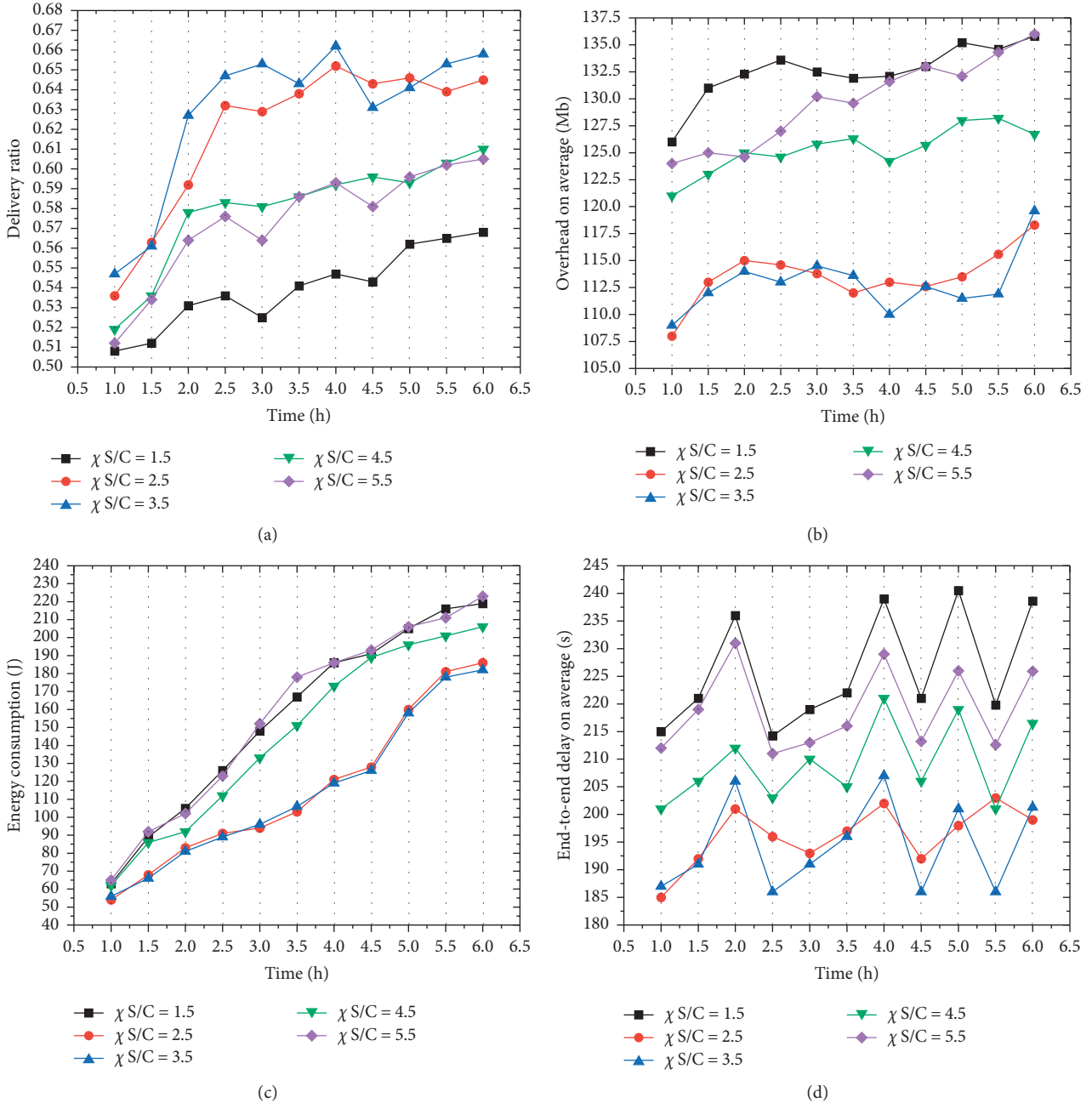


FIGURE 11: The relationship between each parameter value and time under different specific thresholds  $\chi_{SIC}$ .

lost in the process. Especially for the Spray and Wait routing algorithms, we can see that when the number of copies is 30, its delivery ratio is significantly lower than when the number of copies is 15. The delivery ratio of the Spray and Wait routing algorithms (copy = 15) is 0.41–0.47. Therefore, the excessive copying of data is an important reason for the reduction of the delivery ratio of the Spray and Wait routing algorithm. However, in the ICMT algorithm, the transmission of all its packets depends on the cooperation of the cache, which effectively uses the cache space. This achieves the purpose of increasing the delivery ratio. It can be found from Figure 12 that the delivery ratio of the experimental data using the ICMT algorithm is 0.53–0.59, which is 147%

more than SECM. As for the EDTC algorithm proposed in this paper, because when information is transmitted, the nodes in the same community are used to construct a broadcast tree, and the appropriate next-hop node is selected for hierarchical propagation through the coverage. In this way, the reliability and value of the nodes are comprehensively considered, and the node with the highest comprehensive utility value is accurately found, which avoids signal interference when data are propagated in parallel and also guarantee the probability of the message reaching the destination node. Because of this, it has greatly improved the delivery ratio in social networks. It can be seen that its delivery ratio reaches 0.66, which is the highest among these

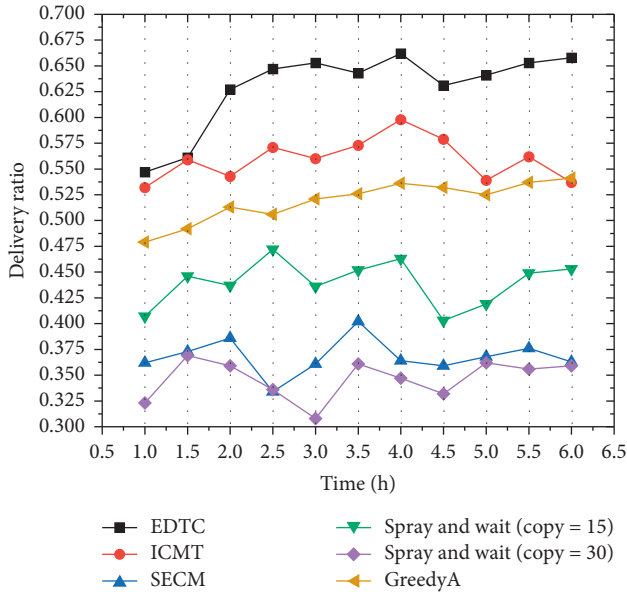


FIGURE 12: The relationship between delivery ratio and time.

algorithms. For the GreedyA algorithm, because it is limited by the number of broadcast links that can be transmitted concurrently, the delivery ratio is lower than the EDTC algorithm. However, because it uses the layer-by-layer scheduling and interference avoidance scheduling methods, it allocates the transmission time slices of the broadcast link, which effectively solves the signal interference problem, so its delivery ratio is stronger than the Spray and Wait and SECM algorithms, which is in a relatively high level.

From Figure 13, we can get the relationship between routing overhead and time. In the transmission algorithms of social networks, if there is a lack of a cache management method to maintain information transmission when a node meets its neighbors, a large amount of redundant data will be received in the cache, thus increasing the overhead. Therefore, in the simulation experiments, the overhead on average obtained by applying the SECM algorithm and the Spray and Wait algorithms is high. For the ICMT algorithm, the overall mean overhead is low and stable, but it increased in the early stage, reaching its peak in 2 hours, and then began to decline and stabilize. Because when using the EDTC algorithm to transmit information, through the constructed broadcast tree, messages can always be delivered to the correct node, thereby avoiding forwarding to unnecessary nodes and reducing the number of copies of forwarded messages in the network. Those improve network congestion and effectively reduce network overhead. Like the EDTC algorithm, the overhead on average of the GreedyA algorithm is also maintained at an ideal level.

Figure 14 shows the connection between energy consumption and stimulation time. In simulation experiments, all energy consumption increases with time. Among them, because the Spray and Wait routing algorithms need to transmit information through spray, its energy consumption is the largest. It can be seen from Figure 14 that, at the 6th hour, its energy consumption reached 540. For the ICMT

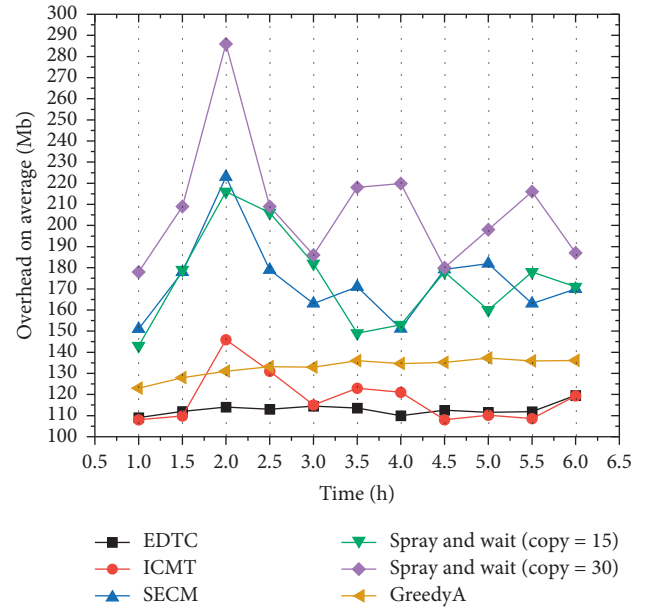


FIGURE 13: Relationship between routing overhead and time.

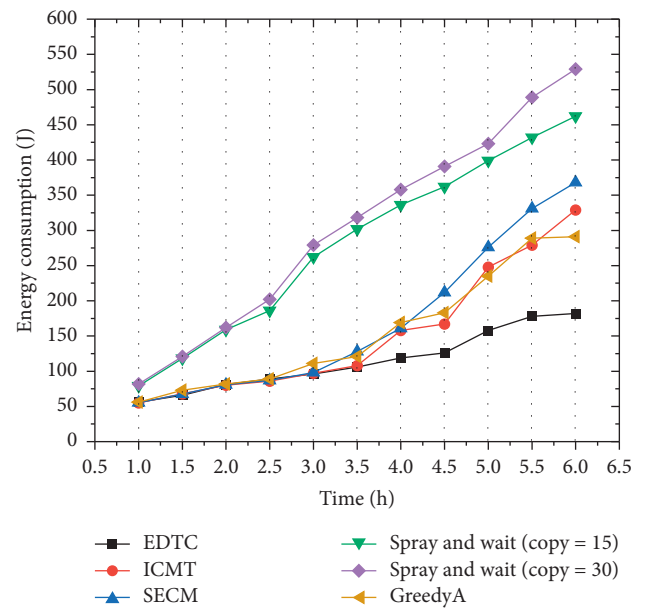


FIGURE 14: Relationship between energy consumption and time.

algorithm, because it can exchange valid data through cooperative nodes, it retains more energy to continue transmission, so its energy consumption data can maintain a low level. The energy consumption data obtained by applying the SECM algorithm is similar to the ICMT algorithm. For the EDTC algorithm, it is different from flooding to achieve the highest delivery rate without any cost for data transmission, resulting in a large amount of energy consumption during data transmission. The EDTC and GreedyA algorithm spreads information transmission by selecting the most appropriate nodes in a hierarchical manner, thereby ensuring the delivery rate and keeping consumption low.

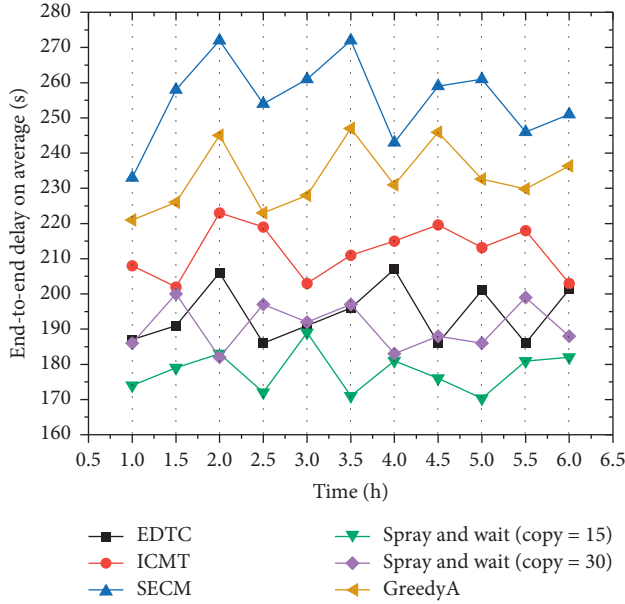


FIGURE 15: Relationship between end-to-end delay on average and time.

In Figure 15, we show the relationship between mean transmission delay and time. It can be seen from Figure 15 that the transmission delay in the social network using the SECM algorithm is very high mainly because the spray step is carried by nodes. The ICMT algorithm uses a cooperative mechanism to achieve reasonable utilization of node cache space. This reduces the propagation delay, but the effect is not very obvious. It is not difficult to find that the best algorithm to reduce the transmission delay among these algorithms is Spray and Wait routing algorithms, which is mainly because neighbors and cooperative nodes are used when data are transferred between nodes, and a large number of shared caches can be used in transmission. According to the broadcast characteristics of wireless signals, nodes only need to broadcast once to transmit data to nodes within coverage. For our proposed EDTC algorithm, it is a broadcast tree constructed according to the rule that the number of covering nodes has the highest priority as the parent node. In addition, the signal interference is avoided through successive interference cancellation technology so as to increase the number of simultaneous transmissions. As shown in Figure 15, its mean transmission delay is maintained at a very low value. Compared with the EDTC algorithm, the GreedyA algorithm has a lower number of broadcast links that can be transmitted concurrently, so its broadcast delay is higher than the EDTC algorithm.

In summary, in the four aspects of delivery ratio, overhead on average, energy consumption, and end-to-end delay on average, we can conclude that the performance of EDTC is better than other four algorithms. However, for the end-to-end delay on average, the performance of EDTC is worse than Spray and Wait routing algorithms.

In social networks, node cache has a great impact on the transmission efficiency of the algorithm. Therefore, we continue to do simulation experiments to test the influence

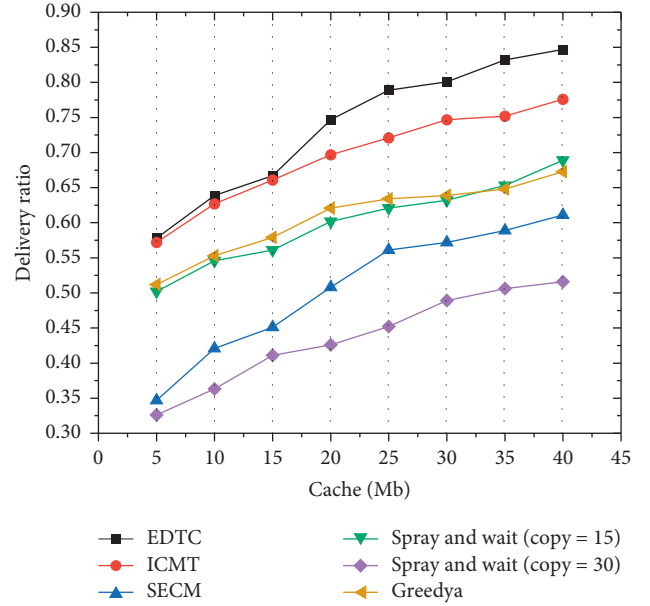


FIGURE 16: Relationship between delivery ratio and cache.

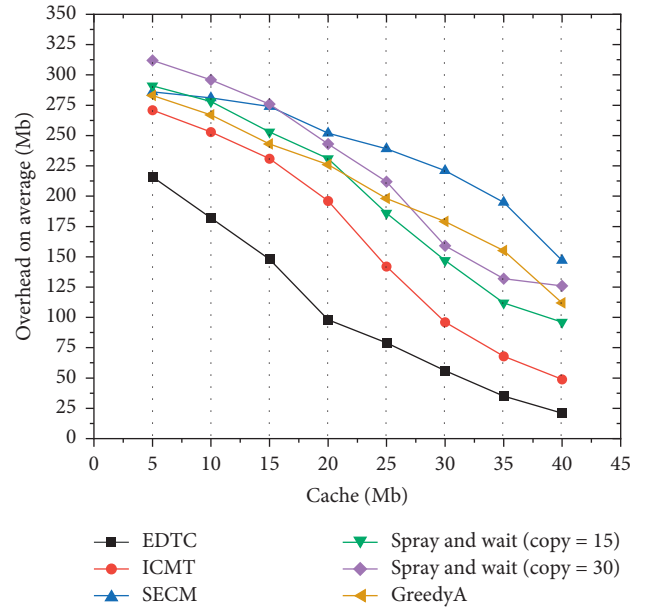


FIGURE 17: Relationship between overhead on average and cache.

of cache on these four parameters. The experimental data that we obtained are shown in Figures 16–19.

In Figure 16, we show the relationship between delivery ratio and cache. When the buffer is small, the network cannot meet the message cache requirements due to the large number of message copies, and the old messages will be quickly squeezed out by the new ones, causing a large number of packets to be dropped. Therefore, the delivery ratio of the five algorithms is not high when the cache is low. But with the increase in the buffer capacity, the transmission rate has increased to varying degrees. Among them, because the Spray and Wait routing algorithms (copy = 30) use the method of

flooding to spread information and require high cache size, its delivery ratio is the lowest. However, for the EDTC and GreedyA algorithm, nodes only need to broadcast once to transmit data to nodes within coverage. Because the GreedyA algorithm is limited by the number of concurrent broadcast links, the delivery ratio is slightly lower. It can be seen from the experimental results that EDTC has the highest delivery ratio. As for the application of ICMT and Spray and Wait routing algorithms (copy = 15) in the process of transmitting information on opportunistic complex networks, with the increase in the node cache, the transportation conditions improved, and the delivery ratio increased by more than 50%.

The association between routing overhead and cache is shown in Figure 17. In general, as the cache increases, the packet loss of nodes in the network becomes smaller, more messages can be successfully transmitted, and the overhead is getting lower and lower. As shown in Figure 17, the overhead of SECM algorithm is the largest because many redundant data are injected by nodes. As the node cache increases, the overhead on average applying the Spray and Wait routing algorithms (copy = 30) is reduced from 310 to 136. Similarly, the overhead of Spray and Wait routing algorithms (copy = 15) is reduced from 281 to 98. But their percentage of decline is much lower than the EDTC and ICMT algorithms. It can be found that, by increasing the node cache, the purpose of decreasing the routing overhead of the nodes in the community can be achieved. As the cache increases, the overhead on average of GreedyA algorithm also decreases from 280 to 117.

The relationship between energy consumption and cache is shown in Figure 18. The experimental results show that, with the increase in node cache, the energy consumption of EDTC algorithm in opportunistic complex networks can be maintained at 48. GreedyA also has a good performance, and energy consumption is stable at 76. The energy consumption of the other three algorithms has increased significantly. Because the Spray and Wait routing algorithms use the ‘‘Spray’’ method and all neighbors receive the data packets, its energy consumption is the largest. In two experiments using the Spray and Wait routing algorithms, the Spray and Wait routing algorithms (copy = 30) with a larger copy data consumes more energy. For the ICMT algorithm, the effect buffer management method can cut down energy consumption. So its energy consumption is less than the Spray and Wait routing algorithms.

As shown in Figure 19, the relationship between the average delivery delay and the cache can be obtained. Figure 19 shows that the mean delay decreases as the node cache increases. For the SECM algorithm, the average delay is generally high, which is mainly because many probability calculation tasks are carried by nodes. For the ICMT algorithm, as the node cache increases, the average delay decreases from 178 to 56. In the EDTC algorithm, the mean delay is almost stable at 41. This shows that the size of the cache has a small impact on the average delay of the EDTC algorithm. In addition, the delay generated by the greedy algorithm is higher, which shows that multiple concurrent links generated by successive interference cancellation techniques can reduce the delay of information transmission in social networks. For the Spray and Wait routing algorithms, although the average delay decreases with the

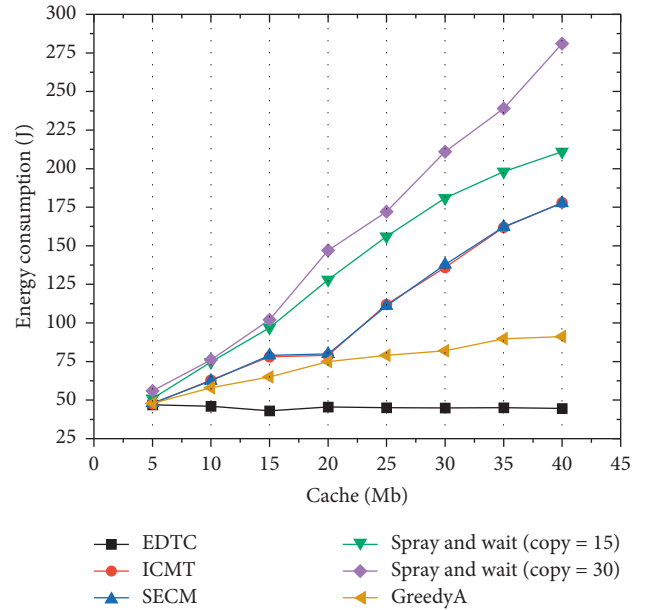


FIGURE 18: Relationship between energy consumption and cache.

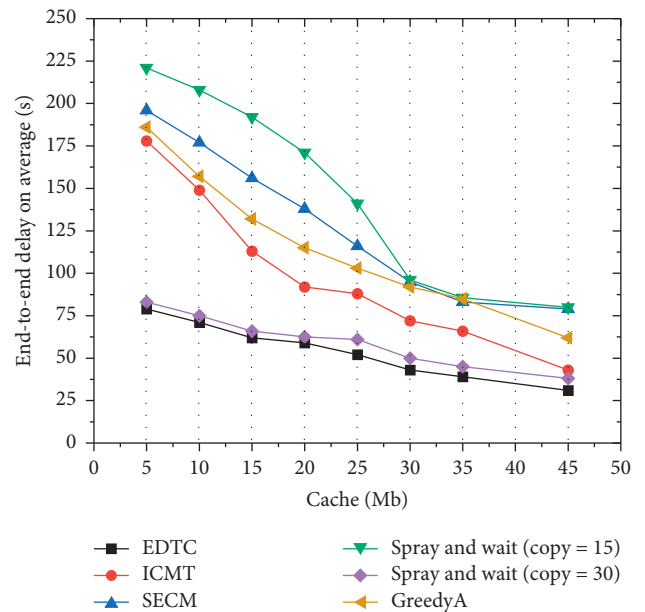


FIGURE 19: Relationship between end-to-end delay on average and cache.

increase in the node cache. However, in the case of copy = 30, the mean delay is significantly lower than the copy = 15.

In actual social networks, the choice of information transfer methods also has a great impact on the performance of the algorithm. So in order to test our proposed EDTC algorithm, in the following simulation experiments, we choose three different models to evaluate the performance of the EDTC algorithm. These three models are SPMBM (shortest path map-based movement), random way point (RWP), and random walk (RM) models [48].

Figure 20 shows the change in deliver ratio in the EDTC algorithm of different mobile models. It can be found from

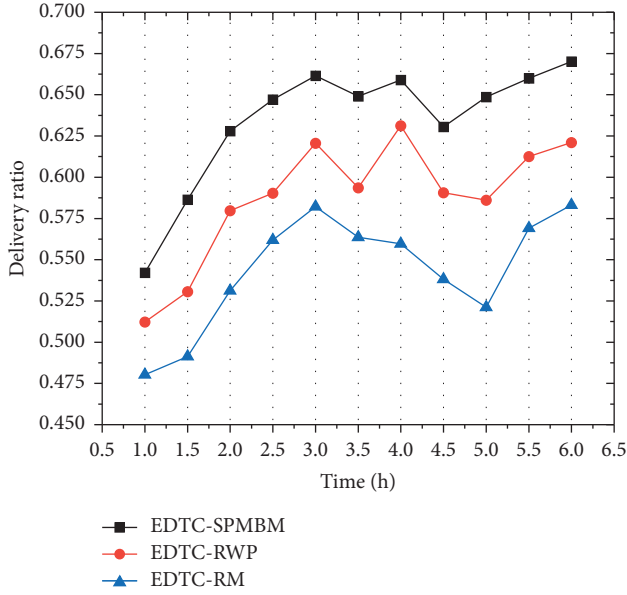


FIGURE 20: Relationship between delivery ratio and time in three mobile models.

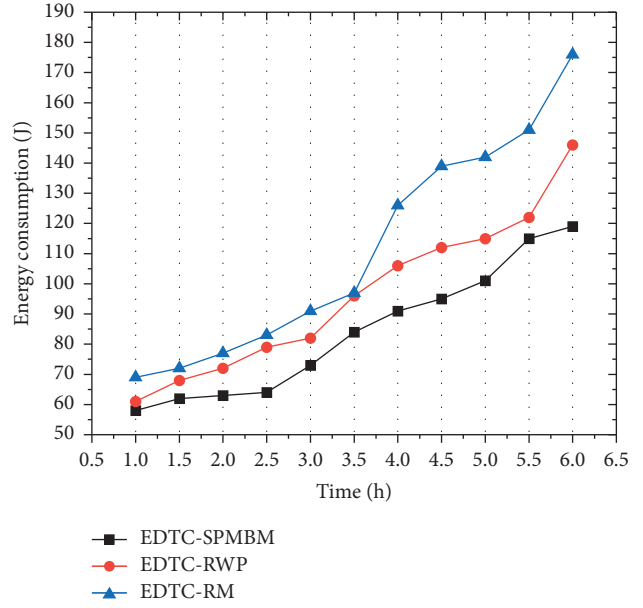


FIGURE 22: Relationship between energy consumption and time in three mobile models.

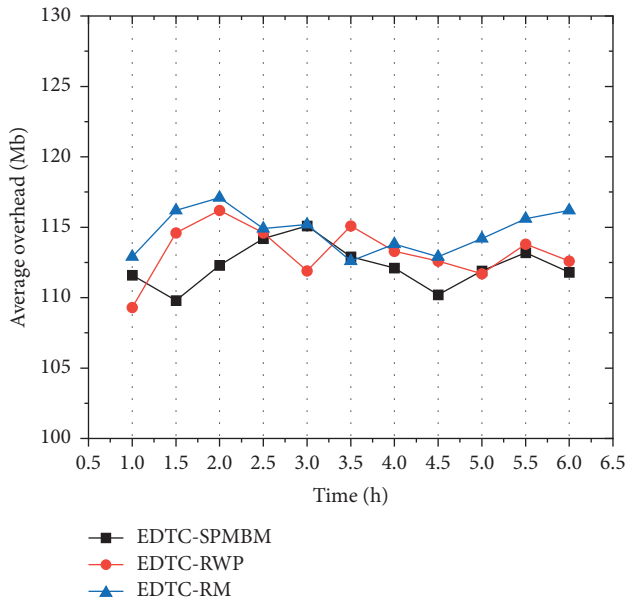


FIGURE 21: Relationship between average delay and time in three mobile models.

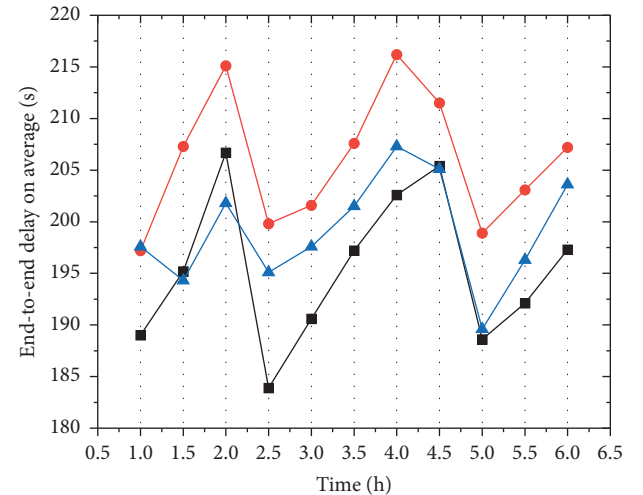


FIGURE 23: Relationship between end-to-end delay on average and time in three mobile models.

Figure 20 that the delivery ratio obtained under the SPMBM model is the highest, finally reaching 0.679. For the experiment using the RM model, the delivery ratio reached the peak 0.583 in the 6th hour. However, in the case of applying the RWP model, the peak 0.626 reached at the 4th hour. In general, the delivery ratio of the EDTC algorithm in the SPMBM model is higher than in RM and RWP models. Moreover, the RWP model is better than the RM model.

We can get the transformation of EDTC algorithm's routing overhead in different models from Figure 21. In general, in the process of applying EDTC algorithm for information transmission under these three models, the overhead

change with time is relatively small, and it remains in the range of 108~117. This experiment shows that the choice of different models has little effect on the delay caused by the EDTC algorithm for information transmission.

Figure 22 shows the difference in energy consumption over time for different models. In general, the energy consumption in all three models increases with time. The difference in energy consumption between the three models is small. The results show that the EDTC algorithm has stable node information transmission performance and does not consume a lot of energy when the model changes.

We can get the information of EDTC algorithm's average delay in different models from Figure 23. The mean delay obtained by transmission under the three models is almost floating between 183 and 216, which shows that the EDTC algorithm can effectively transfer information.

## 5. Conclusion

In this paper, we propose an opportunistic complex networks data transmission and control based on successive interference cancellation techniques. This algorithm performs broadcast link scheduling through layer-by-layer scheduling and interference avoidance scheduling. The link scheduling strategy is designed in conjunction with successive interference cancellation techniques to increase the number of broadcast links that can be transmitted simultaneously. This effectively solves the problem of signal interference during data transmission in an opportunity-complex network. And in the experimental stage, we compared the proposed algorithm with other classic algorithms of opportunistic complex networks. Experimental results show that the algorithm has good transmission ability. And relative to the GreedyA algorithm, because it increases the number of broadcast links that can be transmitted concurrently, it achieves better transmission performance. Subsequently, with the purpose of testing the impact of different mobile models on the EDTC algorithm, we, respectively, tested the algorithm based on the SPMBM, RWP, and RM models. Experimental results show that the algorithm has excellent performance in each model. Applying this algorithm to information transmission in opportunistic complex networks can reduce node energy consumption and propagation delay and greatly improve data transmission efficiency. In the future work, this method can adopt to big data environment to solve the problem in transmission.

## Data Availability

Data used to support the findings of this study are currently under embargo, while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (61672540) and Hunan Provincial Natural Science Foundation of China (2018JJ3299 and 2018JJ3682).

## References

- [1] J. Wu, Z. Chen, and M. Zhao, "Information cache management and data transmission algorithm in opportunistic social networks," *Wireless Networks*, vol. 25, no. 6, pp. 2977–2988, 2019.
- [2] Y. Wang and J. Wu, "Social-tie-based information dissemination in mobile opportunistic social networks," in *Proceedings of the 2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks"*, Madrid, Spain, June 2013.
- [3] C. Jiang, Y. Chen, and K. J. R. Liu, "Evolutionary dynamics of information diffusion over social networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 17, pp. 4573–4586, 2014.
- [4] J. Luo, J. Wu, and Y. Wu, "Advanced data delivery strategy based on multiperceived community with IoT in social complex networks," *Complexity*, vol. 2020, Article ID 3576542, 15 pages, 2020.
- [5] Q. Xu, Z. Su, K. Zhang, P. Ren, and X. S. Shen, "Epidemic information dissemination in mobile social networks with opportunistic links," *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 3, pp. 399–409, 2015.
- [6] Z. Li, F. Xiong, X. Wang, Z. Guan, and H. Chen, "Mining heterogeneous influence and indirect trust for recommendation," *Access IEEE*, vol. 8, pp. 21282–21290, 2020.
- [7] J. Wu, Z. Chen, and M. Zhao, "Effective information transmission based on socialization nodes in opportunistic networks," *Computer networks*, vol. 129, pp. 297–305, 2017.
- [8] K. Zhu, W. Li, X. Fu, and L. Zhang, "Data routing strategies in opportunistic mobile social networks: taxonomy and open challenges," *Computer Networks*, vol. 93, pp. 183–198, 2015.
- [9] A. Socievole, E. Yoneki, F. De Rango, and J. Crowcroft, "ML-SOR: message routing using multi-layer social networks in opportunistic communications," *Computer Networks*, vol. 81, pp. 201–219, 2015.
- [10] J. Wu, Z. Chen, and M. Zhao, "Community recombination and duplication node traverse algorithm in opportunistic social networks," *Peer-To-Peer Networking and Applications*, vol. 13, no. 3, pp. 940–947, 2020.
- [11] C. Jiang, Y. Chen, and K. J. R. Liu, "Graphical evolutionary game for information diffusion over social networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 524–536, 2014.
- [12] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, Article ID 6325654, 12 pages, 2019.
- [13] M. R. Schurgot, C. Comaniciu, and K. Jaffres-Runser, "Beyond traditional DTN routing: social networks for opportunistic communication," *IEEE Communications Magazine*, vol. 50, no. 7, pp. 155–162, 2012.
- [14] D. Zhang, Z. Wang, B. Guo, X. Zhou, and V. Raychoudhury, "A dynamic community creation mechanism in opportunistic mobile social networks," in *Proceedings of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, October 2011.
- [15] J. Wu, Z. Chen, and M. Zhao, "An efficient data packet iteration and transmission algorithm in opportunistic social networks," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–17, 2019.
- [16] R. Gandhi, A. Mishra, and S. Parthasarathy, "Minimizing broadcast latency and redundancy in ad hoc networks," *IEEE/ACM Transactions On Networking*, vol. 16, no. 4, pp. 840–851, 2008.
- [17] S. C.-H. Huang, H.-C. Wu, and S. S. Iyengar, "Multisource broadcast in wireless networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 10, pp. 1908–1914, 2012.
- [18] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5354–5374, 2012.

- [19] D. Zhao, K.-W. Chin, and R. Raad, "Approximation algorithms for broadcasting in duty cycled wireless sensor networks," *Wireless Networks*, vol. 20, no. 8, pp. 2219–2236, 2014.
- [20] F. Xiong, Y. Liu, L. Wang, and X. Wang, "Analysis and application of opinion model with multiple topic interactions," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 8, Article ID 083113, 2017.
- [21] F. Xiong and Z.-Y. Li, "Effective methods of restraining diffusion in terms of epidemic dynamics," *Scientific Reports*, vol. 7, no. 1, pp. 1–14, 2017.
- [22] F. Xiong, Y. Liu, and J. Cheng, "Modeling and predicting opinion formation with trust propagation in online social networks," *Communications in Nonlinear Science and Numerical Simulation*, vol. 44, pp. 513–524, 2017.
- [23] J. Wu, G. Yu, and P. Guan, "Interest characteristic probability predicted method in social opportunistic networks," *IEEE ACCESS*, vol. 7, no. 1, pp. 59002–59012, 2019.
- [24] D. Yu, Q.-S. Hua, Y. Wang, J. Yu, and F. C. M. Lau, "Efficient distributed multiple-message broadcasting in unstructured wireless networks," in *Proceedings of the 2013 Proceedings IEEE INFOCOM*, April 2013.
- [25] J. Wu, X. Tian, and Y. Tan, "Hospital evaluation mechanism based on mobile health for IoT system in social networks," *Computers in Biology and Medicine*, vol. 109, pp. 138–147, 2019.
- [26] S. Pan, R. Hu, S.-F. Fung, G. Long, J. Jiang, and C. Zhang, "Learning graph embedding with adversarial training methods," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2475–2487, 2019.
- [27] S. Pan, R. Hu, G. Long, J. Jiang, L. Yao, and C. Zhang, "Adversarially regularized graph autoencoder for graph embedding," 2018, <http://arxiv.org/abs/1802.04407>.
- [28] L. Wang, Z. Yu, F. Xiong, D. Yang, S. Pan, and Z. Yan, "Influence spread in geo-social networks: a multiobjective optimization perspective," *IEEE Transactions on Cybernetics*, pp. 1–13, 2019.
- [29] L. Wang, Z. Yu, D. Yang, T. Ku, B. Guo, and H. Ma, "Collaborative mobile crowdsensing in opportunistic D2D networks," *ACM Transactions on Sensor Networks*, vol. 15, no. 3, pp. 1–30, 2019.
- [30] H. Liang and T. Baldwin, "A probabilistic rating auto-encoder for personalized recommender systems," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, Melbourne, Australia, 2015.
- [31] H. Chen, X. Wang, S. Pan, and F. Xiong, "Identify topic relations in scientific literature using topic modeling," *IEEE Transactions on Engineering Management*, pp. 1–13, 2019.
- [32] A. Vahdat and D. Becker, "Epidemic routing for partially connected Ad Hoc networks," Technical Report CS-2000-06, Duke University, Durham, NC, USA, 2000.
- [33] H. Lenando and M. Alrfaay, "EpSoc: social-based epidemic-based routing protocol in opportunistic mobile social network," *Mobile Information Systems*, vol. 2018, Article ID 6462826, 8 pages, 2018.
- [34] P. Mundur, M. Seligman, and G. Lee, "Epidemic routing with immunity in delay tolerant networks," in *Proceedings of the MILCOM 2008-2008 IEEE Military Communications Conference*, November 2008.
- [35] F. D. Rango, S. Amelio, and P. Fazio, "Enhancements of epidemic routing in delay tolerant networks from an energy perspective," in *Proceedings of the International Wireless Communications & Mobile Computing Conference*, Sardinia, Italy, July 2013.
- [36] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-Tolerant Networking*, Philadelphia, PA, USA, 2005.
- [37] J. Xue, X. Fan, Y. Cao, J. Fang, and J. Li, "Spray and wait routing based on average delivery probability in delay tolerant network," in *Proceedings of the 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, vol. 2, April 2009.
- [38] W. Huang, S. Zhang, and W. Zhou, "Spray and wait routing based on position prediction in Opportunistic Complex Networks," in *Proceedings of the 2011 3rd International Conference on Computer Research and Development*, vol. 2, March 2011.
- [39] S. Jain, M. Chawla, V. N. G. J. Soares, and J. J. Rodrigues, "Enhanced fuzzy logic-based spray and wait routing protocol for delay tolerant networks," *International Journal of Communication Systems*, vol. 29, no. 12, pp. 1820–1843, 2016.
- [40] S. K. Dhurandher, D. K. Sharma, I. Woungang, and S. Bhati, "HBPR: history based prediction for routing in infrastructure-less Opportunistic Complex Networks," in *Proceedings of the 2013 IEEE 27th International Conference on Advanced Information Networking and Applications (AINA)*, March 2013.
- [41] C. Yu, Z. Tu, D. Yao, F. Lu, and H. Jin, "Probabilistic routing algorithm based on contact duration and message redundancy in delay tolerant network," *International Journal of Communication Systems*, vol. 29, no. 16, pp. 2416–2426, 2016.
- [42] G. Yu, Z. Chen, and J. Wu, "Quantitative social relations based on trust routing algorithm in opportunistic social network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, 2019.
- [43] F. Xu, Q. Xu, Z. Xiong et al., "Intelligent distributed routing scheme based on social similarity for mobile social networks," *Future Generation Computer Systems*, vol. 96, pp. 472–480, 2019.
- [44] X. Tian, J. Yu, L. Ma, G. Li, and X. Cheng, "Distributed deterministic broadcasting algorithms under the sinr model," in *Proceedings of the IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*, April 2016.
- [45] N. I. Miridakis and D. D. Vergados, "A survey on the successive interference cancellation performance for single-antenna and multiple-antenna OFDM systems," *IEEE Communications Surveys & Tutorials*, vol. 15, pp. 312–335, 2012.
- [46] T. Liu, J. Tong, Q. Guo, J. Xi, Y. Yu, and Z. Xiao, "Energy efficiency of uplink massive MIMO systems with successive interference cancellation," *IEEE Communications Letters*, vol. 21, no. 3, pp. 668–671, 2016.
- [47] J. Wu, Z. Chen, and M. Zhao, "SECM: status estimation and cache management algorithm in opportunistic networks," *The Journal of Supercomputing*, vol. 75, no. 5, pp. 2629–2647, 2019.
- [48] J. Wu and Z. Chen, "Data decision and transmission based on mobile data health records on sensor devices in wireless networks," *Wireless Personal Communications*, vol. 90, no. 4, pp. 2073–2087, 2016.