

Wireless Communications and Mobile Computing

Security and Privacy Challenges for Internet-of-Things and Fog Computing 2020

Lead Guest Editor: Ximeng Liu

Guest Editors: Yang Yang, Kim-Kwang Raymond Choo, and Huaqun Wang





Security and Privacy Challenges for Internet-of-Things and Fog Computing 2020

Wireless Communications and Mobile Computing

**Security and Privacy Challenges for
Internet-of-Things and Fog Computing
2020**

Lead Guest Editor: Ximeng Liu


Guest Editors: Yang Yang, Kim-Kwang Raymond
Choo, and Huaqun Wang





Copyright © 2020 Hindawi Limited. All rights reserved.

This is a special issue published in “Wireless Communications and Mobile Computing.” 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

Zhipeng Cai , USA

Associate Editors

Ke Guan , China
Jaime Lloret , Spain
Maode Ma , Singapore

Academic Editors

Muhammad Inam Abbasi, Malaysia
Ghufran Ahmed , Pakistan
Hamza Mohammed Ridha Al-Khafaji ,
Iraq
Abdullah Alamoodi , Malaysia
Marica Amadeo, Italy
Sandhya Aneja, USA
Mohd Dilshad Ansari, India
Eva Antonino-Daviu , Spain
Mehmet Emin Aydin, United Kingdom
Parameshchhari B. D. , India
Kalapaveen Bagadi , India
Ashish Bagwari , India
Dr. Abdul Basit , Pakistan
Alessandro Bazzi , Italy
Zdenek Becvar , Czech Republic
Nabil Benamar , Morocco
Olivier Berder, France
Petros S. Bithas, Greece
Dario Bruneo , Italy
Jun Cai, Canada
Xuesong Cai, Denmark
Gerardo Canfora , Italy
Rolando Carrasco, United Kingdom
Vicente Casares-Giner , Spain
Brijesh Chaurasia, India
Lin Chen , France
Xianfu Chen , Finland
Hui Cheng , United Kingdom
Hsin-Hung Cho, Taiwan
Ernestina Cianca , Italy
Marta Cimitile , Italy
Riccardo Colella , Italy
Mario Collotta , Italy
Massimo Condoluci , Sweden
Antonino Crivello , Italy
Antonio De Domenico , France
Floriano De Rango , Italy


Antonio De la Oliva , Spain
Margot Deruyck, Belgium
Liang Dong , USA
Praveen Kumar Donta, Austria
Zhuojun Duan, USA
Mohammed El-Hajjar , United Kingdom
Oscar Esparza , Spain
Maria Fazio , Italy
Mauro Femminella , Italy
Manuel Fernandez-Veiga , Spain
Gianluigi Ferrari , Italy
Luca Foschini , Italy
Alexandros G. Fragkiadakis , Greece
Ivan Ganchev , Bulgaria
Óscar García, Spain
Manuel García Sánchez , Spain
L. J. García Villalba , Spain
Miguel Garcia-Pineda , Spain
Piedad Garrido , Spain
Michele Girolami, Italy
Mariusz Glabowski , Poland
Carles Gomez , Spain
Antonio Guerrieri , Italy
Barbara Guidi , Italy
Rami Hamdi, Qatar
Tao Han, USA
Sherief Hashima , Egypt
Mahmoud Hassaballah , Egypt
Yejun He , China
Yixin He, China
Andrej Hrovat , Slovenia
Chunqiang Hu , China
Xuexian Hu , China
Zhenghua Huang , China
Xiaohong Jiang , Japan
Vicente Julian , Spain
Rajesh Kaluri , India
Dimitrios Katsaros, Greece
Muhammad Asghar Khan, Pakistan
Rahim Khan , Pakistan
Ahmed Khattab, Egypt
Hasan Ali Khattak, Pakistan
Mario Kolberg , United Kingdom
Meet Kumari, India
Wen-Cheng Lai , Taiwan

Jose M. Lanza-Gutierrez, Spain
Pavlos I. Lazaridis , United Kingdom
Kim-Hung Le , Vietnam
Tuan Anh Le , United Kingdom
Xianfu Lei, China
Jianfeng Li , China
Xiangxue Li , China
Yaguang Lin , China
Zhi Lin , China
Liu Liu , China
Mingqian Liu , China
Zhi Liu, Japan
Miguel López-Benítez , United Kingdom
Chuanwen Luo , China
Lu Lv, China
Basem M. ElHalawany , Egypt
Imadeldin Mahgoub , USA
Rajesh Manoharan , India
Davide Mattera , Italy
Michael McGuire , Canada
Weizhi Meng , Denmark
Klaus Moessner , United Kingdom
Simone Morosi , Italy
Amrit Mukherjee, Czech Republic
Shahid Mumtaz , Portugal
Giovanni Nardini , Italy
Tuan M. Nguyen , Vietnam
Petros Nicolaitidis , Greece
Rajendran Parthiban , Malaysia
Giovanni Pau , Italy
Matteo Petracca , Italy
Marco Picone , Italy
Daniele Pinchera , Italy
Giuseppe Piro , Italy
Javier Prieto , Spain
Umair Rafique, Finland
Maheswar Rajagopal , India
Sujan Rajbhandari , United Kingdom
Rajib Rana, Australia
Luca Reggiani , Italy
Daniel G. Reina , Spain
Bo Rong , Canada
Mangal Sain , Republic of Korea
Praneet Saurabh , India

Hans Schotten, Germany
Patrick Seeling , USA
Muhammad Shafiq , China
Zaffar Ahmed Shaikh , Pakistan
Vishal Sharma , United Kingdom
Kaize Shi , Australia
Chakchai So-In, Thailand
Enrique Stevens-Navarro , Mexico
Sangeetha Subbaraj , India
Tien-Wen Sung, Taiwan
Suhua Tang , Japan
Pan Tang , China
Pierre-Martin Tardif , Canada
Sreenath Reddy Thummaluru, India
Tran Trung Duy , Vietnam
Fan-Hsun Tseng, Taiwan
S Velliangiri , India
Quoc-Tuan Vien , United Kingdom
Enrico M. Vitucci , Italy
Shaohua Wan , China
Dawei Wang, China
Huaqun Wang , China
Pengfei Wang , China
Dapeng Wu , China
Huaming Wu , China
Ding Xu , China
YAN YAO , China
Jie Yang, USA
Long Yang , China
Qiang Ye , Canada
Changyan Yi , China
Ya-Ju Yu , Taiwan
Marat V. Yuldashev , Finland
Sherali Zeadally, USA
Hong-Hai Zhang, USA
Jiliang Zhang, China
Lei Zhang, Spain
Wence Zhang , China
Yushu Zhang, China
Kechen Zheng, China
Fuhui Zhou , USA
Meiling Zhu, United Kingdom
Zhengyu Zhu , China

Contents

An Efficient Pairing-Free Certificateless Searchable Public Key Encryption for Cloud-Based IIoT

Mimi Ma, Min Luo , Shuqin Fan, and Dengguo Feng


Research Article (11 pages), Article ID 8850520, Volume 2020 (2020)

Privacy-Preserving Vehicle Assignment in the Parking Space Sharing System

Tingting Fu, Peng Liu , Kun Liu, and Peng Li



Research Article (13 pages), Article ID 8862652, Volume 2020 (2020)

A Privacy-Preserving Personalized Service Framework through Bayesian Game in Social IoT

Renwan Bi, Qianxin Chen, Lei Chen, Jinbo Xiong , and Dapeng Wu


Research Article (13 pages), Article ID 8891889, Volume 2020 (2020)

Social-Aware Task Allocation in Mobile Crowd Sensing

Weiping Zhu , Wenzhong Guo , and Zhiyong Yu

Research Article (15 pages), Article ID 8822251, Volume 2020 (2020)

A Novel Attack-and-Defense Signaling Game for Optimal Deceptive Defense Strategy Choice

Yongjin Hu, Han Zhang, Yuanbo Guo, Tao Li, and Jun Ma 

Research Article (10 pages), Article ID 8850356, Volume 2020 (2020)

A Host-Based Anomaly Detection Framework Using XGBoost and LSTM for IoT Devices

Xiali Wang  and Xiang Lu 



Research Article (13 pages), Article ID 8838571, Volume 2020 (2020)

TLFW: A Three-Layer Framework in Wireless Rechargeable Sensor Network with a Mobile Base Station

Anwen Wang , Xianjia Meng , Lvju Wang, Xiang Ji, Hao Chen, Baoying Liu, Feng Chen, Yajuan Du, and Guangcheng Yin



Research Article (12 pages), Article ID 3627826, Volume 2020 (2020)

BEHT: Blockchain-Based Efficient Highway Toll Paradigm for Opportunistic Autonomous Vehicle Platoon

Zuobin Ying , Longyang Yi , and Maode Ma


Research Article (13 pages), Article ID 8868656, Volume 2020 (2020)

Achieving Privacy-Preserving Group Recommendation with Local Differential Privacy and Random Transmission

Hanyi Wang , Kun He, Ben Niu , Lihua Yin, and Fenghua Li

Research Article (10 pages), Article ID 8836351, Volume 2020 (2020)

Secure and Intelligent Energy Data Management Scheme for Smart IoT Devices

Tianqi Zhou, Jian Shen , Sai Ji, Yongjun Ren, and Leiming Yan


Research Article (11 pages), Article ID 8842885, Volume 2020 (2020)

A Data Encryption and Fast Transmission Algorithm Based on Surveillance Video

Shi Qiu , Ying Cui , and XianJia Meng 

Research Article (12 pages), Article ID 8842412, Volume 2020 (2020)

Location Privacy-Preserving Method Based on Historical Proximity Location

Xueying Guo, Wenming Wang, Haiping Huang , Qi Li, and Reza Malekian


Research Article (16 pages), Article ID 8892079, Volume 2020 (2020)

An Efficient Searchable Public-Key Authenticated Encryption for Cloud-Assisted Medical Internet of Things

Tianyu Chi, Baodong Qin , and Dong Zheng

Research Article (11 pages), Article ID 8816172, Volume 2020 (2020)

Reliability Evaluation of Generalized Exchanged χ -Cubes Based on the Condition of g -Good-Neighbor

Xiaoyan Li, Hongbin Zhuang, Shuming Zhou, Hongju Cheng, Cheng-Kuan Lin, and Wenzhong Guo 

Research Article (16 pages), Article ID 9793082, Volume 2020 (2020)

Research Article

An Efficient Pairing-Free Certificateless Searchable Public Key Encryption for Cloud-Based IIoT

Mimi Ma,^{1,2,3} Min Luo^{id},⁴ Shuqin Fan,¹ and Dengguo Feng¹

¹State Key Laboratory of Cryptology, Beijing, China

²Key Laboratory of Grain Information Processing and Control (Henan University of Technology), Ministry of Education, Zhengzhou, China

³College of Information Science and Engineering, Henan University of Technology, Zhengzhou, China

⁴School of Cyber Science and Engineering, Wuhan University, Wuhan, China

Correspondence should be addressed to Min Luo; mluo@whu.edu.cn

Received 25 June 2020; Revised 6 August 2020; Accepted 2 December 2020; Published 21 December 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Mimi Ma 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 Industrial Internet of Things (IIoT), as a special form of Internet of Things (IoT), has great potential in realizing intelligent transformation and industrial resource utilization. However, there are security and privacy concerns about industrial data, which is shared on an open channel via sensor devices. To address these issues, many searchable encryption schemes have been presented to provide both data privacy-protection and data searchability. However, due to the use of expensive pairing operations, most previous schemes were inefficient. Recently, a certificateless searchable public-key encryption (CLSPE) scheme was designed by Lu et al. to remove the pairing operation. Unfortunately, we find that Lu et al.'s scheme is vulnerable to user impersonation attacks. To enhance the security, a new pairing-free dual-server CLSPE (DS-CLSPE) scheme for cloud-based IIoT deployment is designed in this paper. In addition, we provide security and efficiency analysis for DS-CLSPE. The analysis results show that DS-CLSPE can resist chosen keyword attacks (CKA) and has better efficiency than other related schemes.

1. Introduction

The gradual maturity of communication technology, especially the emergence of the 5-th generation wireless systems, has greatly promoted the popularization of IoT, which connects everything to the Internet for intelligent identification, tracking, monitoring, etc. [1–3]. Industrial IoT (IIoT) is one of the main application directions of IoT. It can collect the industrial data in real-time via various sensing devices (e.g., global positioning system and radio frequency identification), realize the optimal utilization of resources, improve the quality, and reduce the cost of the product through further analyzing those collected data [4].

IIoT provides new technological guidance for the development of industry and has great application potential. Nowadays, the IIoT market is expanding rapidly [5]. Meanwhile, the number of IIoT devices will also grow several times, inevitably leading to an explosion of the collected industrial data.

In order to effectively manage and utilize these big data, users are more willing to outsource the data to the cloud server provider (CSP), which has powerful data storage and analytical capabilities [6]. In a cloud-based IIoT setting, as presented in Figure 1, massive industrial data is collected through sensor devices and transmitted to the CSP in real-time via the Internet. The CSP further analyzes and mines these data to provide better intelligent services for industrial sectors such as intelligent logistics and manufacturing.

Users enjoy various convenient services offered by CSP; however, their data security and privacy are seriously threatened [7–10]. One reason is that their data is transmitted on an open channel, so an adversary can eavesdrop on the transmitted data to obtain information about an enterprise's production or operations [11]. The other reason is that they will not be able to physically control the outsourced data nor will they be able to fully trust the CSP. For example, the CSP may exploit users' outsourced data to make illegal profits or carry

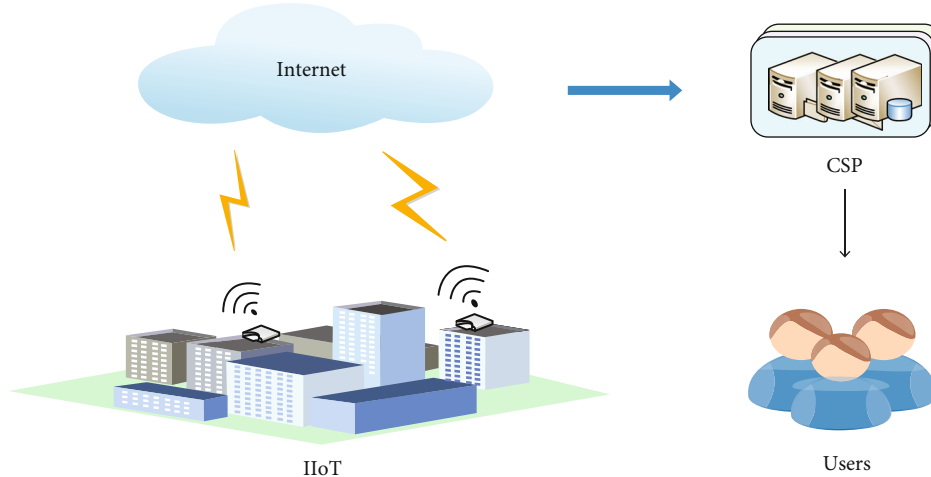


FIGURE 1: Cloud-based IIoT architecture.

out other malicious acts, such as tampering and deleting. Once the confidential data is eavesdropped or destroyed, it may cause unpredictable damage to an enterprise (e.g., huge economic losses). Overall, it is urgent to establish a practical mechanism with security and privacy preservation for IIoT data utilization and management.

Encryption is the most direct approach to guarantee data privacy. Users can first encrypt the confidential data and then submit the ciphertext to CSP. Although traditional encryption is effective in preserving the privacy of IIoT data, it also incurs some troubles in data utilizations, especially the problem of searching over encrypted data. Since the original data structure will be changed once it is encrypted, the search algorithms for plaintext will not be feasible for the encrypted data. To address this issue, the searchable encryption (SE) technology has emerged, which supports search over ciphertext according to keywords [12, 13]. The first symmetric SE (SSE) scheme was presented in [12]. However, SSE suffers from the troublesome key distribution. To resolve this problem, a SE scheme based on a public-key cryptosystem (SPE) was designed in [13]. Since then, various SPE schemes have been designed [14–17]. However, these SPE schemes face cumbersome certificate management or key escrow burden since the inherent designing structure based on public key infrastructure (PKI) and identity-based cryptosystem. The CLSPE schemes [18–20] can overcome these cumbersome burdens. However, most of the previous CLSPE schemes were computationally expensive since the use of many complex pairing operations. Recently, Lu and Li [21] presented a new CLSPE scheme without pairing operation. Unfortunately, we analyze their scheme suffer from user impersonation attack. To improve the security, we design a new DS-CLSPE scheme without pairing for cloud-based IIoT deployment.

1.1. Related Work. The SE technology provides the functionality of searching over ciphertext without losing the confidentiality of original data. The first concrete SE scheme was presented in [12], which was built on a symmetric cryptosystem. Later, various SSE schemes have been designed. A SSE scheme with verifiable functionality (VSSE) was presented

in [22], which can both protect data privacy and provide the verifiability. Recently, Zuo et al. [23] gave two dynamic SSE schemes (i.e., Scheme-A and Scheme-B), which support range queries. The former scheme has forward security property but imposes a heavy storage cost on the client, and the latter scheme reduces the client’s storage but loses the forward security. Later, a novel VSSE scheme with forward security was presented in [24]. However, all SSE schemes face the complex key management issue.

To avoid the key management issue, the concept of SPE was introduced in [13]. A SPE system contains three participants: cloud server (CS), data sender (DS), and data receiver (DR). DS uses DR’s public key to encrypt his/her own data, including files and keywords extracted from files, and sends the encrypted data to CS. DR uses its own private key to produce a trapdoor for the keyword to be retrieved and submits the trapdoor to CS. Then, CS verifies whether the trapdoor matches the ciphertext and returns the successfully matched files to DR. Later, a SPE scheme with a designated server (dSPE) was constructed, in which only the specified server can run the test algorithm [14]. Lin et al. [25] designed a novel blockchain-based system for the secure outsourcing of bilinear pairings to remove the secure channel and the trusted server. To resist inside keyword guessing attack (IKGA), an authenticated SPE scheme was constructed in [15]. Recently, a SPE scheme with forward security was designed in [16] to resist file-injection attacks [26]. However, the above-proposed schemes are built on PKI cryptosystem; they inevitably face the complicated certificate management problem.

To reduce the overhead of managing certificate, an ID-based cryptosystem was introduced, in which the participant’s public key is set to some public information (e.g., name and office number), and the private key is created by a key generation center (KGC) based on the public information [27]. A general framework for transforming a two-level anonymous ID-based encryption (IBE) scheme to an ID-based SPE (IBSPE) scheme was presented in [17]. Recently, Lu et al. [28] constructed an IBSPE with a designated server (dIBSPE), which supports conjunctive keyword search. Li et al. [29] designed two-authenticated dIBSPE schemes based

on symmetric bilinear pairing and asymmetric bilinear pairing, respectively. In their schemes, any adversary cannot run the encryption algorithm to get a valid ciphertext unless it can capture the data sender's private key, and no adversary can perform the test algorithm correctly unless it has the ability to access the specified server's private key. Zhang et al. [30] constructed a proxy-oriented IBSPE scheme based on lattices to resist quantum computer attacks, in which the original data sender delegates his/her own data to a proxy for encryption in order to lower the computation cost of himself. However, all ID-based SPE schemes are plagued by key escrow issues.

In certificateless cryptosystem (CLC), the private key of the participant is jointly created by KGC and the participant itself, which resolves the burden of key escrow and certificate management existing in PKI-based and ID-based cryptosystem [31]. Peng et al. [18] introduced SE technology into the CLC system and proposed a CLSPE scheme with a designated server (dCLSPE). To lower computation overhead in [18], Islam et al. [19] designed a new dCLSPE scheme based on the problems of CDH and BDH. He et al. [20] constructed a novel-authenticated CLSPE scheme against IKGA attacks. However, all previous CLSPE schemes involve bilinear pairing operations, which require high computation overhead. Recently, a CLSPE scheme without pairing operation is designed in [21]. Unfortunately, we analyze scheme [21] cannot resist user impersonation attacks, thus, we develop a new pairing-free CLSPE scheme.

1.2. Research Contributions. An efficient pairing-free DS-CLSPE scheme for cloud-based IIoT is designed in this paper. The main research contributions are listed below:

- (i) First, we analyze Lu et al.'s scheme is subject to user impersonation attack, and their scheme requires a secure channel for trapdoor transmission
- (ii) Second, we give the system model for DS-CLSPE and construct a new DS-CLSPE scheme, which not only eliminates the need for bilinear pairings but also removes the use of secure channel
- (iii) Finally, we present a security analysis for DS-CLSPE and show it can resist the CKA attack. Furthermore, we evaluate the efficiency of DS-CLSPE in terms of computation and communication costs

1.3. Organization of the Paper. The following sections are arranged as below. Section 2 presents some preliminary knowledge. Section 3 shows the analysis of Lu et al.'s scheme. Section 4 gives the detailed construction of our DS-CLSPE scheme. Section 5 and Section 6 show the security proof and the performance analysis of DS-CLSPE, respectively. The last section is mainly to summarize the full paper.

2. Preliminaries

We first present the complexity assumptions used in this paper and then give the system model and formal definition of DS-CLSPE.

2.1. Complexity Assumptions. Suppose \mathbb{G} denotes a q -order cyclic group, and the point $P \in \mathbb{G}$ denotes a generator.

2.1.1. Computational Diffie-Hellman (CDH) Assumption. Given three points $P, aP, bP \in \mathbb{G}$ ($a, b \in \mathbb{Z}_q^*$ are unknown numbers), to figure out abP .

2.2. System Model. As presented in Figure 2, a DS-CLSPE system contains five participants: data sender (DS), data receiver (DR), KGC, front server (FS), and back server (BS). The responsibilities of each participant are described below.

- (i) KGC is responsible for generating system parameters and participants' partial keys
- (ii) DS encrypts his/her own data and then outsources the encrypted data to FS and BS
- (iii) DR submits a trapdoor to FS for querying the encrypted data
- (iv) FS generates an intermediate testing-state ciphertext C_{ITS} according to the received trapdoor and submits C_{ITS} to BS
- (v) BS generates the final test results according to C_{ITS} and returns the test results to DR

2.3. Formal Definition. A DS-CLSPE scheme contains the following algorithms.

- (i) Setup(λ): this algorithm is implemented by KGC. Inputs a security parameter λ , returns the system master key s and public parameters parm
- (ii) PartialKeyGen($\text{parm}, s, \text{ID}_i$): this algorithm is performed by KGC. Inputs parm , s , and the identity $\text{ID}_i (i \in \{\text{DR}, \text{FS}, \text{BS}\})$, outputs the partial key pair (T_i, d_i) for the corresponding participant
- (iii) KeyGen($\text{parm}, \text{ID}_i, T_i, d_i$): each participant generates its own full public/private keys PK_i/SK_i by performing this algorithm
- (iv) Encrypt($\text{parm}, w, \text{ID}_i, \text{PK}_i$): DS executes this algorithm to generate the ciphertext C_w for keyword w
- (v) Trapdoor($\text{parm}, w', \text{SK}_{\text{DR}}, \text{PK}_{\text{FS}}, \text{PK}_{\text{BS}}$): DR performs this algorithm to obtain the trapdoor $T_{w'}$ for w' .
- (vi) FrontTest($\text{parm}, C_w, T_{w'}, \text{SK}_{\text{FS}}$): FS performs this algorithm to generate an intermediate testing-state C_{ITS}
- (vii) BackTest($\text{parm}, C_{ITS}, \text{SK}_{\text{BS}}$): this algorithm is performed by BS. Inputs parm , C_{ITS} , SK_{BS} , outputs "1" if the test succeeds and "0" otherwise

3. Weakness of Lu et al.'s Scheme

We first present Lu et al.'s scheme and then analyze its security weakness.

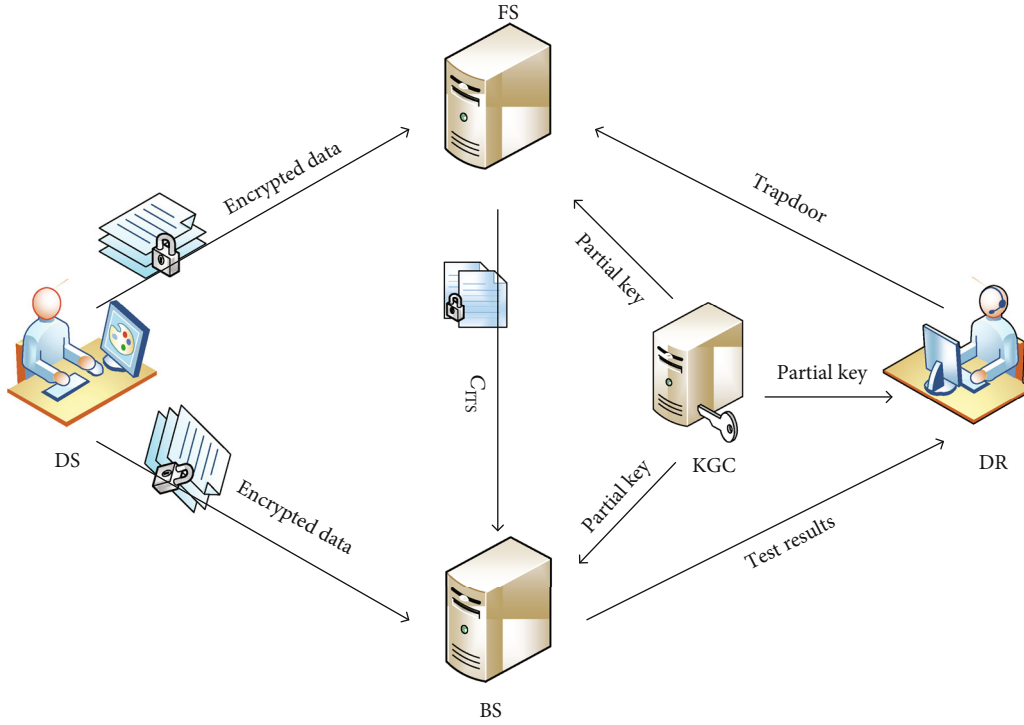


FIGURE 2: The system model for DS-CLSPE.

3.1. Review of Lu et al.'s Scheme. The detailed construction of Lu et al.'s scheme is as follows.

(i) Setup(λ): suppose q is a large prime number and \mathbb{G} is a group with order q . P denotes a generator selected from \mathbb{G} , and $h_1 : \{0, 1\}^* \rightarrow \mathbb{Z}_q^*$, $h_2 : \{0, 1\}^* \rightarrow \mathbb{Z}_q^*$, and $h_3 : \mathbb{G} \rightarrow \mathbb{Z}_q^*$ denote three different hash functions. KGC performs this algorithm as follows

- (1) Selects $s \in \mathbb{Z}_q^*$ at random
- (2) Calculates $P_{\text{pub}} = sP$
- (3) Publishes $\text{parm} = \{\mathbb{G}, P, P_{\text{pub}}, q, h_1, h_2, h_3\}$ and keeps s in secret

(ii) PartialKeyGen($\text{parm}, s, \text{ID}_{\text{DR}}$): KGC selects $t_{\text{DR}} \in \mathbb{Z}_q^*$, computes $T_{\text{DR}} = t_{\text{DR}}P$, $d_{\text{DR}} = t_{\text{DR}} + sh_1(\text{ID}_{\text{DR}})$, and sends the partial key pair $\{T_{\text{DR}}, d_{\text{DR}}\}$ to the receiver

(iii) KeyGen($\text{parm}, \text{ID}_{\text{DR}}, T_{\text{DR}}, d_{\text{DR}}$): the receiver picks $x_{\text{DR}} \in \mathbb{Z}_q^*$, calculates $P_{\text{DR}} = x_{\text{DR}}P$, and sets $\text{PK}_{\text{DR}} = (T_{\text{DR}}, P_{\text{DR}})$, $\text{SK}_{\text{DR}} = (d_{\text{DR}}, x_{\text{DR}})$ as his/her own public/private keys

(iv) Encrypt($\text{parm}, w, \text{PK}_{\text{DR}}$): the sender randomly chooses $r \in \mathbb{Z}_q^*$, calculates $C_w^1 = rP$ and

$$C_w^2 = h_3(rh_2(w)Q_{\text{DR}}), \quad (1)$$

here $Q_{\text{DR}} = T_{\text{DR}} + P_{\text{DR}} + h_1(\text{ID}_{\text{DR}})P_{\text{pub}}$, and sets the ciphertext $C_w = (C_w^1, C_w^2)$.

(v) Trapdoor($\text{parm}, w', \text{SK}_{\text{DR}}$): the receiver computes

$$T_w' = (d_{\text{DR}} + x_{\text{DR}})h_2(w'). \quad (2)$$

(vi) Test(parm, C_w, T_w'): given C_w and T_w' , the server checks

$$C_w^2 \stackrel{?}{=} h_3(T_w' C_w^1). \quad (3)$$

If this equation holds, returns "1"; Otherwise, returns "0".

3.2. Attack on Lu et al.'s Scheme. Lu et al.'s scheme cannot prevent user impersonation attack launched by TypeI adversary \mathcal{A}_1 . During the attack, \mathcal{A}_1 can first forge the private key of the receiver ID_{DR} and then impersonate ID_{DR} to calculate a trapdoor T_w for a challenge keyword w using the forged private key. The detail attack is presented below.

(i) Setup: the challenger \mathcal{C} generates $\{\text{parm}, s\}$ by performing the Setup algorithm

(ii) Queries: \mathcal{A}_1 with identity $\text{ID}_{\mathcal{A}_1}$ can ask the following queries

- (1) Extract Partial Private Key Query: if \mathcal{A}_1 submits this query, then \mathcal{C} returns $\{d_{\mathcal{A}_1}, T_{\mathcal{A}_1}\}$ to \mathcal{A}_1 , where $d_{\mathcal{A}_1} = t_{\mathcal{A}_1} + sh_1(\text{ID}_{\mathcal{A}_1})$ and $T_{\mathcal{A}_1} = t_{\mathcal{A}_1}P$

Upon receiving $\{d_{\mathcal{A}_1}, T_{\mathcal{A}_1}\}$, \mathcal{A}_1 can forge the partial key pair $(d_{\text{DR}}, T_{\text{DR}})$ of the receiver ID_{DR} as follows:

$$T_{\text{DR}}^* = T_{\mathcal{A}_1} \cdot h_1(\text{ID}_{\text{DR}}) \cdot h_1(\text{ID}_{\mathcal{A}_1})^{-1}, \quad (4)$$

$$d_{\text{DR}}^* = d_{\mathcal{A}_1} \cdot h_1(\text{ID}_{\text{DR}}) \cdot h_1(\text{ID}_{\mathcal{A}_1})^{-1}. \quad (5)$$

(2) *Replace Public Key Query*: \mathcal{A}_1 selects $x_{\text{DR}}^* \in \mathbb{Z}_q^*$ randomly and computes $P_{\text{DR}}^* = x_{\text{DR}}^* P$. Then, \mathcal{A}_1 submits this query with $(\text{ID}_{\text{DR}}, \text{PK}_{\text{DR}}^*)$, here $P_{\text{DR}}^* = (T_{\text{DR}}^*, P_{\text{DR}}^*)$. \mathcal{C} will set $\text{PK}_{\text{DR}} \leftarrow \text{PK}_{\text{DR}}^*$

(iii) *Forge Trapdoor*: once \mathcal{A}_1 has successfully forged d_{DR}^* and replaced PK_{DR} , it can forge a trapdoor T_{w_0} for a keyword w_0 with respect to the identity ID_{DR} as below:

$$T_{w_0} = (d_{\text{DR}}^* + x_{\text{DR}}^*) h_2(w_0). \quad (6)$$

(iv) *Challenge*: \mathcal{A}_1 returns the keywords (w_0, w_1) and the identity ID_{DR} as challenge target. Then \mathcal{C} picks up $\mu \in \{0, 1\}$ randomly and returns $C_{w_\mu} \leftarrow \text{Encrypt}(\text{parm}, w_\mu, \text{PK}_{\text{DR}})$ to \mathcal{A}_1

Upon receiving C_{w_μ} , \mathcal{A}_1 performs Test algorithm and gets $\tau \leftarrow \text{Test}(\text{parm}, C_{w_\mu}, T_{w_0}) \in \{0, 1\}$.

(v) *Guess*: \mathcal{A}_1 returns $\mu' = 0$ if $\tau = 1$; otherwise, \mathcal{A}_1 returns $\mu' = 1$. It is easy to see that $\mu' = \mu$, i.e., the advantage of \mathcal{A}_1 is always 1

4. The Proposed DS-CLSPE Scheme

To overcome the weakness of Lu et al.'s scheme and avoid the use of bilinear pairing, we develop a new dual-server CLSPE scheme. The details are described as follows.

(i) *Setup*(λ): suppose q is a large prime number, \mathbb{G} is a group with order q . Let P denote a generator of \mathbb{G} , and $h_0 : \{0, 1\}^* \times \mathbb{G} \rightarrow \mathbb{Z}_q^*$, $h_1 : \{0, 1\}^* \times \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{Z}_q^*$, and $h_2 : \{0, 1\}^* \times \{0, 1\}^* \rightarrow \mathbb{Z}_q^*$ denote three different hash functions. KGC performs the following steps

(1) Chooses $s \in \mathbb{Z}_q^*$ randomly

(2) Calculates $P_{\text{pub}} = sP$

(3) Publishes $\text{parm} = \{P, P_{\text{pub}}, \mathbb{G}, h_0, h_1, h_2\}$ and keeps s secretly

(ii) *PartialKeyGen*($\text{parm}, s, \text{ID}_i$): takes parm , s , and the identity $\text{ID}_i (i \in \{\text{DR}, \text{FS}, \text{BS}\})$ as inputs, KGC performs this algorithm as follows

(1) Selects $t_i \in \mathbb{Z}_q^*$

(2) Calculates $T_i = t_i P$, $d_i = t_i + s\alpha_i \bmod q$, where $\alpha_i = h_0(\text{ID}_i, T_i)$.

(3) Sends $\{T_i, d_i\}$ to the corresponding participant

(iii) *KeyGen*($\text{parm}, \text{ID}_i, T_i, d_i$): this algorithm is performed by the participant $i \in \{\text{DR}, \text{FS}, \text{BS}\}$ as follows

(1) Chooses $x_i \in \mathbb{Z}_q^*$ randomly

(2) Calculates $P_i = x_i P$

(3) Sets $\text{PK}_i = (T_i, P_i)$ and $\text{SK}_i = (d_i, x_i)$

(iv) *Encrypt*($\text{parm}, w, \text{ID}_i, \text{PK}_i$): given ID_i , $\text{PK}_i = (T_i, P_i)$ ($i \in \{\text{DR}, \text{FS}, \text{BS}\}$) and keyword w . DR performs the steps below to produce ciphertext C_w

(1) Selects $r \in \mathbb{Z}_q^*$ randomly

(2) Calculates $C_w^1 = rP$

(3) Calculates

$$C_w^2 = rQ_{\text{FS}} + rQ_{\text{BS}} + h_2(w, \text{ID}_{\text{DR}})Q_{\text{DR}}, \quad (7)$$

where

$$Q_i = T_i + \beta_i P_i + \alpha_i P_{\text{pub}}, \quad (8)$$

$$\alpha_i = h_0(\text{ID}_i, T_i), \quad (9)$$

$$\beta_i = h_1(\text{ID}_i, T_i, P_i). \quad (10)$$

(4) Sets $C_w = (C_w^1, C_w^2)$

The parameters $\{Q_i, \alpha_i, \beta_i\}$ can be published publicly.

(v) *Trapdoor*($\text{parm}, w', \text{SK}_{\text{DR}}, \text{PK}_{\text{FS}}, \text{PK}_{\text{BS}}$): the receiver generates the trapdoor $T_{w'}$ of keyword w' as below

(1) Selects $l \in \mathbb{Z}_q^*$ randomly

(2) Computes $T_{w'}^1 = lP$

(3) Computes

$$T_{w'}^2 = (d_{\text{DR}} + \beta_{\text{DR}} x_{\text{DR}}) h_2(w', \text{ID}_{\text{DR}}) Q_{\text{FS}} + lQ_{\text{BS}}, \quad (11)$$

where Q_i and β_i are computed as above.

(4) Sends $T_{w'} = (T_{w'}^1, T_{w'}^2)$ to the front server

(vi) *FrontTest*($\text{parm}, C_w, T_{w'}, \text{SK}_{\text{FS}}$): given the ciphertext C_w and a trapdoor $T_{w'}$, the front server FS performs the following steps to generate intermediate testing-state C_{ITS}

(1) Selects $\gamma \in \mathbb{Z}_q^*$ randomly

(2) Calculates $C^* = (d_{\text{FS}} + \beta_{\text{FS}} x_{\text{FS}}) C_w^1$

(3) Calculates

$$C_1^* = \gamma(C^* - T_{w'}^1), \quad (12)$$

$$C_2^* = \gamma((d_{FS} + \beta_{FS}x_{FS})(C_w^2 - C^*) - T_{w'}^2). \quad (13)$$

(4) Sends $C_{ITS} = (C_1^*, C_2^*)$ to the back server BS

(vii) BackTest(parm, C_{ITS} , SK_{BS}): the back server checks

$$C_2^* \stackrel{?}{=} (d_{BS} + \beta_{BS}x_{BS})C_1^*. \quad (14)$$

If this equation holds, BS returns “1”; otherwise, returns “0”.

Correctness. Suppose that $w = w'$, then we have

$$C^* = (d_{FS} + \beta_{FS}x_{FS})C_w^1 = rQ_{FS}, \quad (15)$$

$$C_1^* = \gamma(C^* - T_{w'}^1) = \gamma(rQ_{FS} - lP), \quad (16)$$

$$\begin{aligned} C_2^* &= \gamma((d_{FS} + \beta_{FS}x_{FS})(C_w^2 - C^*) - T_{w'}^2) \\ &= \gamma((d_{FS} + \beta_{FS}x_{FS})(rQ_{BS} + h_2(w, ID_{DR})Q_{DR}) - T_{w'}^2) \\ &= \gamma(r(d_{FS} + \beta_{FS}x_{FS})Q_{BS} - lQ_{BS}) \\ &= \gamma(r(d_{BS} + \beta_{BS}x_{BS})Q_{FS} - lQ_{BS}) \\ &= \gamma(d_{BS} + \beta_{BS}x_{BS})(rQ_{FS} - lP) \\ &= (d_{BS} + \beta_{BS}x_{BS})C_1^*. \end{aligned} \quad (17)$$

5. Security Analysis

This section first presents the security model of DS-CLSPE and then gives the formal proof of DS-CLSPE.

5.1. Security Model. Two types of adversaries should be considered in a certificateless cryptosystem [31, 32].

Type I. This adversary is denoted as \mathcal{A}_1 , who has no master key but can replace anyone’s public key.

Type II. This adversary is denoted as \mathcal{A}_2 , who holds the master key but cannot replace anyone’s public key.

As defined in scheme [33], assume that both BS and FS are honest-but-curious, and that they cannot collude. The security model of DS-CLSPE is defined by the following game, i.e., indistinguishability against CKA attack (IND-CKA), which is the interaction between an adversary $\mathcal{A} \in \{\mathcal{A}_1, \mathcal{A}_2\}$ and a challenger \mathcal{C} .

Definition 1 (IND-CKA). The DS-CLSPE scheme is said IND-CKA secure if \mathcal{A} ’s advantage of winning in the following game is negligible.

Game. This game is interacted between \mathcal{A} and \mathcal{C} .

(i) *Setup:* \mathcal{C} generates $\{\text{parm}, s\}$ by executing Setup algorithm. If $\mathcal{A} = \mathcal{A}_1$, \mathcal{C} returns parm to \mathcal{A} ; otherwise ($\mathcal{A} = \mathcal{A}_2$), returns both parm and s

(ii) *Phase 1:* the following oracles can be queried by \mathcal{A}

(a) *CreateUser Queries:* upon receiving this query for ID_i , \mathcal{C} checks whether ID_i has been created. If so, \mathcal{C} outputs PK_i directly. Otherwise, \mathcal{C} performs the algorithm KeyGen to produce the key pair (PK_i, SK_i) and returns PK_i

(b) *PrivateKey Queries:* upon receiving this query for ID_i , \mathcal{C} checks whether ID_i has been created. If so, \mathcal{C} returns SK_i ; otherwise, returns \perp

(c) *PatialPrivateKey Queries:* upon receiving this query for ID_i , \mathcal{C} checks whether ID_i has been created. If so, \mathcal{C} returns partial private key d_i ; otherwise, returns \perp

(d) *ReplacePublicKey Queries:* if $\mathcal{A} = \mathcal{A}_1$, then it can perform these queries. Upon receiving \mathcal{A} ’s query for ID_i with a false public key PK_i^* , \mathcal{C} sets $PK_i \leftarrow PK_i^*$

(e) *Trapdoor Queries:* if \mathcal{A} submits this query for (ID_i, w) , then \mathcal{C} returns the corresponding trapdoor T_w if ID_i has been created; otherwise, returns \perp

(iii) *Challenge:* \mathcal{A} selects identity ID_{ch} and keywords (w_0, w_1) as challenge targets. \mathcal{C} picks up $\sigma \in \{0, 1\}$ at random and returns $C_{w_\sigma} \leftarrow \text{Encrypt}(\text{parm}, w_\sigma, ID_{ch}, PK_{ch}, PK_{FS}, PK_{BS})$ to \mathcal{A}

(iv) *Phase 2:* the oracles defined as in phase 1 can be continuously queried by \mathcal{A}

(v) *Guess:* \mathcal{A} returns $\sigma' \in \{0, 1\}$. We say \mathcal{A} wins the above game if $\sigma' = \sigma$ and below conditions hold: (1) *Trapdoor* queries for w_0 and w_1 have never been submitted by \mathcal{A} ; (2) If $\mathcal{A} = \mathcal{A}_1$, it has never made *PrivateKey* and *PatialPrivateKey* queries for ID_{ch} ; if $\mathcal{A} = \mathcal{A}_2$, it has never made *PrivateKey* queries for ID_{ch} (note that \mathcal{A}_2 can calculate partial private key as it knows master key).

Let $\text{Adv}_{\mathcal{A}}(\lambda) = |\Pr[\sigma' = \sigma] - 1/2|$ express \mathcal{A} ’s advantage in the above-defined game.

5.2. Provable Security

Theorem 2. *The DS-CLSPE scheme can achieve IND-CKA secure if the CDH problem is difficult to solve.*

Theorem 2 can be proofed by the two lemmas below.

Lemma 3. *Let the advantage of \mathcal{A}_1 winning the IND-CKA game be ϵ . Then, we can construct an algorithm \mathcal{C} to calculate the CDH problem with advantage*

$$\varepsilon' \geq \frac{\varepsilon}{(q_T + q_{SK} + q_{PSK} + 1)e}, \quad (18)$$

where e is Euler number and q_{SK} , q_{PSK} , q_T denote the maximum number of PrivateKey queries, PartialPrivateKey queries, and Trapdoor queries, respectively.

Proof. Let q be a large prime number. Given (P, aP, bP) , \mathcal{C} 's goal is to output abP .

- (i) *Setup:* \mathcal{C} generates the public parameters $\text{parm} = \{\mathbb{G}, P, P_{\text{pub}}, h_0, h_1, h_2, q\}$, here $P_{\text{pub}} = aP$. \mathcal{C} selects $x_{\text{FS}}, x_{\text{BS}}, t_{\text{FS}}, t_{\text{BS}} \in \mathbb{Z}_q^*$, computes $P_{\text{FS}} = x_{\text{FS}}P$, $P_{\text{BS}} = x_{\text{BS}}P$, $T_{\text{FS}} = t_{\text{FS}}P$, and $T_{\text{BS}} = t_{\text{BS}}P$, and sets $\text{PK}_{\text{FS}} = (T_{\text{FS}}, P_{\text{FS}})$, $\text{PK}_{\text{BS}} = (T_{\text{BS}}, P_{\text{BS}})$. Finally, \mathcal{C} sends $\{\text{parm}, \text{PK}_{\text{FS}}, \text{PK}_{\text{BS}}\}$ to \mathcal{A}_1
- (ii) *Phase 1:* the following oracles can be queried by \mathcal{A}_1
 - (a) h_0 Queries: a h_0 -list with tuples $(\text{ID}_i, T_i, \alpha_i)$ is maintained by \mathcal{C} . When \mathcal{A}_1 performs h_0 query for (ID_i, T_i) , \mathcal{C} first checks whether $(\text{ID}_i, T_i, \alpha_i)$ is already in h_0 -list. If so, \mathcal{C} returns α_i ; otherwise, \mathcal{C} selects $\alpha_i \in \mathbb{Z}_q^*$ at random, adds $(\text{ID}_i, T_i, \alpha_i)$ into h_0 -list, and returns α_i
 - (b) h_1 Queries: \mathcal{C} maintains a h_1 -list with tuples $(\text{ID}_i, T_i, P_i, \beta_i)$. Upon receiving \mathcal{A}_1 's query for (ID_i, T_i, P_i) , \mathcal{C} searches $(\text{ID}_i, T_i, P_i, \beta_i)$ from h_1 -list. If h_1 -list already contains the searched tuple, \mathcal{C} outputs β_i directly; otherwise, \mathcal{C} chooses $\beta_i \in \mathbb{Z}_q^*$ at random, adds $(\text{ID}_i, T_i, P_i, \beta_i)$ into h_1 -list, and returns β_i
 - (c) h_2 Queries: \mathcal{C} maintains a h_2 -list with tuples $(w_i, \text{ID}_i, h_{2i})$. Upon receiving this query for (w_i, ID_i) , \mathcal{C} searches $(w_i, \text{ID}_i, h_{2i})$ from h_2 -list and returns h_{2i} directly if $(w_i, \text{ID}_i, h_{2i})$ already exists in h_2 -list; otherwise, \mathcal{C} selects $h_{2i} \in \mathbb{Z}_q^*$ at random, adds $(w_i, \text{ID}_i, h_{2i})$ into h_2 -list, and returns h_{2i}
 - (d) *CreateUser Queries:* \mathcal{C} maintains a user-list with tuples $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$. Upon receiving \mathcal{A}_1 's query for ID_i , \mathcal{C} searches user-list for $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$. \mathcal{C} outputs PK_i directly if user-list already includes the searched tuple; otherwise, \mathcal{C} tosses a coin $i \in \{0, 1\}$ at random such that $\Pr[\text{coin}_i = 1] = \delta$ (δ will be computed later) and executes the following steps
 - (1) If $\text{coin}_i = 1$, \mathcal{C} selects two random numbers t_i, x_i from \mathbb{Z}_q^* , sets $\text{PK}_i = (T_i, P_i)$, where $T_i = t_iP$, $P_i = x_iP$. \mathcal{C} adds $(\text{ID}_i, \text{PK}_i, \perp, t_i, x_i, 1)$ into user-list and returns PK_i
 - (2) Otherwise ($\text{coin}_i = 0$), \mathcal{C} selects three numbers x_i, d_i, α_i from \mathbb{Z}_q^* , sets $\text{SK}_i = (d_i, x_i)$

and $\text{PK}_i = (T_i, P_i)$, where $T_i = d_iP - \alpha_iP_{\text{pub}}$, $P_i = x_iP$. \mathcal{C} adds $(\text{ID}_i, \text{PK}_i, \text{SK}_i, \perp, \perp, 0)$ into user-list, adds $(\text{ID}_i, T_i, \alpha_i)$ into h_0 -list, and returns PK_i

- (e) *PrivateKey Queries:* when \mathcal{A}_1 submits this query for ID_i , \mathcal{C} searches $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$ from user-list. If $\text{coin}_i = 0$, \mathcal{C} outputs SK_i ; otherwise, \mathcal{C} aborts the game (this event is denoted as Event_1).
 - (f) *PartialPrivateKey Queries:* when \mathcal{A}_1 submits this query for ID_i , \mathcal{C} searches $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$ from user-list. If $\text{coin}_i = 0$, \mathcal{C} outputs the first part of SK_i , i.e., d_i ; otherwise, \mathcal{C} aborts the game (this event is denoted as Event_2).
 - (g) *ReplacePublicKey Queries:* when \mathcal{A}_1 submits this query with a value PK_i^* , \mathcal{C} replaces PK_i with PK_i^* . Note that this query implies \mathcal{A}_1 must also submit the corresponding private key SK_i^*
 - (h) *Trapdoor Queries:* when \mathcal{A}_1 submits this query for (w, ID_i) , \mathcal{C} searches $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$ from user-list and recovers $(\text{ID}_i, T_i, P_i, \beta_i)$ and $(w_i, \text{ID}_i, h_{2i})$ from h_1 -list and h_2 -list, respectively. If $\text{coin}_i = 0$, \mathcal{C} chooses $l \in \mathbb{Z}_q^*$ and computes $T_w^1 = lP$, $T_w^2 = (d_i + \beta_i x_i)h_{2i}Q_{\text{FS}} + lQ_{\text{BS}}$, where Q_{FS} and Q_{BS} are computed as in the proposed scheme. \mathcal{C} returns $T_w = (T_w^1, T_w^2)$. Otherwise, \mathcal{C} aborts the game (this event is denoted as Event_3).
 - (iii) *Challenge:* \mathcal{A}_1 outputs ID_{ch} and (w_0, w_1) . \mathcal{C} recovers $(\text{ID}_{\text{ch}}, \text{PK}_{\text{ch}}, \text{SK}_{\text{ch}}, t_{\text{ch}}, x_{\text{ch}}, \text{coin}_{\text{ch}})$ and $(\text{ID}_{\text{ch}}, T_{\text{ch}}, P_{\text{ch}}, \beta_{\text{ch}})$ from user-list and h_1 -list, respectively. If $\text{coin}_{\text{ch}} = 1$, \mathcal{C} randomly selects $\sigma \in \{0, 1\}$, $C_{w_\sigma}^2 \in \mathbb{G}$, recovers $(w_\sigma, \text{ID}_{\text{ch}}, h_{2\sigma})$ from h_2 -list, sets $C_{w_\sigma}^1 = bP$, and returns $C_{w_\sigma} = (C_{w_\sigma}^1, C_{w_\sigma}^2)$ to \mathcal{A}_1 . Otherwise, \mathcal{C} ends the game (this event is denoted as Event_4).
- Note that $C_{w_\sigma}^2$ is implicitly defined as $bQ_{\text{FS}} + bQ_{\text{BS}} + h_{2\sigma}Q_{\text{ch}} = b(T_{\text{FS}} + \beta_{\text{FS}}P_{\text{FS}} + \alpha_{\text{FS}}P_{\text{pub}}) + b(T_{\text{BS}} + \beta_{\text{BS}}P_{\text{BS}} + \alpha_{\text{BS}}P_{\text{pub}}) + h_{2\sigma}(T_{\text{ch}} + \beta_{\text{ch}}P_{\text{ch}} + \alpha_{\text{ch}}P_{\text{pub}})$.
- (iv) *Phase 2:* The oracles defined in phase1 can be asked continuously by \mathcal{A}_1
 - (v) *Guess:* \mathcal{A}_1 outputs σ' . If $\sigma' = \sigma$, then \mathcal{C} wins in the above game

At this point, \mathcal{C} can compute the value abP as follows:
 $Z = (\alpha_{\text{FS}} + \alpha_{\text{BS}})^{-1}(C_{w_\sigma}^2 - h_{2\sigma}(T_{\text{ch}} + \beta_{\text{ch}}P_{\text{ch}} + \alpha_{\text{ch}}P_{\text{pub}})) - (t_{\text{FS}} + t_{\text{BS}} + \beta_{\text{FS}}x_{\text{FS}} + \beta_{\text{BS}}x_{\text{BS}})bP = abP$.

Analysis. The advantage of \mathcal{C} winning the game is analyzed below. \mathcal{C} wins in above game if none of the events Event_i ($i = 1, 2, 3, 4$) occur.

From the above proof of Lemma 3,

$$\Pr [\overline{\text{Event}}_1 \wedge \overline{\text{Event}}_2 \wedge \overline{\text{Event}}_3 \wedge \overline{\text{Event}}_4] = \delta(1 - \delta)^{q_{\text{SK}} + q_{\text{PSK}} + q_T}, \quad (19)$$

when $\delta = 1/(q_{\text{SK}} + q_{\text{PSK}} + q_T + 1)$, this value is at its maximum value

$$\begin{aligned} & \Pr [\overline{\text{Event}}_1 \wedge \overline{\text{Event}}_2 \wedge \overline{\text{Event}}_3 \wedge \overline{\text{Event}}_4] \\ &= \left(\frac{1}{q_{\text{SK}} + q_{\text{PSK}} + q_T + 1} \right) \left(1 - \frac{1}{q_{\text{SK}} + q_{\text{PSK}} + q_T + 1} \right)^{q_{\text{SK}} + q_{\text{PSK}} + q_T} \\ &\approx \frac{1}{(q_T + q_{\text{SK}} + q_{\text{PSK}} + 1)e}. \end{aligned} \quad (20)$$

Thus,

$$\begin{aligned} \varepsilon' &\geq \varepsilon \cdot \Pr [\overline{\text{Event}}_1 \wedge \overline{\text{Event}}_2 \wedge \overline{\text{Event}}_3 \wedge \overline{\text{Event}}_4] \\ &= \frac{\varepsilon}{(q_T + q_{\text{SK}} + q_{\text{PSK}} + 1)e}. \end{aligned} \quad (21)$$

Lemma 4. *Let ε denote the advantage of \mathcal{A}_2 winning in IND-CKA game. Then, the algorithm \mathcal{C} can be constructed to solve the CDH problem with advantage*

$$\varepsilon' \geq \frac{\varepsilon}{(q_{\text{SK}} + q_T + 1)e}, \quad (22)$$

where e , q_{SK} , and q_T are defined as Lemma 3.

Proof. Given a CDH instance, i.e., (P, aP, bP) , \mathcal{C} will try to output abP .

- (i) *Setup:* \mathcal{C} selects $s \in \mathbb{Z}_q^*$, calculates $P_{\text{pub}} = sP$ and generates $\text{parm} = \{q, \mathbb{G}, P, P_{\text{pub}}, h_0, h_1, h_2\}$. Then, \mathcal{C} selects $x_{\text{FS}}, x_{\text{BS}}, t_{\text{FS}}$, and $t_{\text{BS}} \in \mathbb{Z}_q^*$, sets $P_{\text{FS}} = x_{\text{FS}}P$, $P_{\text{BS}} = x_{\text{BS}}P$, $T_{\text{FS}} = t_{\text{FS}}aP$, and $T_{\text{BS}} = t_{\text{BS}}aP$, and lets $P_{\text{K}_{\text{FS}}} = (T_{\text{FS}}, P_{\text{FS}})$ and $P_{\text{K}_{\text{BS}}} = (T_{\text{BS}}, P_{\text{BS}})$. Finally, \mathcal{C} sends $\{\text{parm}, s, P_{\text{K}_{\text{FS}}}, P_{\text{K}_{\text{BS}}}\}$ to \mathcal{A}_2
- (ii) *Phase1:* the following oracles can be queried by \mathcal{A}_2
 - (a) h_0, h_1, h_2 *Queries:* when \mathcal{A}_2 submits these oracle queries, \mathcal{C} responds as defined in Lemma 3
 - (b) *CreateUser Queries:* \mathcal{C} maintains a user-list with $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$. Upon receiving this query for ID_i , \mathcal{C} searches $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$ from user-list. If this tuple is already in user-list, \mathcal{C} returns PK_i ; otherwise, \mathcal{C} tosses $\text{coin}_i \in \{0, 1\}$ randomly such that $\Pr[\text{coin}_i = 1] = \delta$ (δ will be computed later) and executes the following steps
 - (1) If $\text{coin}_i = 1$, \mathcal{C} selects t_i, x_i from \mathbb{Z}_q^* , computes $T_i = t_i aP$, $P_i = x_i P$, and sets $P_{\text{K}_i} = (T_i, P_i)$. \mathcal{C} adds $(\text{ID}_i, \text{PK}_i, \perp, \perp, x_i, 1)$ into user-list and returns PK_i

TABLE 1: Notions for some basic operations.

Symbols	Definition
T_{sm}	Runtime for a scalar multiplication in \mathbb{G}
T_{mm}	Runtime for a modular multiplication in \mathbb{Z}_q^*
T_{bp}	Runtime for a bilinear pairing
T_{pa}	Runtime for a point addition in \mathbb{G}
T_H	Runtime for a map-to-point hash
$ \mathbb{G} $	The bit-size of a point in \mathbb{G}
$ \mathbb{Z}_q $	The bit-size of a number in \mathbb{Z}_q
$ \text{PK} $	The bit-size of PK
$ \text{C} $	The bit-size of ciphertext
$ T $	The bit-size of trapdoor

- (2) Otherwise, \mathcal{C} selects x_i, t_i, α_i from \mathbb{Z}_q^* at random, computes $d_i = t_i + s\alpha_i \bmod q$, $T_i = t_i P$, $P_i = x_i P$, and sets $\text{SK}_i = (d_i, x_i)$, $\text{PK}_i = (T_i, P_i)$. \mathcal{C} adds $(\text{ID}_i, \text{PK}_i, \text{SK}_i, \perp, \perp, 0)$ into user-list, adds $(\text{ID}_i, T_i, \alpha_i)$ into h_0 -list, and returns PK_i

- (c) *PrivateKey Queries:* Upon receiving this query for ID_i , \mathcal{C} searches user-list to find $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$. If $\text{coin}_i = 0$, \mathcal{C} outputs SK_i ; otherwise, \mathcal{C} aborts the game (this event is denoted as Event_1)

- (d) *Trapdoor Queries:* when \mathcal{A}_2 submits the query for (w, ID_i) , \mathcal{C} searches $(\text{ID}_i, \text{PK}_i, \text{SK}_i, t_i, x_i, \text{coin}_i)$ from user-list and recovers $(\text{ID}_i, T_i, P_i, \beta_i)$ and $(w_i, \text{ID}_i, h_{2i})$ from h_1 -list and h_2 -list, respectively. If $\text{coin}_i = 0$, \mathcal{C} chooses $l \in \mathbb{Z}_q^*$, computes $T_w^1 = lP$, $T_w^2 = (d_i + \beta_i x_i)h_{2i}Q_{\text{FS}} + lQ_{\text{BS}}$, and returns $T_w = (T_w^1, T_w^2)$ to \mathcal{A}_2 , where Q_{FS} and Q_{BS} are computed as in the proposed scheme; otherwise, \mathcal{C} aborts the game (this event is denoted as Event_2)

- (iii) *Challenge:* \mathcal{A}_2 outputs ID_{ch} and (w_0, w_1) . \mathcal{C} recovers $(\text{ID}_{\text{ch}}, \text{PK}_{\text{ch}}, \text{SK}_{\text{ch}}, t_{\text{ch}}, x_{\text{ch}}, \text{coin}_{\text{ch}})$ from user-list. If $\text{coin}_{\text{ch}} = 1$, \mathcal{C} randomly selects $\sigma \in \{0, 1\}$, $C_{w_\sigma}^2 \in \mathbb{G}$ and recovers $(w_\sigma, \text{ID}_{\text{ch}}, h_{2\sigma})$ from h_2 -list. \mathcal{C} sets $C_{w_\sigma}^1 = bP$ and returns $C_{w_\sigma} = (C_{w_\sigma}^1, C_{w_\sigma}^2)$ to \mathcal{A}_2 . Otherwise, \mathcal{C} aborts the game (this event is denoted as Event_3)

Note that $C_{w_\sigma}^2$ is implicitly defined as $bQ_{\text{FS}} + bQ_{\text{BS}} + h_{2\sigma}Q_{\text{ch}} = b(T_{\text{FS}} + \beta_{\text{FS}}P_{\text{FS}} + \alpha_{\text{FS}}P_{\text{pub}}) + b(T_{\text{BS}} + \beta_{\text{BS}}P_{\text{BS}} + \alpha_{\text{BS}}P_{\text{pub}}) + h_{2\sigma}(T_{\text{ch}} + \beta_{\text{ch}}P_{\text{ch}} + \alpha_{\text{ch}}P_{\text{pub}})$.

- (iv) *Phase2:* the oracles defined in phase 1 can be continuously accessed by \mathcal{A}_2
- (v) *Guess:* \mathcal{A}_2 returns σ' . If $\sigma' = \sigma$, then \mathcal{C} wins in the game

TABLE 2: The performance comparison.

Schemes	Computation cost (ms)				Communication cost (bit)		
	KeyGen	Encrypt	Trapdoor	Test	$ \mathbb{G} $	$ \mathbb{C} $	$ T $
Lu and Li [21]	$2T_{sm} + T_{mm}$	$3T_{sm} + T_{mm}$	T_{mm}	T_{sm}	$2 \mathbb{G} $	$ \mathbb{G} + \mathbb{Z}_q $	$ \mathbb{Z}_q $
Peng et al. [18]	$8T_{sm} + 2T_H$	$4T_{sm} + 3T_H + 3T_{bp}$	$3T_{sm} + T_H$	$T_{sm} + 2T_{pa} + T_{bp}$	$4 \mathbb{G} $	$ \mathbb{G} + \mathbb{Z}_q $	$3 \mathbb{G} $
DS-CLSPE	$6T_{sm}$	$10T_{sm} + 8T_{pa}$	$3T_{sm} + T_{pa}$	$4T_{sm} + 3T_{pa}$	$6 \mathbb{G} $	$2 \mathbb{G} $	$2 \mathbb{G} $

At this point, \mathcal{C} can compute abP as below: $Z = (t_{FS} + t_{BS})^{-1} \cdot (C_{w_r}^2 - h_{2\sigma}(T_{ch} + \beta_{ch}P_{ch} + \alpha_{ch}P_{pub}) - (\beta_{FS}x_{FS} + \beta_{BS}x_{BS} + \alpha_{FS}s + \alpha_{BS}s)bP) = abP$.

Analysis. Now let us analyze \mathcal{C} 's advantage in winning the above game. \mathcal{C} will win the game if Event_1 , Event_2 , and Event_3 do not occur.

From the above proof of Lemma 4,

$$\Pr [\overline{\text{Event}_1} \wedge \overline{\text{Event}_2} \wedge \overline{\text{Event}_3}] = \delta \cdot (1 - \delta)^{q_{sk} + q_T}, \quad (23)$$

when $\delta = 1/(q_{sk} + q_T + 1)$, $\Pr [\overline{\text{Event}_1} \wedge \overline{\text{Event}_2} \wedge \overline{\text{Event}_3}]$ takes its maximum value

$$\begin{aligned} & \Pr [\overline{\text{Event}_1} \wedge \overline{\text{Event}_2} \wedge \overline{\text{Event}_3}] \\ &= \left(\frac{1}{q_{sk} + q_T + 1} \right) \left(1 - \frac{1}{q_{sk} + q_T + 1} \right)^{q_{sk} + q_T} \\ &\approx \frac{1}{(q_{sk} + q_T + 1)e}. \end{aligned} \quad (24)$$

Then, we have

$$\epsilon' \geq \epsilon \cdot \Pr [\overline{\text{Event}_1} \wedge \overline{\text{Event}_2} \wedge \overline{\text{Event}_3}] = \frac{\epsilon}{(q_{sk} + q_T + 1)e}. \quad (25)$$

6. Performance Analysis

This section mainly compares the computation/communication costs of DS-CLSPE with that of Lu and Li [21] and Peng et al. [18]. Let p, q be 512-bit and 160-bit prime numbers, respectively. \mathbb{G} is a cyclic group with order q , which is generated by a point on a super-singular elliptic curve $E(F_p)$. For the convenience of comparison, Table 1 presents the definition of some symbols.

We evaluate the running time of the above basic operations using the MIRACL library [34] and performing on a personal computer (Processor: i5-8250U 1.60 GHz; Memory: 8GB; Operating system: Win10). The evaluation result shows that $T_{sm} = 4.800$ ms, $T_{mm} = 0.003$ ms, $T_{bp} = 13.144$ ms, $T_H = 12.082$ ms, $T_{pa} = 0.025$ ms. Furthermore, the result indicates that the operations of T_H and T_{bp} consume much more time than other operations. Therefore, we should minimize or even avoid using these time-consuming operations to enhance the efficiency of the designed scheme.

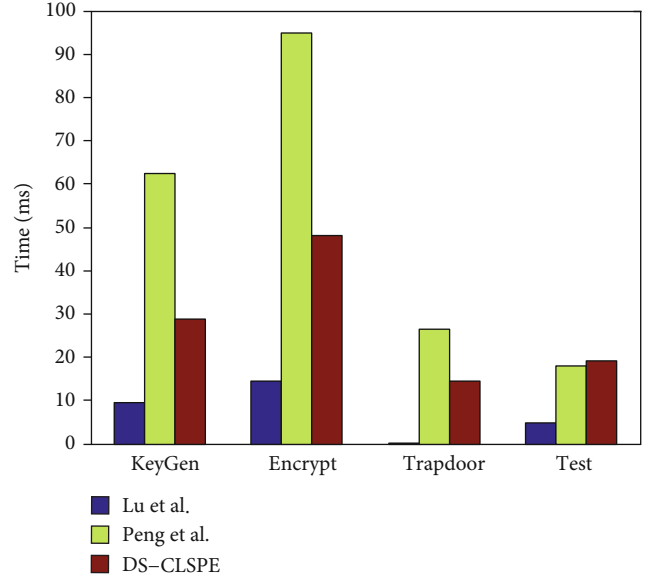


FIGURE 3: Computation cost comparison.

To compare the computation costs, we analyze the proposed DS-CLSPE scheme and schemes [18, 21] in terms of four phases: *KeyGen*, *Encrypt*, *Trapdoor*, and *Test*. Table 2 and Figure 3 present the specific comparison results. In addition, the communication costs of DS-CLSPE and schemes [18, 21] are also presented in Table 2.

From Table 2 and Figure 3, the efficiency of DS-CLSPE is slightly worse than the scheme [21], but DS-CLSPE avoids the security flaws that existed in the scheme [21]. The data security is a primary concern in practical application, so DS-CLSPE is more practical. And in comparison with scheme [18], DS-CLSPE has better performance.

7. Conclusion

As the maturity of IoT and the popularization of sensor devices, IIoT has attracted widespread attention, which can provide users with real-time and reliable intelligent services by collecting and analyzing massive industrial data via the IoT devices. However, some sensitive information may be involved in industrial data, so data security is concerned. To protect data privacy, Lu et al. designed a CLSPE scheme without bilinear pairing operation. Unfortunately, we analyze that their CLSPE scheme cannot prevent user impersonation attacks. To resolve the security flaws, we design an improved pairing-free dual-server CLSPE scheme, i.e., DS-CLSPE. The formal security proof shows that DS-CLSPE

can realize IND-CKA security. Additionally, we evaluate the efficiency of DS-CLSPE, and evaluation results indicate the proposed scheme has better efficiency.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported by the National Natural Science Foundation of China (Nos. 61902111, 61932016, and 61972294), the High-level talent Fund Project of Henan University of Technology (No. 2018BS052), the Project funded by China Postdoctoral Science Foundation (No. 2020M670223), and the National Key Research and Development Program of China (No. 2018YFC1604000).

References

- [1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): a vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.
- [2] S. Pattar, R. Buyya, K. R. Venugopal, S. Iyengar, and L. Patnaik, "Searching for the iot resources: fundamentals, requirements, comprehensive review, and future directions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2101–2132, 2018.
- [3] S. Li, L. Da Xu, and S. Zhao, "5G internet of things: a survey," *Journal of Industrial Information Integration*, vol. 10, pp. 1–9, 2018.
- [4] M. Younan, E. H. Houssein, M. Elhoseny, and A. A. Ali, "Challenges and recommended technologies for the industrial internet of things: a comprehensive review," *Measurement*, vol. 151, p. 107198, 2020.
- [5] MarketsandMarkets, "Industrial iot market by device & technology (sensor, rfid, industrial robotics, dcs, condition monitoring, smart meter, camera system, networking technology), software (plm, mes, scada), vertical, and geography-global forecast to 2023," <https://www.marketsandmarkets.com/Market-Reports/industrial-internet-of-things-market-129733727.html>.
- [6] M. S. Hossain and G. Muhammad, "Cloud-assisted industrial internet of things (IIoT) – enabled framework for health monitoring," *Computer Networks*, vol. 101, pp. 192–202, 2016.
- [7] K. Liang and W. Susilo, "Searchable attribute-based mechanism with efficient data sharing for secure cloud storage," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1981–1992, 2015.
- [8] C. Esposito, A. De Santis, G. Tortora, H. Chang, and K.-K. R. Choo, "Blockchain: a panacea for healthcare cloud-based data security and privacy?," *IEEE Cloud Computing*, vol. 5, no. 1, pp. 31–37, 2018.
- [9] J. Srinivas, A. K. Das, N. Kumar, and J. Rodrigues, "Cloud centric authentication for wearable healthcare monitoring system," *IEEE Transactions on Dependable and Secure Computing*, vol. 17, no. 5, pp. 942–956, 2018.
- [10] D. He, N. Kumar, S. Zeadally, and H. Wang, "Certificateless provable data possession scheme for cloud-based smart grid data management systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1232–1241, 2018.
- [11] A. K. Das, M. Wazid, N. Kumar, A. V. Vasilakos, and J. J. Rodrigues, "Biometrics-based privacy-preserving user authentication scheme for cloud-based industrial internet of things deployment," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4900–4913, 2018.
- [12] X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proceeding 2000 IEEE Symposium on Security and Privacy. S&P 2000*, pp. 44–55, Berkeley, CA, USA, May 2000.
- [13] D. Boneh, G. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Advances in Cryptology - EUROCRYPT 2004. EUROCRYPT 2004. Lecture Notes in Computer Science, vol 3027*, C. Cachin and J. L. Camenisch, Eds., pp. 506–522, Springer, Berlin, Heidelberg, 2004.
- [14] H. S. Rhee, J. H. Park, W. Susilo, and D. H. Lee, "Trapdoor security in a searchable public-key encryption scheme with a designated tester," *Journal of Systems and Software*, vol. 83, no. 5, pp. 763–771, 2010.
- [15] Q. Huang and H. Li, "An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks," *Information Sciences*, vol. 403–404, pp. 1–14, 2017.
- [16] M. Zeng, H.-F. Qian, J. Chen, and K. Zhang, "Forward secure public key encryption with keyword search for outsourced cloud storage," *IEEE Transactions on Cloud Computing*, 2019.
- [17] M. Abdalla, M. Bellare, D. Catalano et al., "Searchable encryption revisited: consistency properties, relation to anonymous ipe, and extensions," in *Advances in Cryptology – CRYPTO 2005. CRYPTO 2005. Lecture Notes in Computer Science, vol 3621*, V. Shoup, Ed., pp. 205–222, Springer, Berlin, Heidelberg, 2005.
- [18] Y. Peng, J. Cui, C. Peng, and Z. Ying, "Certificateless public key encryption with keyword search," *China Communications*, vol. 11, no. 11, pp. 100–113, 2014.
- [19] S. H. Islam, M. S. Obaidat, V. Rajeev, and R. Amin, "Design of a certificateless designated server based searchable public key encryption scheme," in *Mathematics and Computing. ICMC 2017. Communications in Computer and Information Science, vol 655*, D. Giri, R. Mohapatra, H. Begehr, and M. Obaidat, Eds., pp. 3–15, Springer, Singapore, 2017.
- [20] D. He, M. Ma, S. Zeadally, N. Kumar, and K. Liang, "Certificateless public key authenticated encryption with keyword search for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3618–3627, 2017.
- [21] Y. Lu and J. Li, "Constructing pairing-free certificateless public key encryption with keyword search," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 8, pp. 1049–1060, 2019.
- [22] Q. Chai and G. Gong, "Verifiable symmetric searchable encryption for semi-honest-but-curious cloud servers," in *2012 IEEE International Conference on Communications (ICC)*, pp. 917–922, Ottawa, ON, Canada, June 2012.
- [23] C. Zuo, S.-F. Sun, J. K. Liu, J. Shao, and J. Pieprzyk, "Dynamic searchable symmetric encryption schemes supporting range queries with forward (and backward) security," in *Computer Security. ESORICS 2018. Lecture Notes in Computer Science*,

- vol 11099*, J. Lopez, J. Zhou, and M. Soriano, Eds., pp. 228–246, Springer, Cham, 2018.
- [24] Z. Zhang, J. Wang, Y. Wang, Y. Su, and X. Chen, “Towards efficient verifiable forward secure searchable symmetric encryption,” in *Computer Security – ESORICS 2019. ESORICS 2019. Lecture Notes in Computer Science, vol 11736*, K. Sako, S. Schneider, and P. Ryan, Eds., pp. 304–321, Springer, Cham, 2019.
- [25] C. Lin, D. He, X. Huang, X. Xie, and K.-K. R. Choo, “Blockchain-based system for secure outsourcing of bilinear pairings,” *Information Sciences*, vol. 527, pp. 590–601, 2020.
- [26] Y. Zhang, J. Katz, and C. Papamanthou, “All your queries are belong to us: the power of file-injection attacks on searchable encryption,” in *Proceedings of the 25th Security Symposium, USENIX*, pp. 707–720, Austin, TX, USA, August 2016.
- [27] A. Shamir, “Identity-based cryptosystems and signature schemes,” in *Advances in Cryptology. CRYPTO 1984. Lecture Notes in Computer Science, vol 196*, G. R. Blakley and D. Chaum, Eds., pp. 47–53, Springer, Berlin, Heidelberg, 1984.
- [28] Y. Lu, G. Wang, J. Li, and J. Shen, “Efficient designated server identity-based encryption with conjunctive keyword search,” *Annals of Telecommunications*, vol. 72, no. 5-6, pp. 359–370, 2017.
- [29] H. Li, Q. Huang, J. Shen, G. Yang, and W. Susilo, “Designated-server identity-based authenticated encryption with keyword search for encrypted emails,” *Information Sciences*, vol. 481, pp. 330–343, 2019.
- [30] X. Zhang, Y. Tang, H. Wang, C. Xu, Y. Miao, and H. Cheng, “Lattice-based proxy-oriented identity-based encryption with keyword search for cloud storage,” *Information Sciences*, vol. 494, pp. 193–207, 2019.
- [31] S. S. Al-Riyami and K. G. Paterson, “Certificateless public key cryptography,” in *Advances in Cryptology - ASIACRYPT 2003. ASIACRYPT 2003. Lecture Notes in Computer Science, vol 2894*, C. S. Lai, Ed., pp. 452–473, Springer, Berlin, Heidelberg, 2003.
- [32] B. C. Hu, D. S. Wong, Z. Zhang, and X. Deng, “Certificateless signature: a new security model and an improved generic construction,” *Designs, Codes and Cryptography*, vol. 42, no. 2, pp. 109–126, 2007.
- [33] R. Chen, Y. Mu, G. Yang, F. Guo, and X. Wang, “Dual-server public-key encryption with keyword search for secure cloud storage,” *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 4, pp. 789–798, 2015.
- [34] “Shamus software ltd., miracl library,” <http://www.shamus.ie/index.php?page=home>.

Research Article

Privacy-Preserving Vehicle Assignment in the Parking Space Sharing System

Tingting Fu,¹ Peng Liu ,^{1,2} Kun Liu,¹ and Peng Li³

¹School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, China

²Guangdong Provincial Key Laboratory of Information Security Technology, Guangzhou, China

³School of Computer Science and Engineering, University of Aizu, Japan

Correspondence should be addressed to Peng Liu; perrypliu@hotmail.com

Received 17 May 2020; Revised 25 July 2020; Accepted 28 September 2020; Published 17 October 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Tingting Fu 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.

Nowadays, the availability of parking spaces is far behind the quick rising number of cars. Rather than building more lots, a better way is to share private-owned parking spaces. However, this faces the challenge that users are not willing to expose their privacy to the public. To solve this problem, we propose a new architecture for parking space sharing, integrating homomorphic cryptography into the design of a secure protocol for parking space searching and booking. The proposed privacy-preserving matching scheme (PPMS) is constructed in an untrusted third-party service system including two independent entities, namely, a server and an intermediary platform. Via the participant comparison protocol (PCP), a driver can choose from the matching result and be navigated to the parking space near his destination, without knowing any information of the provider and vice versa. In the meanwhile, in order to further improve the efficiency of matching, we also propose a block algorithm based on the longitude and latitude (BABLL), which utilizes a novel partitioning scheme. The feasibility of the architecture is validated through the detailed theoretical analysis and extensive performance evaluations, including the assessment of the resilience to attacks.

1. Introduction

The number of private cars in cities continues to increase against limited parking resources year by year, which leads to the desperate parking phenomenon and brings huge inconvenience to the transportation system. According to our previous work [1], finding a parking space in a city has become the most challenging task. It is estimated that 30% of traffic jams in urban areas every day are caused by vehicles looking for free parking lots. It is one of the major cause of traffic congestion, air pollution, and social anxiety. As shown in Figure 1, it is difficult for a user to find a parking space near the destination, either because a parking lot is full or the space is reserved. In the meantime, a lot of private parking spaces are unoccupied during daytime as owners may have left for work. The obvious spatial and temporal complementarity between the owners and visitors, e.g., rational renting of parking space, could increase the utilization of private parking resources as well as profits of providers. Therefore, to build a parking space sharing system [2], which is able to

match private parking space providers with requestors, is a promising way to solve the problem of parking difficulty.

Currently, the smart vehicle parking system (SVPS) can alert drivers via mobile phones if a parking space is available in a particular area [3], through 5G and edge computing [4]. However, this type of system may not work in an extremely crowded area. In addition, if a parking space is not allocated specifically for the vehicle, it may no longer be available by the time the vehicle arrives. Assigning parking spaces in advance can reduce the time and gasoline spent by drivers looking for parking vacancies, as well as ensure that they meet specified requirements of drivers.

One major challenge of implementing the parking space sharing system is that users are very concerned about privacy disclosure [5] (e.g., location and time). For example, if a malicious user used your parking space, he would know that your parking space is empty for a certain period of time. Next time, he may take over your parking space without notifying you. In the meantime, the sharing service platform may sell your personal information. Therefore, it is necessary to



FIGURE 1: Finding a parking space in big cities.

design an efficient spatial-temporal matching algorithm for the shared parking space system, so that users can be assigned to the nearest place while keeping their privacy protected. Many matching algorithms have been proposed in the sharing economy, such as bike-sharing business. Although these algorithms may be very efficient, the information of users is open to the service provider. As shown in Figure 2, if following the traditional system architecture, both parking space providers and parking space requestors have to upload their information to the server, then wait for the matching results. Therefore, to protect privacy of all participants, a completely trusted platform is required. However, for a variety of reasons, such as sudden external threats, internal adversaries, and insecure platforms, a fully trusted third-party platform is difficult to achieve.

In order to solve the privacy protection problem of the parking space sharing system, we propose to introduce homomorphic cryptography, so that the matching process is calculated over cryptographic data. In the system, we design a privacy-protection matching scheme (PPMS) to meet needs of privacy protection. Specifically, the solution is based on a system architecture with two separate entities: (i) a server and (ii) an intermediate platform, which come from different service providers. Under the architecture we designed, the server is only able to process the encrypted information, while the intermediate platform can only access information disguised with a participant comparison proto-

col [6]. This design ensures no sensitive information will be disclosed to other parties. To implement matching under encryption, we use the additional feature of the Paillier cryptosystem and design an algorithm of parking space matching based on geographical location by longitude and latitude.

To evaluate the performance of the proposed PPMS scheme, we first theoretically analyze the validity and security of PPMS, which measures efficiency by evaluating time and space complexity, and then apply various attack behaviors to verify security resiliency. Finally, we practically implement the system and conduct extensive experiments on various scenarios. Experimental results show that PPMS can work efficiently while preventing information leakage.

Our contributions in this paper are as follows:

- (i) We are the first to focus on the privacy-preserving problem in the parking space sharing system and try to allocate appropriate parking spaces under anonymous situations
- (ii) We propose a scheme called privacy-protection matching scheme (PPMS) to realize the matching with encrypted information, which is based on a new block algorithm (BABLL) utilizing longitude and latitude
- (iii) We evaluate the effectiveness, efficiency, and security of the proposed PPMS scheme by both theoretical

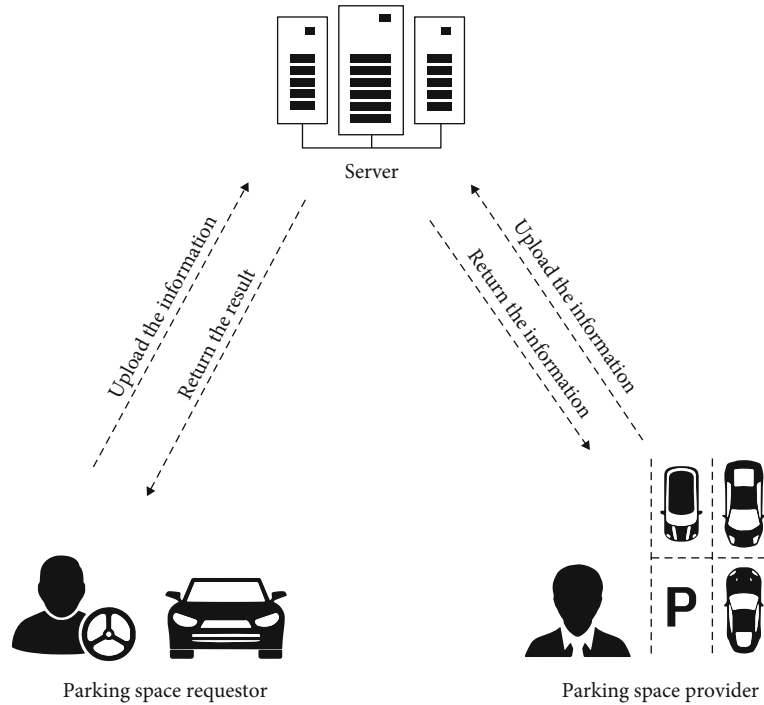


FIGURE 2: Traditional system architecture.

analysis and practical implementation. The results show that our scheme totally meets the needs of parking space sharing systems

The rest of this paper is organized as follows: in Section 2, literature review is conducted. We provide a brief review of the preliminary background of homomorphic cryptography in Section 3. In Section 4, we present our benchmark solution, the privacy protection matching scheme (PPMS) for shared parking systems, as well as analysis, performance evaluation, and considerations. The security analysis together with the performance evaluation are discussed in Section 5. We draw conclusion of this work in Section 6.

2. Related Work

Quite a few studies have focused on solving parking problem in big cities, for example, in [7], Xu et al. propose an integrated auction and market design method for the parking space sharing and allocation problem. There is an IoT-based platform which implements price-compatible top trading cycles and chains (PC-TTCCs) mechanism among agents. The platform's parking spaces are reassigned via a one-sided Vickrey-Clarke-Groves (O-VCG) auction. Another work taking the price into account is proposed in [8]. Zhao et al. introduce a management framework of shared parking resource in terms of time and spatial dimension [9]. Their emphasis is on the nonpreemptive utilization of parking resources. The uncertainties of P-users' and O-users' arrival and departure are simulated in the intelligent parking management system (IPMS). A practical case of parking space sharing in Suzhou is studied, and some suggestions are given by the author [10]. Some work focuses on the design of the

infrastructure or architecture, e.g., in [11], it develops time shared parking management system based on BaaS cloud infrastructure in JAVA. The shared allocation model of residential area and adjacent business district in Nantong is studied in [12]. In [13], fog computing and roadside cloud are utilized to find a vacant spot, then the matching theory is applied to solve the parking problem. In [14], the research of a shared parking system can be divided into three aspects: system structure analysis, system prediction, and privacy protection. A campus parking sharing method is developed in [15], and daytime and nighttime users can share a common campus parking lot. To ensure the QoS, the number of reserved parking spaces and a probability-based access method for premium cars are studied. However, in the above existing work, more or less, one important factor is neglected, i.e., the privacy of participants.

In the related area, many researches have been done in privacy-related problems e.g., in recommendation domain, Qi et al. consider the spatial-temporal information of the QoS data and locality-sensitive hashing (LSH), and propose a location-aware and time-aware recommendation approach considering privacy concerns [16]. Xiao protects users' privacy when they perceive data through secret sharing and multiparty security calculation to prevent leakage [17]. L-PPTD [18], a privacy-preserving truth discovery framework for mobile crowd sensing systems with two noncolluding cloud platforms, is proposed. By adopting the additively homomorphic cryptosystem, it achieves the protection of each worker's sensory data and reliability information. Tang et al. also propose a framework following the same two-server model from the best-known prior solution [19]. They leverage Yao's Garbled Circuit (GC) for data comparison. In [20], the authors present two homomorphic Paillier

encryption-based reliable and privacy-preserving truth discovery schemes, for stable and frequently moving users, respectively. Li proposes a privacy-preserving reinforcement learning framework for a patient-centric dynamic treatment regime [21]. LoPrO [22] is a location privacy-preserving online algorithm which can differentially guarantee the EV location information privacy by leveraging the Laplace mechanism. Much of above work is built on protecting privacy by blending it with the additional data. To reduce users' tuning time, the privacy-preserving spatial index (PSI) method is proposed in [23]. For Internet of Vehicle (IoV), there remains a challenge to avoid privacy conflicts for computation offloading. An edge computing-enabled computation offloading method, named ECO, with privacy preservation for IoV, is proposed in [24], to address this challenge.

Many encryption schemes have been developed to support privacy-preserving. For example, ref. [25] proves the effectiveness of homomorphic encryption. In [6], Zhang proposes a comparison method in the case of encryption. Furthermore, searchable encryption can be very useful in cloud computing [26]. Besides addition and comparison, many other operations can be conducted in the case of encryption, such as in [27], and the authors propose a practical and privacy-preserving data aggregation scheme that can compute arbitrary aggregation functions without a TA. On one hand, the scheme can ensure users' anonymity and privacy protection, while on the other, the scheme is efficient in enabling participants to join or leave the system dynamically.

A blockchain-based privacy-preserving system is briefly introduced in [28]. However, only the user ID is anonymized while other information is open to everyone. Otherwise, it is not possible for users to find their trading partners. In this paper, we are going to design and implement such a system with state-of-the-art encryption technologies.

3. Preliminary

In the proposed scenario, sensitive information must be encrypted during storing, transmission, and processing. Especially for the processing, all operation have to be done without revealing any of original contents. Sorting is one of the most important functions to effectively implement the matching process, because it determines the winning position. The cryptosystem needs to be function-effective and time-efficient. According to the above analysis, we chose the Paillier cryptosystem, which supports homomorphic encryption, to implement our design. In this section, we briefly introduce the security and attack model, concept of homomorphic cryptography, and the participant comparison protocol (PCP).

3.1. Security and Attack Model. In this design, we would like to provide strong privacy protection for all participants in the parking space sharing system. Therefore, regarding the security model, we assume there is no trustworthy entity except the government approved agency, such as CA. Malicious attackers hide in the server, the intermediate platform, requestors, and providers. They are trying to act like a sham,

to steal user private information, to cheat in the transaction, and to destroy services. To resist these adversarial actions, this paper considers the following attack model:

3.1.1. DDoS Attack. In this situation, the malicious user tries to expend system resources but not to conduct any actual transactions, e.g., proposing a lot of parking requests but finally cancel them.

3.1.2. Multiple Reservation Attack. In this case, multiple malicious users try to book the same parking space so as to disrupt normal functionality.

3.1.3. Adversarial Behavior Attack. Some users may damage the parking space on purpose or possess the space violating the agreed time period.

3.1.4. Statistical Attack. A common attack from malicious users is statistical attack, i.e., using a lot of known plaintext sent to the server or the platform to guess the corresponding ciphertext.

3.1.5. Personator Attack. Malicious users may intercept and capture the normal user's encrypted ID and use it to cheat in the transaction.

3.1.6. Theft of Private Information Attack. Malicious users may use all means to reveal the private information of participants, such as ID, location, time preference, distance, and other sensitive data.

3.2. Homomorphic Encryption. Traditionally, no matter data is encrypted during transmission or storing, and to get meaningful results, it must be decrypted before processing. As more and more tasks are being offloaded to a server or a cloud, it is unavoidable to expose the plaintext of the corresponding data. Instead, the ability to perform calculation on encrypted ciphertext will ensure that the processor can never reveal the content of the original data. Homomorphic cryptography is the encryption algorithm that can achieve the goal, where the operation performed on the ciphertext, once decrypted, matches the corresponding operation performed on the original plaintext. The concept of homomorphic encryption can be presented as Eq. (1),

$$\forall m_1, m_2 \in M, E_K(m_1) \odot_c E_K(m_2) \rightarrow E_K(m_1 \odot_M m_2), \quad (1)$$

where M is the set of plaintext, m is any plaintext in M , E is the encryption function, K is the encryption key, and \odot_M and \odot_c stand for some operators in M and the corresponding set of ciphertext C . Regarding the ideal homomorphic encryption scheme, untrusted servers can only view, operate on, and return encrypted data, thus protecting users' privacy.

With regard to the development of homomorphic encryption, the first homomorphic encryption scheme implements only partial homomorphic encryption (PHE), which means that possible homomorphic operations are limited to only one operation such as addition or multiplication. In 2009, the first fully homomorphic scheme (FHE) that would allow combined computation of the encrypted data without decryption was proposed by Gentry [29]. The

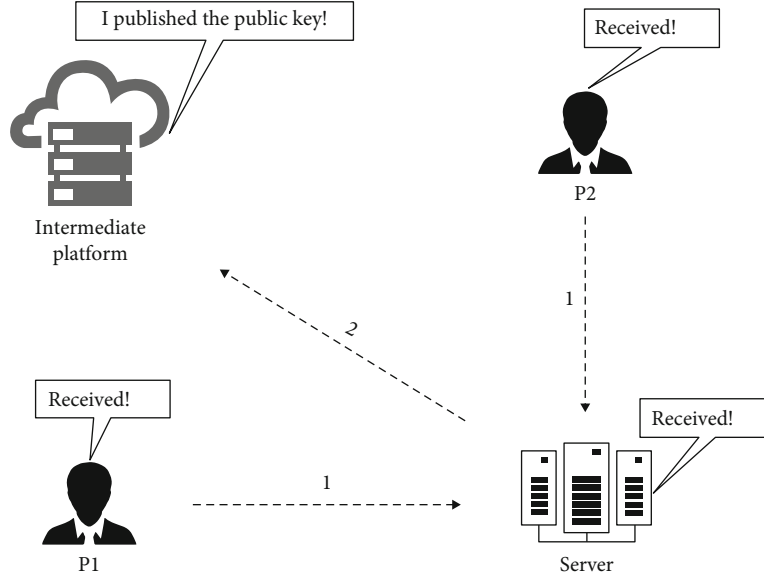


FIGURE 3: Participant comparison protocol.

scheme uses grid-based cryptography and guidance mechanism. Extensions to this foundation include LTV and GSW. Despite these advances, in most cases, current FHE implementations are not suited to the efficiency requirements of speed and space. As a compromise, recent trend is to use feature-reduced homomorphic encryption that has a limited number and depth of FHE operations but still yields good performance.

In the proposed parking space sharing system, we adopt and implement a Paillier cryptosystem (PC), which is an asymmetric encryption system based on the problem of calculating the n^{th} residue and is considered computationally difficult. PC is a partial homomorphic scheme that supports ciphertext addition. That is, $E((m_1) + (m_2)) = E(m_1) \cdot E(m_2)$, where E is the encryption function, and m_1 and m_2 are plaintexts. Since PC can achieve semantic security, it is impossible to determine the original plaintext in polynomial time. That is, all of data is encrypted using PC before being transmitted to an untrusted server for processing. The server will perform required operations on encrypted data and also return the encrypted outcome. Users decrypt the outcome to get the final result. Ideally, privacy is protected because the plaintext of data is not exposed to any untrusted entities.

3.3. Participant Comparison Protocol. To assign the nearest parking space to the requesting driver, we need to compare the distance of candidate places. However, distance is also part of privacy needed to be protected. Therefore, the comparison must be done on ciphertexts. Participant comparison protocol allows us to perform size comparisons without decryption. It is used to handle the returned results.

Figure 3 shows the process how participant comparison protocol works on our system. We divide it into 2 main phases as corresponding to the numbers in Figure 3.

Phase 1: the intermediate platform generates Paillier's key pair and publishes the public key. The private key is used by the platform to decrypt the ciphertext and extract the

information so that it must not be leaked. The platform keeps the private key with itself at all times to ensure its safety. Participants P1 and P2 use the public key to encrypt their values and send them to the server.

Phase 2: the server uses its key pairs to interfere with the received information. The detail is shown in Eq. (2) (where Inf represents the values that the participants need to compare, and x and y represent the key pairs generated by the server). The server sends the interference information to the intermediate platform. The intermediate platform uses its own private key to decrypt the interference information, as shown in Eq. (3) and then directly compares the size.

$$PK(Inf \cdot x + y) = PK(Inf)^x \cdot PK(y), \quad (2)$$

$$Inf \cdot x + y = SK(PK(Inf)^x \cdot PK(y)). \quad (3)$$

In Eq. (2) and Eq. (3), $PK()$ stands for the encryption function using the public key of the intermediate platform, and $SK()$ stands for the decryption function using the secret key of the intermediate platform. x and y are two large integers generated by the server. Inf is the information sent by users. In this design, the server can only see the encrypted information, and the intermediate platform sees only the padded information. Furthermore, neither of them would be able to uncover the original information during the calculation. In the meantime, it is possible for the intermediate platform to use the padded information to compare and sort when a proper homomorphic encryption system is selected.

4. Privacy-Preserving Matching Scheme

In this section, we will illustrate the proposed privacy-preserving matching scheme by first introducing the architecture and then the workflow, finally the detailed algorithm.

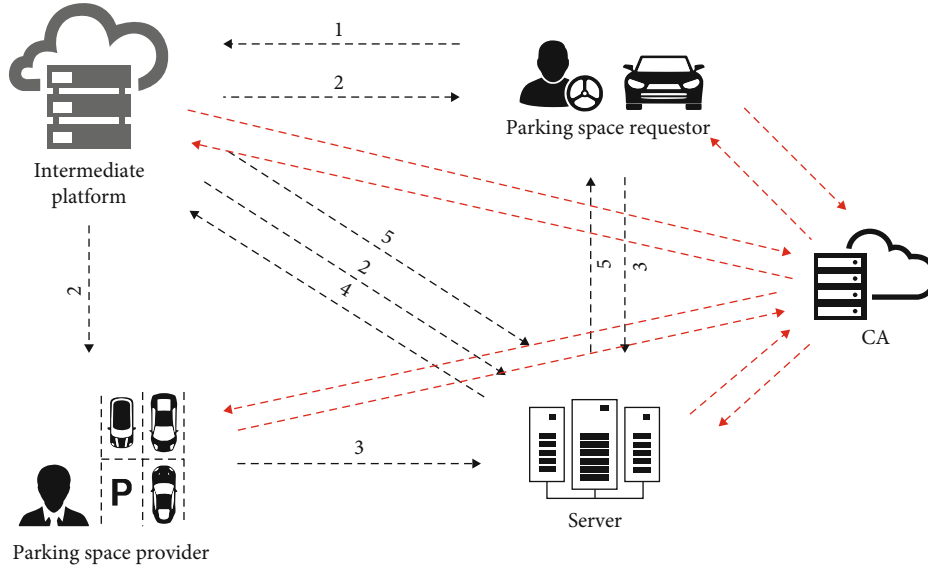


FIGURE 4: System architecture and workflow.

4.1. System Architecture. Figure 4 illustrates the system architecture of the privacy-preserving matching scheme, which is composed of five entities: the requestor, the intermediate platform, the space provider, the server, and the CA. These entities use modern networking technologies to communicate with each other, such as wireless area network, wide area network, cellular network, and vehicular ad-hoc network [30]. Each entity is explained in detail below.

4.1.1. Requestor. The requestor is the user looking for a parking space. He needs the logic location and access code of the desired parking space near his destination. He should have no access to the privacy information such as the provider's identity, available hours of the parking space, and the physical location (in the time of matching and selecting). Regarding the physical location, it is allowable to the requestor when he is guided to and really uses the parking space. The major actions which the requestor must take are to register in the system, provide asymmetric key pairs, and upload the encrypted requirement (by the public key of the intermediate platform) to the server.

4.1.2. Provider. The provider is willing to share his parking space to gain some profit. The information he needs to provide including location, available time, price, and payee information. He must register in the system and send the said information to the server in the form of ciphertext. Later, the requestor is allowed to update above information following the specified procedure.

4.1.3. Server. The server is a party providing privacy preserving in the matching system. It initializes the new matching process based on the requestor's request and generates a large integer pair (i.e., x and y) to encapsulate the information so that neither the intermediary platform nor the system participant can decrypt the ciphertext. The major actions that the server must take are to obstruct the encrypted requirement

from the requestor and remove the obstructive signal from the result returned from the intermediate platform. The final candidate set will be sent to the requestor for selection and confirmation.

4.1.4. Intermediate Platform. The intermediate platform is isolated from the server, and they are always from different service providers. They work under semitrustful condition and abide by the rules of the proposed system. The actions it takes are to distribute the Paillier cryptosystem-based public key, receive obstructed data from the server and execute matching algorithm, encrypt the result with the requestor's public key, and return it to the server.

4.1.5. CA. CA is the authority only responsible for identity verification. It is usually set up by the government. For requestors and providers, they send their real identity to CA for pseudonyms. The intermediate platform and the server can verify the identity of the requestor and the provider by pseudonyms through CA.

Specifically, the server and the intermediate platform can access the public keys of requestors from the database, and as long as the results are determined, the information can be guaranteed to be returned to the parking space provider.

4.2. Detailed Workflow. To realize the privacy-preserving parking space sharing in the target scenario, some interactive steps must be followed in the system. As the number shown in Figure 4, we divide the requestor-provider matching process into five main phases. All participants are assumed having registered in the system, e.g., their identities can be verified by the official authority. The parking space providers store information about their available provisions in the server and update it from time to time. The server can not disclose these information because they are encrypted by the public key of the intermediate platform. Some useful symbols in the algorithms are given in Table 1.

TABLE 1: Notations.

U	The requestor
P_i	The i^{th} provider
T_U, T_{P_i}	The time information of the requestor and providers
L_U, L_{P_i}	The location information of the requestor and providers
R	The result set
SK_U, PK_U	Secret-public key pair generated by the requestor
SK_I, PK_I	Secret-public key pair generated by the intermediate platform
B_j	The j^{th} block.
$PK()$	Encryption using public key
$SK()$	Decryption using secret key
x, y	Key pair generated by the server
$()^{pk}$	The information is encrypted

Phase 1: request posting. The requestor who wants to secure a parking space near his destination should post a request to the intermediate platform as early as possible.

Phase 2: new process launching. During the initialization of the intermediate platform, it generates Paillier key pairs (PK_I, SK_I) and makes the public key (PK_I) available to the server and all registered participants. The intermediate platform launches a new process when it receives a new request from a requestor and informs the server that it is ready for the obstructed data.

Phase 3: information collecting. After getting the notification from the intermediate platform, the requestor uses the key PK_I to encrypt his parking information including location and time as given in Eq. (4). The requestor then sends the encrypted information ($\text{inf}_U^{\text{pk}_I}$) to the server. However, he needs to let the server know the subarea of the destination so as to reduce computational cost. The information of provisions ($\text{inf}_{P_i}^{\text{pk}_I}$) of providers is already stored in the server using the key PK_I as given in Eq. (5), as well as the ID of subarea they belong to. Once the server knows which subarea the requestor is going to, it fetches the corresponding enciphered information and uses his key pairs (x and y) to twice-encrypt the information as given in Eq. (6). Finally, the server sends these information ($\text{inf}_S^{\text{pk}_I}$) to the intermediate platform. In Eq. (7), it represents the process of using x and y for secondary processing. The specific reason why they are equal is as shown in Eq. (3).

$$\text{inf}_U^{\text{pk}_I} = PK_I((T_U, L_U)), \quad (4)$$

$$\text{inf}_{P_i}^{\text{pk}_I} = PK_I((T_{P_i}, L_{P_i})), \quad (5)$$

$$\left(\text{inf}_U^{\text{pk}_I}\right)_{x,y} = PK_I(x \cdot (L_U, T_U) + y), \quad (6)$$

$$\text{inf}_S^{\text{pk}_I} = \left(\text{inf}_U^{\text{pk}_I}, \text{inf}_{P_i}^{\text{pk}_I}\right)_{x,y} \quad (7)$$

Phase 4: assignment calculating in this step, the intermediate platform receives the encrypted information ($\text{inf}_S^{\text{pk}_I}$) from the server, which is under double encryption of both the platform and the server. The platform uses its own private key to unlock the cryptograph previously encrypted with its public key and get the ciphertext (still under encryption of the large integer pair by the server) as given in Eq. (8). This ensures that the platform is not able to read the original user data. Then, the intermediate platform calculates and finds the right assignment result ($R \cdot x + y$) on these ciphertext (the specific calculation process will be analyzed in detail in the next section). The requestor generates the Paillier key pairs (PK_U, SK_U) and publishes his public key (PK_U) to the intermediate platform, with which the intermediate platform uses to encrypt the assignment result information ($PK_U(R \cdot x + y)$) and sends it to the server, which ensures that the server is unable to read the original user data.

$$\text{inf}_S = SK_I\left(\text{inf}_S^{\text{pk}_I}\right) = (\text{inf}_U, \text{inf}_{P_i}, \text{inf}_{B_i})_{x,y}. \quad (8)$$

Phase 5: results returning when the server gets the result information ($PK_U(R \cdot x + y)$), it uses its own pair keys (x and y) to decrypt the result information ($PK_U(R)$) using Eq. (9). At this stage, the result is still in ciphertext so that the server cannot read it. Finally, the server sends this information ($PK_U(R)$) to the requestor. The requestor uses his private key to get the result (R). From the result, the requestor should choose one desired parking space within a limited time period. If he fails to choose a parking space in time (time out) or cancels the result for three times, the requestor will be banned for a few while. This can stop DDoS attack to some extent. Upon the selected parking space is booked, the status will be updated to reserved, so that multiple reservation attack can be prevented. Moreover, two transactions are going to be conducted, i.e., a deposit will be paid to the intermediate platform, and the rent has to be paid to the space provider. To avoid exposing identifications, the payment goes through the blockchain-based strategy as proposed in [31]. When the requestor leaves the parking space in accord with the signed contract, the deposit will be returned by the intermediate platform. Otherwise, the deposit will be confiscated to punish inappropriate behaviors.

During above process, the operation of homomorphic addition satisfies our design for result and information delivery, and it removes x and y that can be realized using

$$PK_u(R) = \sqrt[\ast]{(PK_U(R \cdot x + y)/PK_U(y))}, \quad (9)$$

where PK_u is the requestor's public key, and R is the result of the calculation.

4.3. Block Algorithm Based on Longitude and Latitude. In order to solve the matching problems in privacy-preserving parking space sharing, in this part, the block algorithm based on latitude and longitude (BABLL) theory is shown in Alg. 1. Next, a detailed introduction of the calculation in the intermediate platform will be given.

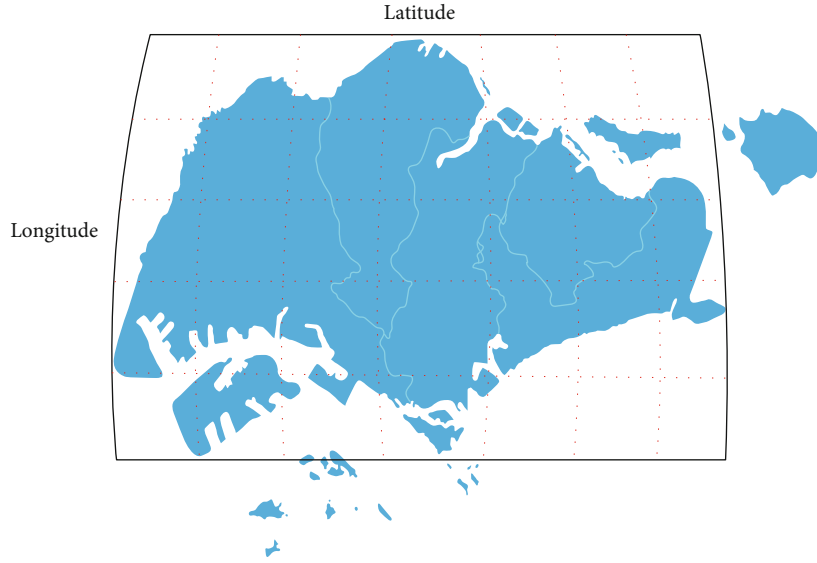


FIGURE 5: Divide the map of Singapore into blocks.

Neither the intermediate platform nor the server knows the actual location of each parking space; therefore, it may incur a large number of comparisons which is not efficient for the system. In this paper, we propose to divide the service area into subareas (i.e., blocks). Basically, in a city, the external trapezoid of this city may be drawn using small scales of latitude and longitude. However, it is a trade-off between the security and efficiency. Fine-grained division will bring higher efficiency but lower security, as in a small block, and there may be only a few parking spaces, so that other users may dope out the information of the provider. In this paper, we divide the trapezoid according to the length of the latitude and longitude to reduce the operation time. Figure 5 shows how we divide this trapezoid block on a rough map of Singapore. The length is about 7 km along the line of latitude and is 5.2 km for longitude.

From the previous section, we already know that we can implement the comparison and sorting mechanism in the case of encryption, and we can already specifically determine which block the requestor's parking place is. According to the requestor's preference, we need to further shrink the large block to a smaller subblock, such as the one surrounded by blue dashed lines in Figure 6. For example, the requestor may start with the small subblock such as 200 meters by 200 meters. If no available parking space is found, then the subblock can be gradually increased. Because the distance of a small segment along latitude at a certain longitude is determined, then the intermediate platform can match the available parking spaces within a certain range of the destination without knowing the exact location and the range. The calculation only needs additions, subtractions, and comparisons on ciphertexts which are supported by homomorphic encryption. The locations of these parking spaces form the result set, which will be returned to the requestor. Below, we divide the process into two situations:

The first case: the search area is within a block. Since all the data belonging to the same block as the destination has

been uploaded to the intermediate platform, the search of parking spaces can limit in this block as long as the requestor's ideal range does not exceed the boundary. As shown in Figure 6, the red dots represent free parking spaces while the black dot is where the requestor wants to go. The blue box stands for the parking range that the requestor prefers.

The second case: the search area is across two or four blocks, and more data of adjacent blocks needs to be obtained from the server during matching. As shown in Figure 7, at the beginning, only the data from the block on the right is uploaded to the intermediate platform. When the requestor preferred range extends to another block, the intermediate platform will get stuck during calculation and then post a request to load parking information from the block on the left. The intermediate platform only asks for the adjacent blocks without knowing the actual block number. The server selects the block according to the request.

The matching process takes place in the intermediate platform. Here, the detailed workflow of the block algorithm based on longitude and latitude is illustrated in Algorithm 1. It mainly consists of three primary phases.

Phase 1: data fetching. The intermediate platform fetches all information ($T_U * x + y$, $T_{P_i} * x + y$, $L_U * x + y$, $L_{P_i} * x + y$ and $B_j * x + y$) from the server.

Phase 2: choose the right block. The intermediate platform makes sure whether the location (L_U) with given range is in the block based on two pairs of latitude and longitude information. If no, the intermediate platform needs to require the server to provide additional block(s).

Phase 3: find the right result set. The intermediate platform searches within the blocks to find all shared parking spaces satisfying the time preference and finally returns the sorted matching result (in descending order of distance). Note that only parking spaces located within the range given by the requestor will be checked, and the result will be encrypted by the public key of the requestor before returning it to the server.

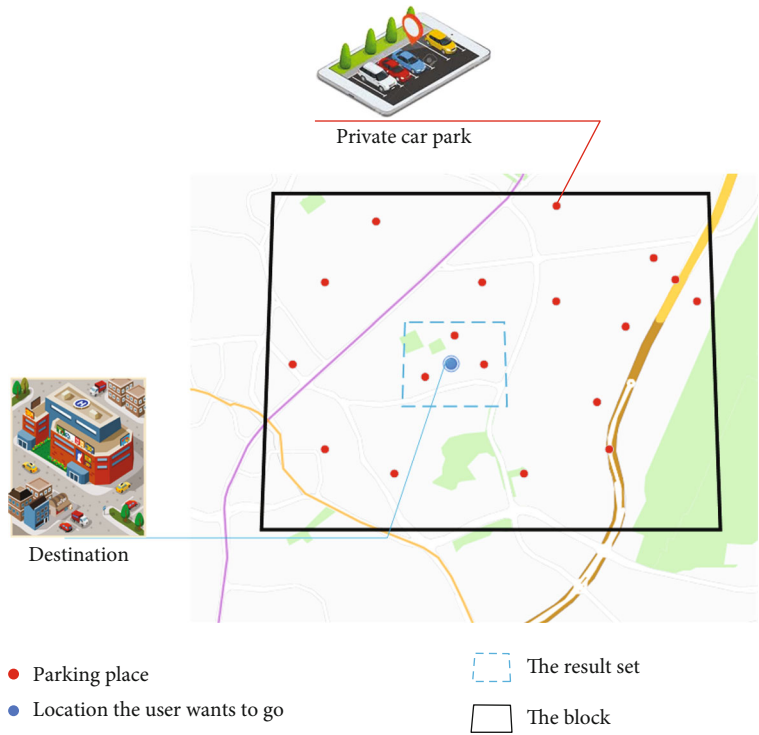


FIGURE 6: Search within a block.

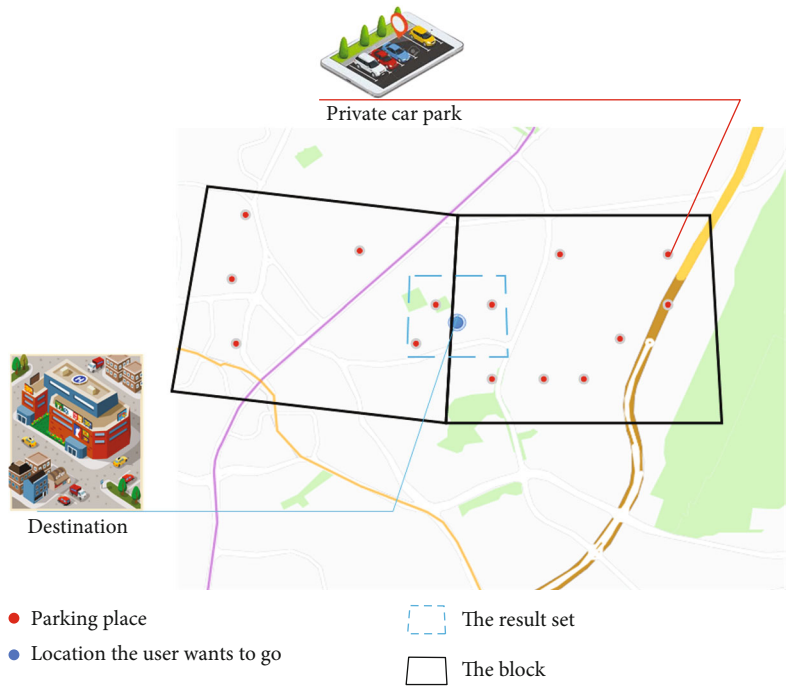


FIGURE 7: Search across blocks.

5. Performance Evaluation

In this section, we are going to illustrate the performance evaluation by first giving security analysis and then presenting the numerical results. Three state-of-the-art algorithms, i.e., IoT-UPM [7], campus parking [15], and blockchain-

based [28], are compared in terms of security features and computation efficiency.

5.1. Security Analysis. Security is the first priority in the proposed privacy-preserving matching scheme (PPMS) and is analyzed from the following three aspects:

```

1: Fetch the information ( $PK_p(T_U \cdot x + y)$ ,  $PK_p(T_{P_i} \cdot x + y)$ ,
 $PK_p(L_U \cdot x + y)$ ,  $PK_p(L_{P_i} \cdot x + y)$  and  $PK_p(B_j \cdot x + y)$ )
from the server; use  $SK_p$  to decrypt the information into
( $T_U \cdot x + y$ ,  $T_{P_i} \cdot x + y$ ,  $L_U \cdot x + y$ ,  $L_{P_i} \cdot x + y$  and  $B_j \cdot x + y$ )
Input:  $T_U \cdot x + y$ ,  $T_{P_i} \cdot x + y$ ,  $L_U \cdot x + y$ ,  $L_{P_i} \cdot x + y$  and
 $B_j \cdot x + y$ ;
Output:  $R_S$ ;
2: if trapezoid of  $L_U$  is at the boundary of  $B_j$  then
3:   require adjacent blocks and merge them into  $B_j$ 
4: end if
5: for  $T_{P_i}$  ( $i=1$  to  $m$ ) in  $B_j$  do
6:   if  $T_U$  matches with  $T_{P_i}$  then
7:     get  $T_{P_i}$  and merge into  $L_p$ 
8:   end if
9: end for
10: sort  $L_p$  by distance and form the result set( $R_S$ )
11: return  $R_S$ 

```

ALGORITHM 1 BABLL.

Aspects 1, manipulated message: in the PPMS, all information that involves communication between entities is encrypted to prevent eavesdropping from adversaries. In the Paillier cryptosystem, even for the same plaintext, the encryption result will be totally different each time with the same public key. Thus, statistical attacks are avoided.

Aspects 2, authentication: public certificate authority (CA) server is an essential part of the VANET; in this paper, we also assume that all participants need to register with the CA and get a few pseudonyms. The intermediate platform and the server can use these pseudonyms to verify the identity of participants. Unauthorized access will be immediately denied. The pseudonyms will be updated at a given period of time. Therefore, the system is robust under DDoS or replay attack. CA is only responsible for ID authentication, and it does not access any data of requestors and parking space providers.

Aspects 3, privacy disclosure: true IDs of participants (requestors and providers) are protected by the public CA using pseudonyms. Location, time preference, distance, and other sensitive data are calculated and matched in the form of encryption, so as to ensure no privacy leakage. Since the server will change big integers x and y constantly, there is little chance for the intermediate platform to predict the blocks, not to mention the actual location.

In addition to the above security analysis, we compare the proposed method (PPMS+BABLL) with the other three methods in terms of different security attributes, as shown in Table 2.

“N/A” means in the corresponding method, the data is not needed for making decision. For user identification, in the blockchain-based method, super peers and privileged users can access them. Many other attributes are evaluated as partial because privacy data (such as price and location) is open to all other peers, and only the ID is protected. In fact, many attack methods can connect the ID with the data via

machine learning or big data analysis. In many cases, the attack can happen without knowing the ID, e.g., a free parking space could be maliciously occupied without necessarily knowing the owner’s ID. Campus parking must be running under honesty model, such that parking permission status $X_i(k+1)$ and parking probability $p_i(k)$ for user i at time k is trustworthy. Moreover, some sensitive information, such as destination of requestors and parking duration of requestors, are only hidden from other users. Since there is only one parking lot, the location of the provider is not able to be protected. Regarding Iot-UPM, it is an auction-based method, so these sensitive information needs to be uploaded to the cloud for assignment. Compared with other methods, they either only need to accept the requestor to a fixed parking lot or users themselves are responsible for searching preferred parking space through open data, and our work has implemented high strength in privacy protection and excellent performance in resource allocation.

5.2. Numerical Results. To verify the performance of the proposed PPMS, we choose Singapore as the target city and randomly generate the data sets. At first, 100 requestors are simulated to demand parking spaces and then extend to 1000; finally, to 10000, so that the results can be observed under different network scale. The number of providers are 5000, 10000, and 35000. Each data set is tested for multiple times to eliminate the effect of odd locations. The experiment is running on a computer with 2.6 GHz Core i5 CPU and 8 GB memory.

As shown in Table 3, the computational complexity of each operation is listed. In Paillier cryptography, the length of the key determines the security strength. The duration of a single operation remains constant once the length is determined. As can be seen from the table, the time complexity of each password-related operation is $O(1)$. The sorting operation under encryption has a time complexity of $O(n \log n)$, where n is the number of elements in the sorting list. For each operation, we list the count used by each entity in the remaining columns of Table 3. Since the time complexity is no more than $O(n \log n)$, we consider the proposed scheme efficient.

In Table 4, we show the storage cost of different key lengths. Theoretically, a longer asymmetric key can provide greater security. However, the storage overhead is also significant. Using a 2048-bit key (which is sufficient in the near future), each ciphertext and signature require 512 bytes storage space. Then, to protect the 8-byte data, an overhead of 1536 bytes occurs including two signatures and one ciphertext.

We now analyze the effectiveness of the improved PPMS solution in terms of time and space efficiency, in comparison with one state-of-the-art-related schemes, a price-compatible top trading cycles, and chains mechanism [7] (IoT-UPM for short). We have implemented our PPMS and BABLL prototype in Java, including the Paillier cryptosystem following the workflow described in Section 4. The communication cost is omitted since the amount of data transmitted is relatively small compared with the bandwidth of 4G or WiFi.

First, we evaluate the time performance of the system. The key length that we use is 2048 bits, which provides sufficient security. The numbers of parking requests are set to

TABLE 2: Security and privacy comparison.

Attributes	PPMS+BABLL	Blockchain-based [28]	Campus parking [15]	IoT-UPM [7]
User identification	Yes, via CA	Partial	Partial	No
Destination of requestors	Yes	Partial	No	No
Parking duration of requestors	Partial	No	Partial	No
Cost function of requestors	N/A	N/A	Yes	N/A
Location of providers	Yes	Partial	No	No
Availability period of providers	Yes	Partial	No	No
Price of providers	Yes	Partial	N/A	No
Burst requests	Yes	No	Yes	No
Efficiency	Medium	Low	High	Medium

TABLE 3: Composition of operations in each entity.

	$O()$	Server	Provider	Requestor	IP
Key gen.	$O(1)$	0	0	1	1
Encryption	$O(1)$	1	2	$2n$	0
Decryption	$O(1)$	0	0	1	$n + 2$
Padding	$O(1)$	$n + 2$	0	0	0
Sorting	$O(n \log n)$	1	0	0	n
CA op	$O(1)$	2	2	2	n

TABLE 4: Per item space cost over key length.

Key length	512 bits	1024 bits	2048 bits
Key size	192 bytes	384 bytes	768 bytes
Ciphertext	128 bytes	256 bytes	512 bytes
Signature	128 bytes	256 bytes	512 bytes
Message	392 bytes	776 bytes	1542 bytes

TABLE 5: System Performance on time consumptions.

	100	1000	10000
Total time	2573 s	25238 s	255722 s
Avg. time	25.73 s	25.24 s	25.57 s
Success rate	100%	100%	100%

100, 1000, and 10000, respectively. As shown in Table 5, we record the processing time from request posting to result set returning. The average waiting time is around 25 seconds which is acceptable as the success rate is keeping at 100%. We can also find that from 100 request to 10000 requests, the waiting time is very similar, because the major part of time is spent on encrypting not matching. The time efficiency is compared with IoT-UPM [7] and blockchain-based [28]. For BCOS adopted by the blockchain-based method, we set the ratio of privileged nodes to 0.5%. Figure 8 shows that the proposed PPMS-BABLL outperforms the time efficiency of IoT-UPM and blockchain-based under encryption. Compared to the scheme in IoT-UPM, the processing time in PPMS is significantly reduced. Regarding the blockchain-based method, with the increasing of participators, the cost of processing rises very quickly. When the number of partic-

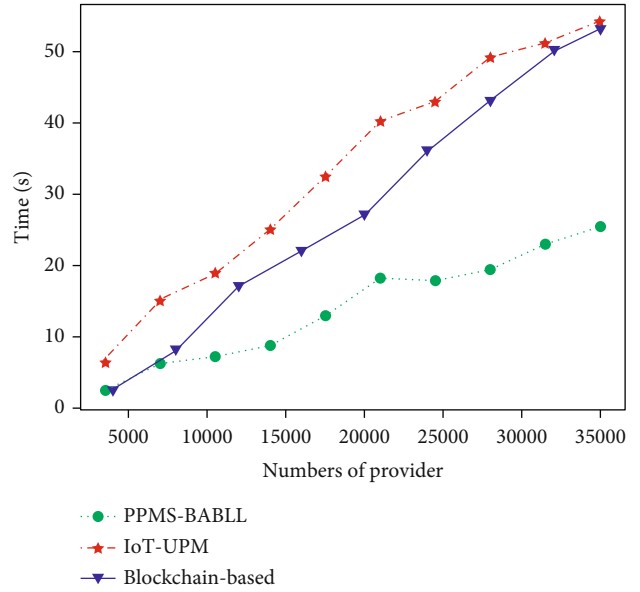


FIGURE 8: Comparison of time efficiency.

TABLE 6: System performance.

	Total number	Completed	Terminated
Process	10	10	0

ipators is small, the blockchain-based method can achieve similar performance with our method, but it grows higher than PPMS-BABLL later to ensure security. This is due to the proposed PPMS-BABLL only fetches a small part of information for comparison and sorting in the intermediate platform.

We then evaluate the safety measurements of PPMS. In particular, we generate manipulated messages and simulate the attacks mentioned in Section 4 to verify the security of PPMS. We repeat experiments for 100 times where the total number of providers is set to 35000 and 200 for requestors. We manipulate 10% of the messages and record the rejection rate as well as the time if the process terminate due to attack. Table 6 illustrates the result of manipulating message security. As can be seen, all manipulated messages are rejected, while normal messages are accepted. Again, all 10 processes are completed correctly without termination. Therefore, it

shows that the system not only has a strong ability to resist such attacks but also has lower processing time.

6. Conclusions

In this paper, a privacy-preserving vehicle assignment problem in the parking space sharing system is studied. A homomorphic encryption-based privacy protection matching scheme (PPMS) is designed. In order to reduce the overhead of the proposed PPMS, a block algorithm based on the longitude and latitude (BABLL) is proposed. Through the security analysis, the scheme is proved to be able to protect the privacy of sensitive information such as location, time, identity of both requestors, and space providers. The scheme is also robust against attacks, e.g., DDoS or replay. We implement the prototype system and conduct comparative experiments. The results show that the proposed scheme can ensure very good success rate of matching with high time efficiency. In addition, the system resists multiple rounds of practical attacks, while maintains normal operations.

Data Availability

Data available on request, please contact the corresponding author Peng Liu (perryliu@hotmail.com).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is supported by the Opening Project of Guangdong Provincial Key Laboratory of Information Security Technology (Grant No. 2017B030314131) and the Natural Science Foundation of China (Grant No. 61601157).

References

- [1] P. Liu, B. Xu, G. Dai, Z. Jiang, and J. Wu, "MDP: Minimum delay hot-spot parking," *Journal of Network and Computer Applications*, vol. 87, pp. 210–222, 2017.
- [2] S. Marcu and A. Florea, "Smart parking system - another way of sharing economy provided by private institutions," in *2018 Thirteenth International Conference on Digital Information Management (ICDIM)*, pp. 18–23, Berlin, Germany, 2018.
- [3] A. Shahzad, J.-y. Choi, N. Xiong, Y.-G. Kim, and M. Lee, "Centralized Connectivity for Multiwireless Edge Computing and Cellular Platform: A Smart Vehicle Parking System," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 7243875, 23 pages, 2018.
- [4] S. Wan, X. Li, Y. Xue, W. Lin, and X. Xu, "Efficient computation offloading for Internet of Vehicles in edge computing-assisted 5G networks," *The Journal of Supercomputing*, vol. 76, no. 4, pp. 2518–2547, 2020.
- [5] R. Liu, Y. Yang, D. Kwak, D. Zhang, L. Iftode, and B. Nath, "Your Search Path Tells Others Where to Park," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, 2017no. 3.
- [6] L. Zhang, X.-Y. Li, Y. Liu, and T. Jung, "Verifiable private multiparty computation: ranging and ranking," in *2013 Proceedings IEEE INFOCOM*, pp. 605–609, Turin, Italy, 2013.
- [7] X. T. R. Kong, S. X. Xu, M. Cheng, and G. Q. Huang, "IoT-enabled parking space sharing and allocation mechanisms," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 4, pp. 1654–1664, 2018.
- [8] B. Tan, K. Kang, S. Xu, T. Qu, and C. Li, "First- and second-price sealed-bid auctions applied to private parking slot sharing," in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, pp. 1–5, Zhuhai, China, 2018.
- [9] P. Zhao, H. Guan, and P. Wang, "Data-Driven Robust Optimal Allocation of Shared Parking Spaces Strategy Considering Uncertainty of Public Users' and Owners' Arrival and Departure: An Agent-Based Approach," *IEEE Access*, vol. 8, pp. 24182–24195, 2020.
- [10] L. Yuan, "Study on the design of car-sharing parking space in the background of smart citya case study of suzhou," in *2019 International Conference on Robots Intelligent System (ICRIS)*, pp. 382–385, Haikou, China, 2019.
- [11] J. Ma, L. Zhang, and Y. Xu, "Time-shared parking mechanism and application based on baas cloud infrastructure," in *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA)*, pp. 533–536, Beijing, China, 2017.
- [12] X. Liu, H. Hu, and Z. Deng, "The Shared Allocation Model of Night Parking Spaces between Residential Areas and Adjacent Business Districts," in *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*, pp. 527–532, Dublin, Ireland, 2019.
- [13] O. T. T. Kim, N. Dang Tri, V. D. Nguyen, N. H. Tran, and C. S. Hong, "A shared parking model in vehicular network using fog and cloud environment," in *2015 17th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 321–326, Busan, South Korea, 2015.
- [14] D. Kondor, H. Zhang, R. Tachet, P. Santi, and C. Ratti, "Estimating Savings in Parking Demand Using Shared Vehicles for Home-Work Commuting," *IEEE Transactions on Intelligent Transportation Systems*, no. 8, pp. 2903–2912, 2018.
- [15] W. Griggs, J. Y. Yu, F. Wirth, F. Hausler, and R. Shorten, "On the design of campus parking systems with qos guarantees," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1428–1437, 2016.
- [16] L. Qi, X. Zhang, S. Li, S. Wan, Y. Wen, and W. Gong, "Spatial-temporal data-driven service recommendation with privacy-preservation," *Information Sciences*, vol. 515, pp. 91–102, 2020.
- [17] M. Xiao, J. Wu, S. Zhang, and J. Yu, "Secret-sharing-based secure user recruitment protocol for mobile crowdsensing," *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, 2017, Atlanta, GA, USA, 2017.
- [18] C. Miao, L. Su, W. Jiang, Y. Li, and M. Tian, "A lightweight privacy-preserving truth discovery framework for mobile crowd sensing systems," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, Atlanta, GA, USA, 2017.
- [19] X. Tang, C. Wang, X. Yuan, and Q. Wang, "Non-interactive privacy-preserving truth discovery in crowd sensing applications," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, Honolulu, HI, USA, 2018.

- [20] C. Zhang, L. Zhu, C. Xu, X. Liu, and K. Sharif, "Reliable and Privacy-Preserving Truth Discovery for Mobile Crowdsensing Systems," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2019.
- [21] X. Liu, R. Deng, K. R. Choo, and Y. Yang, "Privacy-preserving reinforcement learning design for patient-centric dynamic treatment regimes," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–1, 2019.
- [22] D. An, Q. Yang, W. Yu, D. Li, and W. Zhao, "LoPrO: Location Privacy-preserving Online auction scheme for electric vehicles joint bidding and charging," *Future Generation Computer Systems*, vol. 107, pp. 394–407, 2020.
- [23] D. Song, M. Song, and K. Park, "A Privacy-Preserving Spatial Index for Spatial Query Processing," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 2067047, 9 pages, 2018.
- [24] X. Xu, Y. Xue, L. Qi et al., "An edge computing-enabled computation offloading method with privacy preservation for internet of connected vehicles," *Future Generation Computer Systems*, vol. 96, pp. 89–100, 2019.
- [25] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proceedings of the 41st annual ACM symposium on Symposium on theory of computing - STOC '09*, pp. 169–178, 2009.
- [26] Q. Liu, G. Wang, and J. Wu, "An efficient privacy preserving keyword search scheme in cloud computing," in *2009 International Conference on Computational Science and Engineering*, vol. 2, pp. 715–720, Vancouver, BC, Canada, 2009.
- [27] C. Xu, L. Zhang, L. Zhu et al., "Aggregate in my way: Privacy-preserving data aggregation without trusted authority in ICN," *Future Generation Computer Systems*, vol. 111, pp. 107–116, 2020.
- [28] J. Hu, D. He, Q. Zhao, and K.-K. R. Choo, "Parking management: a blockchain-based privacy-preserving system," *IEEE Consumer Electron. Mag.*, vol. 8, no. 4, pp. 45–49, 2019.
- [29] R. Gennaro, C. Gentry, and B. Parno, "Non-interactive Verifiable Computing: Outsourcing Computation to Untrusted Workers," in *Advances in Cryptology - CRYPTO 2010*, p. 547, Springer, Berlin, Heidelberg, 2010.
- [30] P. Liu, Y. Ding, and T. Fu, "Optimal ThrowBoxes assignment for big data multicast in VDTNs," *Wireless Networks*, 2019.
- [31] S. R. Pokhrel, Y. Qu, S. Nepal, and S. Singh, "Privacy-Aware Autonomous Valet Parking: Towards Experience Driven Approach," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2020.

Research Article

A Privacy-Preserving Personalized Service Framework through Bayesian Game in Social IoT

Renwan Bi,¹ Qianxin Chen,¹ Lei Chen,² Jinbo Xiong¹ ,¹ and Dapeng Wu³

¹Fujian Provincial Key Laboratory of Network Security and Cryptology, College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China

²College of Engineering and Computing, Georgia Southern University, GA 30458, USA

³Chongqing Key Laboratory of Optical Communication and Networks, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Correspondence should be addressed to Jinbo Xiong; jbxiong@fjnu.edu.cn

Received 25 June 2020; Revised 10 September 2020; Accepted 4 October 2020; Published 17 October 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Renwan Bi 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.

It is enormously challenging to achieve a satisfactory balance between quality of service (QoS) and users' privacy protection along with measuring privacy disclosure in social Internet of Things (IoT). We propose a privacy-preserving personalized service framework (Persian) based on static Bayesian game to provide privacy protection according to users' individual security requirements in social IoT. Our approach quantifies users' individual privacy preferences and uses fuzzy uncertainty reasoning to classify users. These classification results facilitate trustworthy cloud service providers (CSPs) in providing users with corresponding levels of services. Furthermore, the CSP makes a strategic choice with the goal of maximizing reputation through playing a decision-making game with potential adversaries. Our approach uses Shannon information entropy to measure the degree of privacy disclosure according to the probability of game mixed strategy equilibrium. Experimental results show that Persian guarantees QoS and effectively protects user privacy despite the existence of adversaries.

1. Introduction

The rapid development of cloud computing and big data technologies has greatly promoted work productivity and life quality. Along with such advancement, there come frequent user privacy disclosures that have attracted wide attention from academia and the industry [1]. In recent years, thanks to the marriage of wireless technologies [2] and mobile communications, social networks (SNs) have become an indispensable part of life [3]. Social networks enable communications and services far beyond instant messaging compared to traditional messaging services and applications [4]. The content of transmission has also become more diverse, including text, voice, image, video, and other multimedia data [5]. Data owners (e.g., mobile and smart device users) enjoy personalized services by gaining various application privileges while data collectors (e.g., service providers and application developers) obtain vast amounts of personal sensitive and security-critical data through privileged interfaces

[6]. Such user data become attractive targets of attacks and are subject to serious privacy disclosures [7, 8].

According to the "China Privacy Risk Index Analysis Report" published by the trusted institutions in 2018, mobile Internet applications in the social category have an average of 11,014 users per App, and the average amount of data acquired can reach up to 21.24 pieces/user, which is the most among all categories of Apps. At present, the number of user data leakages increased by 15.46% from 2018, and the privacy risk index increased by 26.66% [9]. Users inevitably leave a trail of footprints in the real world while accessing online services from a mobile device [10]. For example, people share various information on Twitter, and even when the original blog is deleted, relevant comments remain on the web [11]. Additionally, 267 million Facebook users' information, including names, gender, email addresses, and social identity, are stolen in April 2020 and sold on the dark web [12]. Thus, protecting user's privacy and security is critically important yet challenging in social networks.

Social networking aspects, in recent years, have been extended to the Internet of Things (IoT) that autonomously build social relationships for smart devices to discover new objects and their services [13, 14]. The marriage of IoT and social network enables advanced and deep interactions among people and between people and the environment. Such advancement leads to the emergence of social IoT [13], where social approaches are employed for managing large volumes of user data with connected IoT devices [6, 15]. This can result in a greater challenge for user privacy in social IoT. Efficient and effective IoT nodal interactions rely on the establishment of trustworthy relationship among nodes [15, 16]. This is particularly important in helping overcome the perceptions of uncertainty and privacy risk [17, 18].

Two main reasons contribute to vast amounts of user privacy disclosure in social IoT [19]: technical deficiencies and economic interest conflicts among all participants. Therefore, the privacy protection of social IoT users should also be analyzed from these two aspects. At present, from the technical perspective, user privacy is protected mainly through anonymity [20, 21], differential privacy [22, 23], network access control [24, 25], and ecosystem [26] in social IoT.

Anonymity protection [20] hides private data in a data block so that other users are incapable of associating a user's real identity information with the collected data blocks. This is also a common problem of anonymous protection schemes, and it is difficult to defend against background knowledge attacks. Differential privacy protection [22, 27] adds small amounts of Laplace noise into the original data prior to publishing the data for added fuzziness where other users are incapable of distinguishing between the real data and the fuzzy data. Zhang et al. [27] attach importance to the social connection of users, consider the existence of untrusted service providers and malicious attackers, and propose and implement an effective IoT service with differential privacy protection. Despite the high data utility, it is difficult to implement personalized privacy at the user level according to the user's security requirements using differential privacy [23]. Network access control [24, 28] decides whether to grant authorized access by analyzing the credibility and closeness among visitors. While it implements privacy protection according to the wishes of data owners, it lacks an effective privacy measurement scheme.

Among the notable research works on the security and privacy of social IoT [19], Frustaci et al. addressed that security and privacy issues are a great challenge for IoT and yet they are also enabling factors to create a "trust ecosystem" [26]. Particularly, in their discussions of the importance of trust, excellent flexibility is considered as a critical factor to deal with changeable security conditions and personalized security requests. Users or nodes having defined personalized security and privacy policies should be facilitated to help in decision-making [26].

Considering big data privacy protection from the perspective of users' interests and economics, the existing literatures [29] mainly describe the benefits and costs of participants by employing game theory, simulate the rational selection process, and formulate the optimal privacy protection scheme through Nash equilibrium [21]. To some extent,

these methods make up for the defects of technical schemes; yet how to balance the data utility and efficiency of privacy protection is still a difficult open issue.

In order to tackle this problem, this paper proposes the Persian framework, aiming at providing personalized services to social IoT users on the basis of protecting user privacy. Particularly, in order to resist against adversaries with background knowledge, static Bayesian game theory is applied to the strategic struggle between CSP and adversaries. The contributions in this paper are summarized as follows:

- (i) *Implementing User Classification.* Aiming at the difficult problem of users' discrete attribute classification, we adopt the fuzzy uncertainty reasoning method to classify users according to the membership function and expert rules
- (ii) *Defining Trust and Security Responsibilities of the CSP.* We construct a trust management center (TMC) that supervises the CSP's behavior and evaluates each service. TMC employs the incremental update strategy to manage the reputation of the CSP, so as to avoid CSP from proactively disclosing user privacy
- (iii) *Achieving a Satisfactory Balance between Quality of Service (QoS) and Privacy Protection.* We use the mixed strategy equilibrium to explain the correctness of the CSP making services strategies against different types of adversaries. Moreover, we utilize Shannon information entropy to measure the privacy disclosure, thereby providing a theoretical basis for users' privacy protection in social IoT. Experimental results show that the Persian framework achieves the correct user privacy classification and trust assessment, while privacy disclosure is limited to a low degree

The rest of this paper is organized as follows: The related work is overviewed in Section 2. We illustrate the preliminaries, including fuzzy uncertainty reasoning and Bayesian game theory in Section 3. The system model and security model are introduced in Section 4. In Section 5, we present the modules of the Persian framework in detail, including user classification, trust management, Bayesian game, and privacy measurement. Experiments and evaluation are described in Section 6, and the study is concluded in Section 7.

2. Related Work

In social IoT, protecting user privacy has been a research hotspot due to frequent user privacy disclosures. Existing literatures mainly present from the two aspects of technology and economic interests. Anonymity, differential privacy, access control, and trust management are often adopted to protect users' privacy in social IoT.

Liu et al. [30] proposed a k -anonymous algorithm, which generates an initial weighted social network and reduces the adjustment of relation weight through the sorting process. It improves anonymity efficiency and resists against

Neighborhood attacks. However, this approach does not significantly improve the utility of anonymous data. Xie and Zheng [31] proposed a differential social network anonymous algorithm satisfying k -anonymity and l -diversity. For the key nodes and general nodes, the proposed algorithm uses different types of anonymous operations to transfer anonymous objects from privacy attributes to anonymized sensitive attributes. Based on implementing privacy protection, the proposed algorithm improves the utility of anonymous data. Furthermore, an indicator (UL) is constructed to measure the data utility loss. Chen et al. [32] proposed a classification data clustering scheme based on rough entropy and DBSCAN clustering algorithm, which effectively balances data utility and anonymity performance of mobile social networks. Nonetheless, it falls short on formal security analysis against the attacker.

Li et al. [33] proposed an MB-CI strategy for protecting the edge weights of social networks, which retains most of the shortest paths under the premise of satisfying differential privacy, effectively reducing the error caused by noise and improving the accuracy of published data. At the same time, it effectively resists against the consistent reasoning attacks on data records without user-level privacy protection.

Wang et al. [34] proposed a data publishing algorithm (RescueDP) satisfying differential privacy to protect real-time and spatiotemporal crowd-sourced data in social networks. They also proposed an enhanced neural network algorithm, which accurately predicts statistical data with added noise, thereby improving the utility of published data. Huang et al. [35] proposed a differential privacy protection method (PBCN) based on clustering and noise, aiming at achieving a “trade-off” between data availability and privacy protection level. Jahid et al. [36] proposed an encryption-based access control architecture (EASiER) to address privacy disclosure in online social networks. It transfers access control from the social network provider to users and implements fine-grained access control for dynamic social contacts using attribute-based encryption. Hu et al. [24] constructed a multiparty access control (MPAC) model and proposed a specific multiparty strategy specification scheme and strategy evaluation mechanism to protect the shared data associated with multiple users in online social networks against collusion attacks.

The existing literature also proposed a number of personalized privacy protection schemes. Cai et al. [37] proposed a data disinfection method for centralized processing of user configuration files and relationships among users. By controlling the set of user attributes and the relationship among users to hide sensitive information, the proposed method resists against the set inference attack in the process of data publishing in social networks. Cai et al. [38] proposed a privacy-preserving scheme for interactive messaging by leveraging user credibility and social behaviors, which guarantees the privacy protection in the process of information exchange through information confusion and sensitive attribute substitution. In order to solve the trust difficulties, Sharma et al. [39] proposed a novel solution in the form of fission computing. The proposed solution relies on the edge-crowd integration for maintenance of trust and preser-

vation of privacy rules in Social IoT, using crowdsources as mini-edge servers and entropy modeling for defining trust between the entities.

Additional literature also considers privacy protection from the perspective of interests and puts forward a number of effective schemes and models utilizing game theory. Jin et al. [40] applied game theory to trajectory privacy protection. For any two sensing nodes in the network, this method selects the best strategy through the Bayesian game analysis to resist against the dishonest attacks of internal nodes, thus protecting the trajectory privacy of users. Hu et al. [41] proposed a multiparty control game, which extends the research on strategy selection among rational controllers in multiparty access control. The Nash equilibrium is used to explain the optimal strategy selection state, and no controller has any valid reason or authority to change its settings to deviate from the equilibrium, which solves the privacy conflict of collaborative data sharing in online social networks from the perspective of interests. Wu et al. [29] proposed an extended game model to solve the problem of privacy and utility equilibrium in the publishing of multicorrelated privacy data, which solves the differential privacy parameters according to Nash equilibrium, thereby improving data utility. Shan et al. [42] proposed a forwarding control mechanism for social networks based on game theory. By calculating the game revenue matrix of the publisher and forwarder and comparing the probability of dishonest forwarding with the threshold set by the publisher, this approach protects the privacy of publishing content according to the personalized privacy requirements of the publisher. Xiong et al. [21, 23] also actively applied game theory to the privacy protection of the application environment.

Xiong et al. [43] conducted a comprehensive survey on the privacy measurement and quantification of big data. Serjantov and Danezis [44] used the Shannon entropy to describe the effective size of anonymous sets. Lin et al. [45] employed mutual information to measure privacy disclosure under the data protection mechanism. Diaz et al. [46] utilized conditional entropy to describe adversaries’ observation ability and indirectly measured the level of protection mechanism. Additionally, Chen et al. [47] proposed an information surprise indicator to measure the surprise degree that still exists after an adversary acquires user attributes.

3. Preliminaries

3.1. Fuzzy Uncertainty Reasoning. Fuzzy reasoning [48] is a method of uncertainty reasoning, which is suitable for any situation where the input fluctuates in a specific range. Also, the output is also fuzzy, rather than precise. The fuzzy concept is regarded as a membership degree [49], reflecting the closeness of input or output with a fuzzy set in the universe. If the membership degree equals 1, it means that the variable values belong to the fuzzy set completely; if the membership degree equals 0, it means that no elements belong to the fuzzy set absolutely. A membership value in $(0, 1)$ means that some elements, but not all, belong to the fuzzy level to some extent. The membership function replaces the positive or negative results with the fuzzy evaluation results, which is helpful for

considering the influence of multiple factors. Generally, the membership function of the gradient type is widely accepted, as shown in Figure 1. Upon finding the membership, the rule activation is performed. The fuzzy output is generated by the activation of finite rules.

3.2. Bayesian Game Theory. The static Bayesian game (SBG) model [50] is also known as static incomplete information game. The type set of all participants is known. Any participant can only infer the probability that other participants belong to a certain type at a certain time, but cannot determine other participants' type and cannot determine the relevant action strategies or benefit. Furthermore, all participants choose action strategies simultaneously in the game. Even if there are differences in the order of choosing, the participants who choose the strategy posterior do not have the knowledge of the selected strategy. SBG model [50] can be represented by a quintuple, $SBG = (\Gamma, T, P, S, U)$, which is described as follows:

- (1) Participant set is $\Gamma = \{1, 2, \dots, n\}$, where $n \geq 2$, because it is meaningless to discuss a game with only one participant. Any participant $i (i \in \Gamma)$ is a rational decision-maker with the ability of independent selection, whose goal is to maximize their expected benefit and choose action strategies
- (2) Participant type set is $T = \{T_1, \dots, T_n\}$, where T_i represents the participant i 's types, $i \in \Gamma$, and $|T_i| \geq 2$. If each participant has only one candidate type (i.e., $\forall i \in \Gamma$ and $|T_i| = 1$), at the point, the static incomplete information game will become the static complete information game
- (3) The probability set of participants' inference about the types of other participants is $P = \{P_1 < t_{-1} | t_1 >, \dots, P_n < t_{-n} | t_n >\}$, where $P_i < t_{-i} | t_i >$ represents the probability of participant i 's inference about the types of other participants. Meanwhile, t_i represents participant i 's type, and t_{-i} (i.e., $\{t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n\}$) represents all participants' type other than participant i
- (4) The participants' strategy set is $S = \{S_1(T_1), \dots, S_n(T_n)\}$, where $S_i(T_i) (i \in \Gamma)$ (i.e., $S_i = \{s_i(1; t_i), \dots, s_i(m; t_i)\}$) represents participant i 's strategy set and $s_i(j; t_i) (j \in \{1, \dots, m\})$ represents participant i 's strategy
- (5) The benefit function of participants is $U = \{U_1(T_1), \dots, U_n(T_n)\}$, where $U_i(T_i) (i \in \Gamma)$ represents the participant i 's benefit. Since $U_i = u_i(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n; t_i)$, participant i 's benefit is related to its own type and the strategy chosen by other participants

4. System Model and Security Model

We introduce important notations and descriptions in this paper, as shown in Table 1.

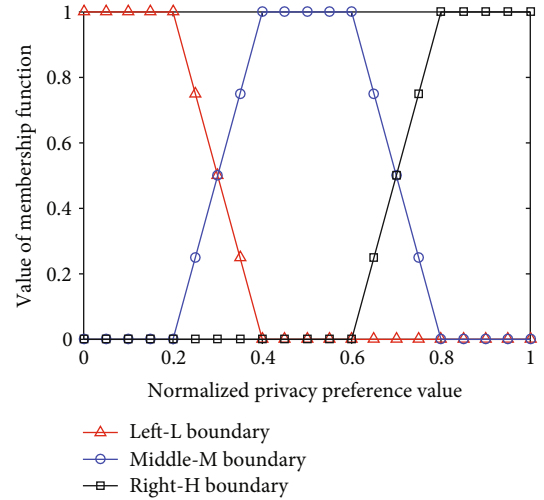


FIGURE 1: Membership function.

TABLE 1: Main notations and descriptions.

Notations	Descriptions
m	Trust depth
λ	Trust penalty factor
μ	Benefit factor
k	Background factor
θ	Trust threshold
δ_j^l	The reputation of CSP in j -th service
σ_j	Average reputation of CSP in previous j services
s	Action strategy set
u	Action benefit set
U	Total benefit of participant
P^*	Probability under mixed strategy equilibrium
H	Privacy information entropy

4.1. System Model. In social IoT, we mainly focus on how a CSP provides personalized services for users. The system model includes four entities: Users, CSP, TMC, and adversaries (\mathcal{A}), as shown in Figure 2.

- (i) *Users* own multiattribute data, and obtain personalized services in exchange of providing private data and individual preferences to CSP
- (ii) *CSP* is the back-end server of various applications in social IoT, which obtains user data through the application privilege interface and provides personalized services according to users' individual preference. Meanwhile, CSP plays static Bayesian game with adversaries and makes strategies
- (iii) *TMC* is responsible for supervising CSP's behaviors, including managing and updating CSP's reputation.

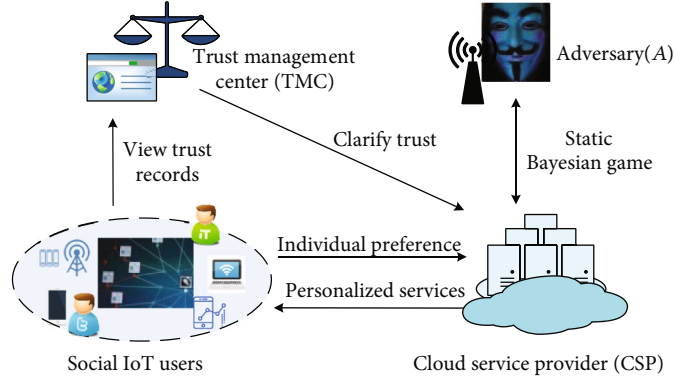


FIGURE 2: System model.

Social IoT users choose whether to trust CSP based on the reputation provided by TMC

- (iv) \mathcal{A} are malicious individuals or organizations in social IoT to obtain user private data by compromising communication links between users and CSP. Meanwhile, \mathcal{A} play strategic game with the CSP to maximize their own benefits

4.2. *Security Model.* We consider that the proposed Persian framework is implemented in a semitrusted security model [51–53]. CSP is considered as an honest-but-greedy entity. On the one hand, it is supervised by the TMC and strictly implements protocols. On the other hand, it also hopes to obtain tremendous benefits through one-off privacy trafficking. TMC is regarded as a fully-trusted entity, which is in charge of managing the reputation of CSP without possessing any private information. \mathcal{A} use rational judgment to attack adaptively. Their owned background knowledge can increase the probability of obtaining privacy. If \mathcal{A} believes that there is no benefits from launching an attack, or if the CSP’s strategy choices are indistinguishable, the Persian framework is considered secure.

5. The Construction Modules of Persian Framework

In this section, we illustrate the construction of the Persian framework in detail, including user classification, trust management, Bayesian game, and privacy measurement module.

5.1. *Overview of Persian Framework.* Above all, we explain the basic principles of the Persian framework, as illustrated in Figure 3. On the user side, users in social IoT score for each attribute according to their subjective security requirements. Having received the normalized multiattribute scores, the user classification module obtains individual privacy preferences based on fuzzy uncertainty reasoning. Before CSP provides personalized services to users, it is necessary to establish trust relationship, which is the responsibility of TMC with notarization. At the same time, there is a strategic game relationship between CSP and \mathcal{A} , and the hybrid strategy equilibrium results are used in

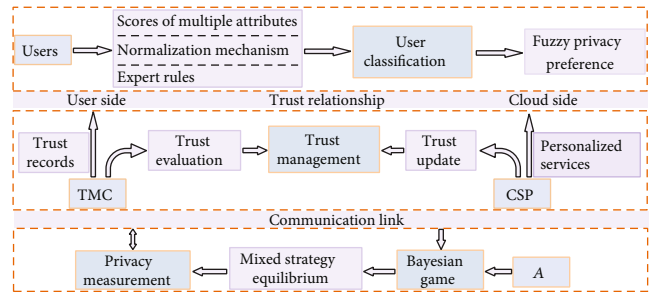


FIGURE 3: Overview of Persian framework.

privacy measurement module. Finally, TMC evaluates the services provided by CSP and performs trust update synchronously. Before each request for a service, users check the average reputation of CSP from TMC to determine whether to open data access to CSP. In essence, these entities constrain each other to provide users with secure personalized services.

5.2. *Classifying User Privacy Level.* Users in social IoT have multiple attributes of privacy data, mainly including natural attributes, social attributes, and behavioral attributes. Natural attributes are users’ own identity information, such as names, ages, typically independent from external factors. Social attributes, such as occupational status, marriage status, among others, are a feature of users’ integration into the society affected by social factors. Behavioral attributes represent users’ pursuit of individual preferences and lifestyle, such as shopping preferences, and habits. Different users have different individual security requirements for each private attribute. In this section, based on users’ individual privacy preference [54], we employ fuzzy uncertainty reasoning to classify users as the basis for the CSP to provide the corresponding level of service.

5.2.1. *Normalizing User Attribute.* In order to measure users’ security requirements for private data, we define DP, degree of privacy preference, $(DP_i = \{Name_i, Age_i, Occu_i, Marr_i, Shop_i\}, i = 1, \dots, 5)$. Through a comprehensive investigation, we use users’ natural attributes (Name and Age), social

TABLE 2: DP of user attribute.

Users	Name	Age	Occu	Marr	Shop
u_1	8	2	9	10	4
u_2	7	3	4	5	7
u_3	5	6	7	7	3
u_4	2	8	1	6	8
u_5	9	1	7	4	3

TABLE 3: NDP of user attribute.

Users	Name	Age	Occu	Marr	Shop
u_1	0.86	0.14	1	1	0.2
u_2	0.71	0.29	0.38	0.17	0.8
u_3	0.43	0.71	0.75	0.5	0
u_4	0	1	0	0.33	1
u_5	1	0	0.75	0	0

attributes (Occu and Marr), and behavioral attributes (Shop) as references. Anonymous users give a score (0–10 points) according to their subjective privacy requirements. We randomly selected five questions, as shown in Table 2.

Definition 1. Degree of privacy preference (DP) is used to measure users’ attention to private data. The lower DP is, the lower the users’ attention to data will be; otherwise, the higher the users’ attention to data will be.

Since fuzzy reasoning requires the input to be numerical data within the interval [0,1], we adopt linear function to normalize DP. Taking the j -th attribute as an example, the normalization process is shown in Formula (1). After the same treatment, the normalized DP (NDP) is shown in Table 3.

$$\text{NDP}_{ij} = \frac{\text{DP}_{ij} - \min(\text{DP}_{ij})}{\max(\text{DP}_{ij}) - \min(\text{DP}_{ij})}, i = 1, \dots, 5. \quad (1)$$

5.2.2. Fuzzy Uncertainty Reasoning. Here, we use fuzzy reasoning of Mamdani [55] type to classify users. The advanced expert rules are shown in Table 4. The input fuzzy sets (Name, Age, Occu, Marr, and Shop) are composed of “high” and “low,” which are represented by symbols “H” and “L.” The output fuzzy sets (NDP) are composed of “high,” “medium,” and “low,” which are represented by symbols “H,” “M,” and “L.” We use u_3 ’s attribute vector in Table 3 ([0.43, 0.71, 0.75, 0.5, 0]) as the fuzzy input. We can then obtain the membership degree of each input attribute to the fuzzy level through calculating the membership function, as illustrated in Table 5. Clearly, four rules are activated, namely,

- 13. NDP = L; $\min(0.67, 1, 1, 0.5, 1) = 0.5$;
- 15. NDP = M; $\min(0.67, 1, 1, 0.5, 1) = 0.5$;
- 29. NDP = M; $\min(0.33, 1, 1, 0.5, 1) = 0.33$;
- 31. NDP = H; $\min(0.33, 1, 1, 0.5, 1) = 0.33$;

The more satisfied the preceding part is, the stronger the rule will be, and the more instructive output will be. Since logic “and” is the link among the conditions in the preceding part, the strength of the four rules is determined by the “minimum value” method. Finally, we employ the central average defuzzy method to calculate the fuzzy output, and obtain the NDP’s approximation result equaling 0.462 through computing Formula (2). Therefore, u_3 obviously belongs to the M level.

$$\text{NDP} = \frac{\sum \bar{y} \cdot \mu(\bar{y})}{\sum \mu(\bar{y})} = 0.462. \quad (2)$$

In Formula (2), \bar{y} represents the maximum of fuzzy level interval, and $\mu(\bar{y})$ represents NDP’s membership value about the fuzzy level.

5.3. Trust Management. Users in social IoT submit data to CSP for personalized service, resulting in losing control of their personal data. In order to provide users with a satisfactory service experience, the trust for CSP needs to be clarified. Trust management [56] is to evaluate the target entity by referring to its historical behavior and reputation in social IoT. When social IoT users request to interact with a CSP, the service policy adopted by the CSP corresponds to a specific reputation value. CSP improves its reputation by providing good QoS. In turn, the reputation provides the basis for users to choose a CSP. We consider service behavior for J times and the reputation function of CSP as shown in the Formula.

$$\delta_j^l = \begin{cases} -10 + \lambda & \text{if } l = -1 \\ 0 & \text{if } l = 0 \\ 1 & \text{if } l = L \\ 2 & \text{if } l = M \\ 3 & \text{if } l = H \end{cases}. \quad (3)$$

In (3), $j = \{0, 1, \dots, J\}$, and $l \in \{-1, 0, L, M, H\}$. If $l = -1$, CSP actively discloses user privacy; if $l = 0$, CSP denies service. Other conditions indicate that CSP provides low, medium, and high QoS, respectively. Specially, we introduce a trust penalty factor λ , which represents the reputation penalty that CSP suffers from betraying trust. The construction function of λ is as below:

$$\lambda = - \sum_j m^{j-1} \mathbb{I}\{\delta_j^l > 0\}, \quad (4)$$

where $|\lambda| \leq m$ is satisfied in any case, and $\mathbb{I}\{\cdot\}$ is a two-value function. If the logic is true, then $\mathbb{I}\{\cdot\} = 1$; otherwise, $\mathbb{I}\{\cdot\} = 0$. Obviously, the deeper the trust relationship of the CSP betrayal is, the greater the reputation penalty will be.

Combined with the above reputation function, it makes sense to think of trust as a threshold. When the CSP’s reputation value is greater than the threshold, the user considers the CSP to be credible. To be more realistic, we consider

TABLE 4: Expert rules.

Num	Name	Age	Occu	Marr	Shop	NDP	Num	Name	Age	Occu	Marr	Shop	NDP
1	L	L	L	L	L	L	17	H	L	L	L	L	L
2	L	L	L	L	H	L	18	H	L	L	L	H	M
3	L	L	L	H	L	L	19	H	L	L	H	L	M
4	L	L	L	H	H	M	20	H	L	L	H	H	H
5	L	L	H	L	L	L	21	H	L	H	L	L	M
6	L	L	H	L	H	M	22	H	L	H	L	H	H
7	L	L	H	H	L	M	23	H	L	H	H	L	H
8	L	L	H	H	H	M	24	H	L	H	H	H	H
9	L	H	L	L	L	L	25	H	H	L	L	L	L
10	L	H	L	L	H	L	26	H	H	L	L	H	M
11	L	H	L	H	L	L	27	H	H	L	H	L	M
12	L	H	L	H	H	M	28	H	H	L	H	H	H
13	L	H	H	L	L	L	29	H	H	H	L	L	M
14	L	H	H	L	H	M	30	H	H	H	L	H	H
15	L	H	H	H	L	M	31	H	H	H	H	L	H
16	L	H	H	H	H	H	32	H	H	H	H	H	H

TABLE 5: Input attribute vector fuzzification.

Fuzzy level	Membership				
	Name	Age	Occu	Marr	Shop
L	0.67	0	0	0.5	1
H	0.33	1	1	0.5	0

the trust depth m , where the reputation of CSP is only related to the completion of the previous m services. The CSP's reputation on the j th service is shown in the Formula.

$$\sigma_j = \begin{cases} (\delta_1^l + \dots + \delta_{j-1}^l) / (j-1), & \text{if } 0 < j \leq m, \\ (\delta_{j-m}^l + \dots + \delta_{j-1}^l) / m, & \text{if } j > m. \end{cases} \quad (5)$$

When $j = 1$, we initialize σ_j to a small positive number. If $\sigma_j > \theta$, user will trust the CSP. In this way, CSP will not be willing to take the initiative to disclose user privacy for reputation. Moreover, TMC also needs to store and update the reputation of CSP for the next service. If the reputation is updated according to Formula (5), we need to calculate and store the average of m reputations. In order to reduce computation and storage overhead, we propose an incremental update strategy as shown in Formula (6). We only need to store two reputations (i.e., σ_{j-1} and δ_{j-1}^l). Another advantage is that users can only check the last time's service reputation of CSP, preventing users from completely rejecting the CSP because of occasional disclosure behaviors.

$$\sigma_j = \begin{cases} (\sigma_{j-1} \cdot (j-1) + \delta_{j-1}^l) / j, & \text{if } 1 < j < m, \\ (\sigma_{j-1} \cdot (m-1) + \delta_{j-1}^l) / m, & \text{if } j \geq m. \end{cases} \quad (6)$$

5.4. Static Bayesian Game. In addition to preventing CSP from voluntarily disclosing users' privacy, it is also necessary to resist theft attacks by potential adversaries (A). Therefore, we consider constructing a two-party static Bayesian game (SBG) [50] between the CSP and A to protect user privacy from the perspective of interests.

- (1) We consider a strategic game between the CSP and A. Participants set can be formalized as $\Gamma = \{\text{CSP}, \mathcal{A}\}$
- (2) We consider two types of adversaries, denoted as $T_A = \{\mathcal{A}_{yk}, \mathcal{A}_{nk}\}$, where \mathcal{A}_{yk} represents the adversary with background knowledge, and \mathcal{A}_{nk} represents the adversary without background knowledge. \mathcal{A} 's types set is public knowledge while CSP has only one type
- (3) We mainly consider the probability of CSP inferring A's type, then use $P_{yk}(\mathcal{A}_{yk} | \text{CSP})$ and $P_{nk}(\mathcal{A}_{nk} | \text{CSP})$ to represent the probability that CSP infers \mathcal{A} to be \mathcal{A}_{yk} and \mathcal{A}_{nk} , respectively. In this game, \mathcal{A} 's type is known only by him/herself. Thus, it is private knowledge, while joint probability $P(\text{CSP}, \mathcal{A}_{yk})$ and $P(\text{CSP}, \mathcal{A}_{nk})$ are public knowledge
- (4) The strategy set of CSP denotes $S_{\text{CSP}} = \{s_{YP}, s_{NP}\}$, where s_{YP} represents CSP providing services and s_{NP} represents CSP denying services. Note that $Y P \in \{\text{LP}, \text{MP}, \text{HP}\}$, indicating that the CSP provides low, medium, and high QoS, respectively. Meanwhile, the strategy set of \mathcal{A} denotes $S_{\mathcal{A}} = \{s_{YA}, s_{NA}\}$, where s_{YA} represents A choosing to attack and s_{NA} represents \mathcal{A} choosing not to attack. The strategy set of participants is determined before the game, regarded as public knowledge, while the strategy chosen in the game is private knowledge

TABLE 6: Benefit matrix of A and CSP.

\mathcal{A}	CSP	
	YP	NP
s_{YA}^{nk}	$(u_{YA}^{nk}, u - 1)$	$(u_c^{\mathcal{A}}, u_{NP})$
s_{YA}^{yk}	$(u_{YA}^{yk}, u - 1)$	$(u_c^{\mathcal{A}}, u_{NP})$
s_{NA}	(u_{NA}, u_{YP})	(u_{NA}, u_{NP})

- (5) The benefit function of CSP denotes $U_{CSP} = \{C, u_{NP}\}$, where u_{YP} and u_{NP} represent CSP's benefits choosing to provide services and deny services, respectively. The benefit function of \mathcal{A} denotes $U_A = \{u_{YA}^{yk}, u_{YA}^{nk}, u_{NA}\}$, where u_{YA}^{yk} and u_{YA}^{nk} represent \mathcal{A} 's benefits choosing to attack, respectively, and u_{NA} represents \mathcal{A} 's benefit choosing not to attack. Obviously, participants' benefits strongly depends on the participants' types and selected strategies

Upon received user's NDP level, the CSP provides the corresponding QoS. If NDP is H, then the CSP provides low-level services with low service quality, which comes with a low risk of user privacy disclosure, thereby meeting the high-security requirements of users. If NDP is M, then the CSP provides middle-level services. If NDP is L, then the CSP provides high-level services with high quality of service, yet with an increased possibility of user privacy disclosure. Next, we construct a game benefit matrix as shown in Table 6.

Since the CSP provides different QoS according to users' individual preferences, it will gain different reputation benefits, as shown in the Formula.

$$u_{YP} = \begin{cases} 1, & \text{if YP = LP,} \\ 2, & \text{if YP = MP,} \\ 3, & \text{if YP = HP.} \end{cases} \quad (7)$$

Particularly, the benefit of the CSP due to denial of service is $u_{NP} = 0$, and the loss of the CSP due to attack from adversary is $u_{-1} = -3$. On the other hand, \mathcal{A} 's benefit consists of three parts: basic benefit, attack cost, and extra incentive. Thus, we can determine that A's benefit is shown in the Formula.

$$\begin{cases} u_{YA}^{nk} = u_b^{\mathcal{A}} + u_c^{\mathcal{A}} + u_{YP}^2 \cdot \mu, \\ u_{YA}^{yk} = u_{YA}^{nk} \cdot (1 + k). \end{cases} \quad (8)$$

In (8), $u_b^{\mathcal{A}}$ represents \mathcal{A} 's basic benefit, $u_c^{\mathcal{A}}$ represents \mathcal{A} 's attack cost, and the other represents extra incentive refer to benefit factor μ . This means that the higher the QoS provided by CSP, the higher the data quality submitted by users, and the greater the benefits gained from A successfully attacking. When \mathcal{A} chooses not to attack, $u_{NA} = 0$. Additionally, k is

regarded as background factor, and is used to increase \mathcal{A} 's benefit.

In order to facilitate the analysis of the incomplete information game, we use the Harsanyi transformation to introduce a virtual participant "nature" (N). N randomly selects both participants' types. $P\langle \text{CSP}, \mathcal{A}_{yk} \rangle$ and $P\langle \text{CSP}, \mathcal{A}_{nk} \rangle$ are public knowledge, where $P\langle \text{CSP}, \mathcal{A}_{yk} \rangle + P\langle \text{CSP}, \mathcal{A}_{nk} \rangle = 1$. $P_{yk}\langle \mathcal{A}_{yk} | \text{CSP} \rangle$ and $P_{nk}\langle \mathcal{A}_{nk} | \text{CSP} \rangle$ represent the probabilities of the CSP inferring type \mathcal{A} , respectively, which can be obtained by Bayesian formula:

$$\begin{cases} P_{yk}\langle \mathcal{A}_{yk} | \text{CSP} \rangle = P\langle \text{CSP}, \mathcal{A}_{yk} \rangle / P\langle \text{CSP} \rangle = P\langle \text{CSP}, \mathcal{A}_{yk} \rangle, \\ P_{nk}\langle \mathcal{A}_{nk} | \text{CSP} \rangle = P\langle \text{CSP}, \mathcal{A}_{nk} \rangle / P\langle \text{CSP} \rangle = P\langle \text{CSP}, \mathcal{A}_{nk} \rangle. \end{cases} \quad (9)$$

For simplicity, we use P_{yk} and P_{nk} to replace $P_{yk}\langle \mathcal{A}_{yk} | \text{CSP} \rangle$ and $P_{nk}\langle \mathcal{A}_{nk} | \text{CSP} \rangle$, respectively. Then, we use P_{YP} and P_{NP} to represent the probability of CSP choosing YP strategy and NP strategy (i.e., $P_{YP} + P_{NP} = 1$), respectively. Furthermore, we use P_{YA} and P_{NA} to represent the probability of A choosing YA strategy and NA strategy, respectively, (i.e., $P_{YA} + P_{NA} = 1$). Next, we calculate the benefits of the CSP choosing YP and NP strategies, as shown in the Formula.

$$\begin{cases} u_{YP} = P_{YA} \cdot u_{-1} + P_{NA} \cdot u_{YP}, \\ u_{NP} = 0. \end{cases} \quad (10)$$

The benefits of \mathcal{A} choosing YA strategy and NA strategy are shown in the Formula.

$$\begin{cases} u_{YA} = [P_{YP} \cdot u_{YA}^{yk} + P_{NP} \cdot u_c^{\mathcal{A}}] \cdot P_{yk} + [P_{YP} \cdot u_{YA}^{nk} + P_{NP} \cdot u_c^{\mathcal{A}}] \cdot P_{nk}, \\ u_{NA} = 0. \end{cases} \quad (11)$$

We can obtain the CSP's and \mathcal{A} 's benefit, respectively, from the Formula.

$$\begin{cases} U_{CSP}(P_{YA}^*, P_{YP}) = u_{YP} \cdot P_{YP} + u_{NP} \cdot P_{NP}, \\ U_{\mathcal{A}}(P_{YA}, P_{YP}^*) = u_{YA} \cdot P_{YA} + u_{NA} \cdot P_{NA}. \end{cases} \quad (12)$$

In the static Bayesian equilibrium state, CSP's benefit function U_{CSP} can reach the maximum regardless of what \mathcal{A} chooses. Also, \mathcal{A} 's benefit function $U_{\mathcal{A}}$ can reach the maximum regardless of what CSP chooses. Essentially, it means that the strategies of participants are indistinguishable and can be solved by simultaneous equations.

$$\begin{cases} \frac{\partial U_{CSP}(P_{SA}^*, P_{MP})}{\partial P_{MP}} = 0, \\ \frac{\partial U_{\mathcal{A}}(P_{SA}, P_{MP}^*)}{\partial P_{SA}} = 0. \end{cases} \quad (13)$$

Therefore, we can obtain the mixed strategy Bayesian equilibrium (P_{YA}^*, P_{YP}^*) , as shown in the Formula.

$$\begin{cases} P_{YA}^* = \frac{u_{YP}}{u_{YP} - u_{-1}}, \\ P_{YP}^* = \frac{-u_c^{\mathcal{A}}}{(u_{YA}^{yk} - u_{YA}^{nk}) \cdot P_{yk} + (u_{YA}^{mk} - u_c^{\mathcal{A}})}. \end{cases} \quad (14)$$

5.5. Privacy Disclosure Measurement. In social IoT, users exchange personalized services from the CSP by providing private data, which will inevitably have the risk of being leaked over communication links. Shannon information entropy [57] is used to measure privacy disclosure, as shown in the Formula below.

$$H(X) = - \sum_{i=1}^n P_i \cdot \log(P_i), \quad (15)$$

where $0 \leq P_i \leq 1$, $\sum_{i=1}^n P_i = 1$.

From A's point of view, the probability P_{YP} and P_{NP} can be inferred. As described in Formula (16), the greater $H(\mathcal{A})$ is, the closer P_{YP} and P_{NP} are, the higher the indistinguishability of A to CSP's strategy, and the lower degree of privacy disclosure. Otherwise, the higher is the degree of privacy disclosure. For instance, if $P_{YP} = 0.5$, the information entropy is 1, indicating that A is completely confused about the service decision of the CSP.

$$H(\mathcal{A}) = -P_{YP} \cdot \log(P_{YP}) - P_{NP} \cdot \log(P_{NP}). \quad (16)$$

In social IoT, a single game obviously cannot satisfy the user requirement. Therefore, we consider the finite static Bayesian game for J times. Similarly, the privacy disclosure measurement in a long term can be described by Formula (17). It can be used to evaluate the privacy disclosure status of J times of service.

$$H^J(\mathcal{A}) = - \frac{1}{J} \sum_{j=1}^J P_{YP}^j \log(P_{YP}^j) - P_{NP}^j \log(P_{NP}^j). \quad (17)$$

6. Experiment and Evaluation

In order to better illustrate the feasibility of the Persian framework, we consider that its performance is influenced by three factors: user classification, trust assessment, and privacy disclosure measurement. The experiments were carried out on a workstation with a 3.30 GHz quad-core processor, 8GB memory, and Windows 7 64-bit operating system to simulate and analyze on the Matlab R2016a platform.

6.1. User Classification. In the user classification module, Fuzzy Toolbox [58] is used to conduct fuzzy uncertainty reasoning for users, and the result is used to verify the theoretical calculation. Assuming that the input user is u_3 in Table 2, we synthesize the rule constraints of all input attributes. As shown in Figure 4, rules (13)(15)(29)(31) are

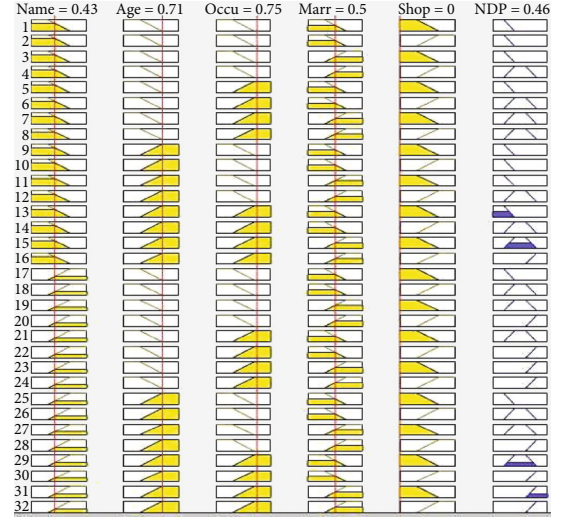


FIGURE 4: Simulation result of Fuzzy Toolbox.

activated, and the output NDP is 0.46. It can be seen that u_3 belongs to the M level, which is consistent with the theoretical calculation.

6.2. Trust Assessment. We evaluate the CSP's trust based on the QoS provided by CSP for J times, and update the CSP's visual trust to users. Users in Social IoT need to check the reputation of the CSP prior to submitting private data to the CSP. Only when CSP's reputation is greater than the trust threshold (i.e., $\theta = 0$), users are willing to trust the CSP and share their privacy. Based on the individual preferences of 30 users (i.e., $J = 30$), we quantified the CSP's active disclosure behavior and indirectly clarified the trust of the CSP.

Figure 5(a) shows that the CSP provides three levels of QoS according to users' individual privacy preferences, and then obtains three ratings of reputation, namely, 1, 2, and 3. It is worth noting that the CSP will be subject to severe reputational penalties if it voluntarily discloses users' privacy in the process of providing services. The CSP chooses to disclose privacy when $j = 3$, and it loses the reputation of 12 units. The attendant consequences are disastrous for the CSP, resulting in a significant decline in service delivery rates. On the other hand, we consider the impact of trust depth on the reputation of the CSP. Figure 5(b) shows that dishonest behaviors of the CSP will lead to the decline of the visible reputation of multiround services. As the depth of trust relationship (i.e., m) increases, the spread of the reputation penalty becomes more serious. Regardless of the situation, there is no reason for the curve in SBG for J times. The CSP is used to actively disclose user privacy in order to maintain visible reputation with users.

6.3. Privacy Disclosure Measurement. It is certain that the CSP provides different levels of QoS according to individual preferences. As a result, the behavioral strategy of A and the CSP may change driven by interests. Figure 6(a) shows the relationship between Nash equilibrium and QoS levels. With the improvement of service quality, the probability of attack is gradually increased because of the temptation of

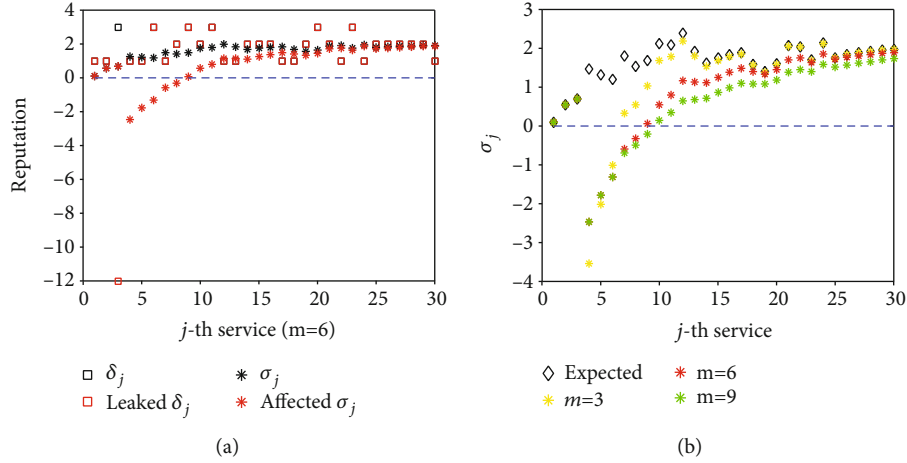


FIGURE 5: Trust assessment results. (a) The impact of privacy disclosure on reputation; (b) the impact of trust depth on reputation.

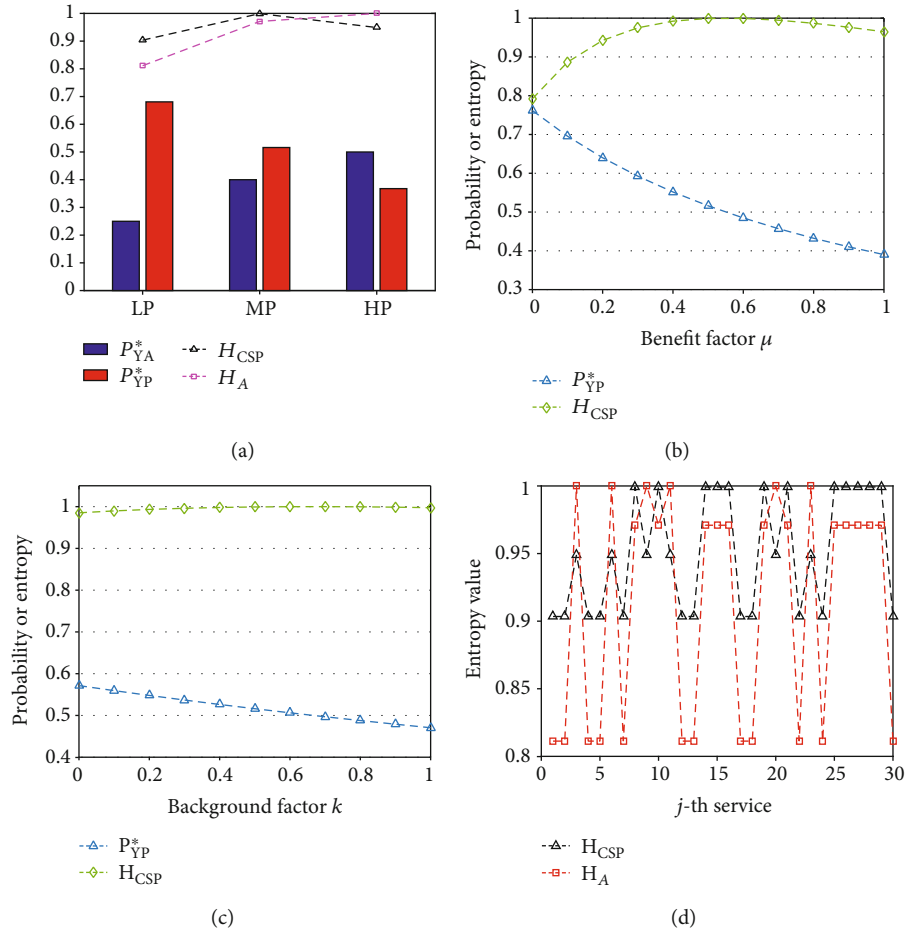


FIGURE 6: Nash equilibrium results. (a) The relationship between Nash equilibrium and QoS levels; (b) the relationship between Nash equilibrium and benefit factor; (c) the relationship between Nash equilibrium and background factor; (d) Nash equilibrium.

high data quality, and the CSP tends to choose denying service due to potential attacks. Despite the fact that the CSP provides high QoS, the result of information entropy has declined slightly, just below 1. This suggests that A is still confused about the decision-making of the CSP, and the risk

of user privacy disclosure remains at a relatively low level. Further, we explain the relationship between Nash equilibrium and two internal incentive factors μ and k . From Figure 6(b), the attack benefit u_{YA} of adversaries increases with μ , and the probability P_{YP}^* of CSP to provide the service

gradually decreases. When P_{yp}^* is close to 0.5, the service information entropy H (CSP) reaches the maximum, and the privacy protection level reaches its peak. Figure 6(c) shows a similar trend in relationships. The probability of a successful attack by adversaries increases with k , and as a result, CSP tends to refuse to provide services. Likewise, the intensity of privacy protection reaches maximum while k gets close to 0.5.

Additionally, we observe the Nash equilibrium change in the CSP service delivery for J times. Figure 6(d) shows that the information entropy is maintained at a high level. According to Formula (17), we can calculate the average privacy disclosure of 0.9509. As long as the CSP does not actively disclose the user's privacy, the confusion of \mathcal{A} about the CSP's decision will not be decreased. Also, we have clarified the trust in the CSP in Section 5.3. Based on the above, the Persian framework can effectively provide personalized services while keeping privacy disclosure to a minimum level.

7. Conclusion

There are frequent occurrences of user privacy disclosure in social IoT, drawing wide attention in academia and the industry. A few achievements have been acquired, but a number of key techniques are still in need. On the one hand, users have to share their privacy to the CSP to exchange application privileges, so as to enjoy personalized services. On the other hand, users are reluctant to disclose their privacy. Find a satisfactory balance between QoS and privacy protection under the premise of ensuring personalized services is the main contribution of this paper. We proposed a privacy-preserving personalized service framework (Persian) through a static Bayesian game. In this framework, users independently infer their privacy preferences combined with offline fuzzy reasoning. The trust of the CSP is supervised by TMC to ensure normal service operations. The CSP provides users with personalized service according to users' individual preferences. Furthermore, we employ the game mixing strategy equilibrium to achieve privacy protection from the perspective of interests. Meanwhile, we measure privacy disclosure by using information entropy under the proposed framework.

The future work is to further expand the fuzzy reasoning with neural network and consider additional user attributes. We will also consider more types of adversaries and constantly optimize the proposed model to achieve better comprehensiveness and efficiency for privacy protection.

Data Availability

The data used in this paper comes from the comprehensive questionnaire investigation.

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 grants 61872088, 61872090, U1905211, and 61702105, in part by the Natural Science Foundation of Fujian Province under grant 2019J01276, in part by the Guizhou Provincial Key Laboratory of Public Big Data Research Fund under grant 2019BDKFJJ004, in part by the Science and Technology Research Program of Chongqing Municipal Education Commission under grant KJQN201801316, in part by the Innovation and Entrepreneurship Demonstration Team of Yingcai Program of Chongqing under grant CQYC201903167, and in part by Scientific and Technological Research Program of Chongqing Municipal Education Commission under grant KJZD-K201901301.

References

- [1] J. Ni, X. Lin, and X. S. Shen, "Toward edge-assisted internet of things: from security and efficiency perspectives," *IEEE Network*, vol. 33, no. 2, pp. 50–57, 2019.
- [2] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, "Channel state information prediction for 5g wireless communications: a deep learning approach," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 227–236, 2020.
- [3] G. Liu, Q. Yang, H. Wang, and A. X. Liu, "Three-valued subjective logic: a model for trust assessment in online social networks," *IEEE Transactions on Dependable and Secure Computing*, p. 1, 2019.
- [4] D. Wu, B. Liu, Q. Yang, and R. Wang, "Social-aware cooperative caching mechanism in mobile social networks," *Journal of Network and Computer Applications*, vol. 149, p. 102457, 2020.
- [5] H. Wang, M. Hempel, D. Peng, W. Wang, H. Sharif, and H. H. Chen, "Index-based selective audio encryption for wireless multimedia sensor networks," *IEEE Transactions on Multimedia*, vol. 12, no. 3, pp. 215–223, 2010.
- [6] J. Xiong, L. Chen, M. Z. Bhuiyan et al., "A secure data deletion scheme for iot devices through key derivation encryption and data analysis," *Future Generation Computer Systems (Early Access)*, vol. 111, pp. 741–753, 2020.
- [7] D. Wu, J. Yan, H. Wang, and R. Wang, "User-centric edge sharing mechanism in software-defined ultra-dense networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 7, pp. 1531–1541, 2020.
- [8] D. Wu, X. Han, Z. Yang, and R. Wang, "Exploiting transfer learning for emotion recognition under cloud-edge-client collaborations," *IEEE Journal on Selected Areas in Communications*, p. 1, 2020.
- [9] X. Meng, *China privacy risk index analysis report*, 2020, <http://www.tiuchina.com/yjbg/346.html>.
- [10] H. Wang, C. Gao, Y. Li, Z.-L. Zhang, and D. Jin, "Revealing physical world privacy leakage by cyberspace cookie logs," *IEEE Transactions on Network and Service Management*, p. 1, 2020.
- [11] D. Keküllüoğlu, W. Magdy, and K. Vaniea, "Analysing privacy leakage of life events on twitter," in *12th ACM Conference on Web Science*, pp. 287–294, Southampton, United Kingdom, 2020.

- [12] R. Sobers, "107 Must-Know Data Breach Statistics for 2020," 2020, http://www.sogou.com/link?url=hedJjaC291M8s0HmWVMlqTtIfarcvd6pInEXJCF-KNim9pVlRx94EYkHAapUdxCOVW_atjSnNqI.&query=2020+data+leakage+survey.
- [13] M. S. Roopa, S. Pattar, R. Buyya, K. R. Venugopal, S. S. Iyengar, and L. M. Patnaik, "Social internet of things (siot): foundations, thrust areas, systematic review and future directions," *Computer Communications*, vol. 139, pp. 32–57, 2019.
- [14] J. Xiong, J. Ren, L. Chen et al., "Enhancing privacy and availability for data clustering in intelligent electrical service of iot," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1530–1540, 2019.
- [15] Z. Lin and L. Dong, "Clarifying trust in social internet of things," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 2, pp. 234–248, 2018.
- [16] T. Cheng, G. Liu, Q. Yang, and J. Sun, "Trust assessment in vehicular social network based on three-valued subjective logic," *IEEE Transactions on Multimedia*, vol. 21, no. 3, pp. 652–663, 2019.
- [17] Q. Jiang, Z. Chen, J. Ma, X. Ma, J. Shen, and D. Wu, "Optimized fuzzy commitment based key agreement protocol for wireless body area network," *IEEE Transactions on Emerging Topics in Computing*, p. 1, 2019.
- [18] Q. Jiang, N. Zhang, J. Ni, J. Ma, X. Ma, and K. K. Choo, "Unified biometric privacy preserving three-factor authentication and key agreement for cloud-assisted autonomous vehicles," *IEEE Transactions on Vehicular Technology*, p. 1, 2020.
- [19] J. Xiong, R. Ma, L. Chen, Y. Tian, L. Lin, and B. Jin, "Achieving incentive, security, and scalable privacy protection in mobile crowdsensing services," *Wireless Communications and Mobile Computing*, vol. 2018, 12 pages, 2018.
- [20] C. Liu, Y. Tian, J. Xiong, Y. Lu, Q. Li, and C. Peng, "Towards attack and defense views to k-anonymous using information theory approach," *IEEE Access*, vol. 7, pp. 156025–156032, 2019.
- [21] J. Xiong, M. Zhao, M. Bhuiyan, L. Chen, and Y. Tian, "An ai-enabled threeparty game framework for guaranteed data privacy in mobile edge crowdsensing of iot," *IEEE Transactions on Industrial Informatics*, p. 1, 2019.
- [22] C. Dwork, *Differential privacy*, Springer, NewYork, NY, USA, 2011.
- [23] J. Xiong, R. Ma, L. Chen et al., "A personalized privacy protection framework for mobile crowdsensing in iiot," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4231–4241, 2020.
- [24] H. Hu, G.-J. Ahn, and J. Jorgensen, "Multiparty access control for online social networks: model and mechanisms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 7, pp. 1614–1627, 2012.
- [25] Q. Li, H. Zhu, J. Xiong, R. Mo, Z. Ying, and H. Wang, "Fine-grained multiauthority access control in iot-enabled mhealth," *Annals of Telecommunications*, vol. 74, no. 7-8, pp. 389–400, 2019.
- [26] M. Frustaci, P. Pace, G. Aloï, and G. Fortino, "Evaluating critical security issues of the iot world: present and future challenges," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2483–2495, 2018.
- [27] L. Zhang, X. Zhu, X. Han, and J. Ma, "Differentially privacy-preserving social iot," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Xi'an, China, China, 2019.
- [28] Q. Li, J. Ma, R. Li, J. Xiong, and X. Liu, "Provably secure unbounded multi-authority ciphertext-policy attribute-based encryption," *Security and Communication Networks*, vol. 8, no. 18, pp. 4098–4109, 2015.
- [29] X. Wu, T. Wu, M. Khan, Q. Ni, and W. Dou, "Game theory based correlated privacy preserving analysis in big data," in *IEEE Transactions on Big Data*, 2017.
- [30] C.-G. Liu, I.-H. Liu, W.-S. Yao, and J. S. Li, "K-anonymity against neighborhood attacks in weighted social networks," *Security and Communication Networks*, vol. 8, no. 18, pp. 3864–3882, 2015.
- [31] Y. Xie and M. Zheng, "A differentiated anonymity algorithm for social network privacy preservation," *Algorithms*, vol. 9, no. 4, p. 85, 2016.
- [32] Z. G. Chen, H. S. Kang, S. N. Yin, and S. R. Kim, "An efficient privacy protection in mobility social network services with novel clustering-based anonymization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, Article ID 275, 2016.
- [33] X. Li, J. Yang, Z. Sun, and J. Zhang, "Differential privacy for edge weights in social networks," *Security and Communication Networks*, vol. 2017, 10 pages, 2017.
- [34] Q. Wang, Y. Zhang, X. Lu, Z. Wang, Z. Qin, and K. Ren, "Real-time and spatiotemporal crowd-sourced social network data publishing with differential privacy," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 591–606, 2016.
- [35] H. Huang, D. Zhang, F. Xiao, K. Wang, J. Gu, and R. Wang, "Privacy-preserving approach pbcn in social network with differential privacy," *IEEE Transactions on Network and Service Management*, vol. 17, no. 2, pp. 931–945, 2020.
- [36] S. Jahid, P. Mittal, and N. Borisov, "Easier: encryption-based access control in social networks with efficient revocation," *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, pp. 411–415, 2011.
- [37] Z. Cai, Z. He, X. Guan, and Y. Li, "Collective data-sanitization for preventing sensitive information inference attacks in social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 577–590, 2018.
- [38] Y. Cai, S. Zhang, H. Xia, Y. Fan, and H. Zhang, "A privacy-preserving scheme for interactive messaging over online social networks," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6817–6827, 2020.
- [39] V. Sharma, I. You, D. N. K. Jayakody, and M. Atiquzzaman, "Cooperative trust relaying and privacy preservation via edge-crowdsourcing in social internet of things," *Future Generation Computer Systems*, vol. 92, pp. 758–776, 2019.
- [40] X. Jin, N. Pissinou, S. Pumpichet, C. A. Kamhoua, and K. Kwiat, "Modeling cooperative, selfish and malicious behaviors for trajectory privacy preservation using bayesian game theory," in *38th Annual IEEE Conference on Local Computer Networks*, pp. 835–842, Sydney, NSW, Australia, 2013.
- [41] H. Hu, G. J. Ahn, Z. Zhao, and D. Yang, "Game theoretic analysis of multiparty access control in online social networks," *Proceedings of the 19th ACM symposium on Access control models and technologies*, pp. 93–102, 2014.
- [42] F. Shan, H. Li, and H. Zhu, "Game theory based forwarding control method for social network," *Journal on Communications*, vol. 39, no. 3, pp. 172–180, 2018.

- [43] J. Xiong, M. S. Wang, and Y. L. Tian, "Research progress on privacy measurement for cloud data," *Journal of software*, vol. 29, no. 7, pp. 1963–1980, 2018.
- [44] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," *International Workshop on Privacy Enhancing Technologies*, pp. 41–53, 2002.
- [45] Z. Lin, M. Hewett, and R. B. Altman, "Using binning to maintain confidentiality of medical data," *Proceedings of the AMIA Symposium*, p. 454, 2002.
- [46] C. Diaz, C. Troncoso, and G. Danezis, "Does additional information always reduce anonymity?," *Proceedings of the 2007 ACM workshop on Privacy in electronic society*, pp. 72–75, 2007.
- [47] T. Chen, A. Chaabane, P. U. Tournoux, M. A. Kaafar, and R. Boreli, "How much is too much? Leveraging ads audience estimation to evaluate public profile uniqueness," in *International Symposium on Privacy Enhancing Technologies Symposium*, pp. 225–244, Springer, Berlin, Heidelberg, 2013.
- [48] J. Xiong, X. Chen, Q. Yang, L. Chen, and Z. Yao, "A task-oriented user selection incentive mechanism in edge-aided mobile crowdsensing," *IEEE Transactions on Network Science and Engineering*, p. 1, 2019.
- [49] N. Werro, *Fuzzy Classification of Online Customers*, Springer, New York, NY, USA, 2015.
- [50] D. Fudenberg and J. Tirole, *Game Theory*, Massachusetts, Cambridge, 1991.
- [51] J. Xiong, Y. Zhang, L. Lin, J. Shen, X. Li, and M. Lin, "ms-posw: a multiserver aided proof of shared ownership scheme for secure deduplication in cloud," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 3, 2019.
- [52] X. Liu, R. H. Deng, P. Wu, and Y. Yang, "Lightning-fast and privacy-preserving outsourced computation in the cloud," *Cybersecurity*, vol. 3, no. 1, p. 17, 2020.
- [53] J. Xiong, R. Bi, M. Zhao, J. Guo, and Q. Yang, "Edge-assisted privacy-preserving raw data sharing framework for connected autonomous vehicles," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 24–30, 2020.
- [54] P. Zhang, C. Peng, and C. Hao, "Privacy protection model and privacy metric methods based on privacy preference," *Computer Science*, vol. 45, no. 6, pp. 130–134, 2018.
- [55] M. Blej and M. Azizi, "Comparison of mamdani-type and sugeno-type fuzzy inference systems for fuzzy real time scheduling," *International Journal of Applied Engineering Research*, vol. 11, no. 22, pp. 11071–11075, 2016.
- [56] A. Mehmood, I. Natgunanathan, Y. Xiang, G. Hua, and S. Guo, "Protection of big data privacy," *IEEE Access*, vol. 4, pp. 1821–1834, 2016.
- [57] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Chichester, West Sussex, United Kingdom, 2012.
- [58] B. Narasimhan and A. Malathi, "A fuzzy logic system with attribute ranking technique for risk-level classification of cahd in female diabetic patients," in *2014 International Conference on Intelligent Computing Applications*, pp. 179–183, Coimbatore, India, 2014.

Research Article

Social-Aware Task Allocation in Mobile Crowd Sensing

Weiping Zhu ¹, Wenzhong Guo ^{1,2,3} and Zhiyong Yu^{1,2,3}

¹College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350116, China

²Fujian Provincial Key Laboratory of Networking Computing and Intelligent Information Processing, Fuzhou University, Fuzhou 350116, China

³Key Laboratory of Spatial Data Mining and Information Sharing, Ministry of Education, Fuzhou 350003, China

Correspondence should be addressed to Wenzhong Guo; guowenzhong@fzu.edu.cn

Received 26 March 2020; Revised 4 September 2020; Accepted 13 September 2020; Published 14 October 2020

Academic Editor: Kim-Kwang Raymond Choo

Copyright © 2020 Weiping 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.

Task allocation is a significant issue in crowd sensing, which trades off the data quality and sensing cost. Existing task allocation works are based on the assumption that there is plenty of users available in the candidate pool. However, for some specific applications, there may be only a few candidate users, resulting in the poor completion of tasks. To tackle this problem, in this paper, we investigate the task allocation problem with the assistance of social networks. We select a subset of users; if a user can not complete the task, he can propagate the task information to his friends. The object of this problem is to maximize the expected number of completed tasks. We prove that the task allocation problem is an NP-hard and submodular problem and then propose a native greedy selection (NGS) algorithm, which selects the user with maximum margin gain in each round. To improve the efficiency of the NGS algorithm, we further propose a fast greedy selection algorithm (FGS), which selects the user who can actually complete the maximum number of tasks. Experimental results show that although FGS gets slightly worse results in terms of the expected number of completed tasks, it can greatly reduce the running time of seed selection.

1. Introduction

In recent years, with the popularity of mobile devices and the development of various communication technologies, mobile crowd sensing (MCS) [1–3] has emerged as a promising sensing paradigm, in which mobile users leverage their carry-on devices to collect and upload the sensing data. By collecting and analyzing sensing data from a group of users, the platform can provide ubiquitous services. Due to its advantages such as low deployment cost and wide spatio-temporal coverage, numerous practical MCS applications have been rolled out in many areas, such as environmental monitoring [4], smart transportation [5], and emergency alarming [6].

A mobile crowd sensing process can be divided into four stages: task creation, task allocation, task execution, and data integration [7]. Particularly, task allocation is an important stage, which trades off the sensing quality and sensing cost [3, 8]. Different from traditional wireless sensor networks [9], task allocation in MCS should select eligible

users to complete tasks. So far, there have been numerous related works on task allocation [10–12]. Among these works, the platform directly selects users from a large pool of candidate users to maximize the sensing quality under limited cost or minimize the sensing cost while guaranteeing certain data quality.

With the development of MCS, there may be plenty of applications using this scheme to collect data. For some specific applications, there may not be enough available candidate users. The reasons can be listed as follows: on the one hand, in the early stage of a new application, there are a few registered users available to allocate tasks. On the other hand, due to the privacy concern, users do not open the location service, so the platform cannot get their locations. This situation makes the platform hard to recruit enough users, and the existing candidate users can complete only a small number of tasks, resulting in a low completion ratio of tasks.

Fortunately, with the development of social networks, “word-of-mouth” effect has played an important role in propagating and sharing information, giving an alternative

way to allocate mobile sensing tasks. The social relationship is built if there is a connection between two users. If users know the service is requested by his friends, he may be glad to participate in the sensing tasks. There are several advantages of allocating tasks with the assistance of social networks. From the perspective of the platform, it only needs to recruit a few seeds to complete or propagate the task information, which can reduce the sensing cost. Furthermore, users may be glad to help their friends to complete sensing tasks due to their friendship. So the sensing enthusiasm will be increased compared to the traditional “push” task allocation model. From the perspective of users who have not registered in the platform, they do not need to share their real-time locations, thus reducing the risk of privacy leakage. Inspired by this, this paper attempts to allocate the sensing tasks with the assistance of social networks. Concretely, we select a small number of users to allocate the tasks. If the users can reach the target location at the required time, then, the task can be completed by the selected users. Otherwise, the selected users propagate the task information to his friends, who may be in the target location at the required time, different from [13, 14], which select seeds to propagate the task information in the social network to maximize the spatiotemporal coverage of an area. This paper considers that seeds propagate task information to their friends only when the task is not completed.

Considering some location-based tasks in the platform, for example, monitoring the traffic dynamic information of a specific road, reporting the air pollution of the target location, this type of task requires users to arrive in the target location to collect data. However, the platform only knows a little part of users’ location information. To allocate these tasks, an effective way for the platform is selecting some users as seeds to complete or propagate these tasks. These seeds not only complete these tasks that they can complete but also propagate the task information that they cannot complete to their friends and expect their friends can help to complete the sensing tasks. For the latter case, the completion probability of these tasks depends on the probability of friends in the target location and the probability that the seeds propagate the task information to him.

The task allocation problem mentioned above is similar to the location aware influence maximization problem in social networks to some extent, which is to select k users in the target location to maximize the expected number of influenced users. However, there are mainly two differences between these two problems: (1) location constraints of seeds and their influenced users. In the influence maximization problem, both seeds and influenced users should be in the target location, while in the task allocation problem, seeds are not necessarily in the target location but influenced users should be in the target location (in case only propagate once). (2) performance quantification of seeds. In the influence maximization problem, a seed is more valuable if he can influence more additional users, while in the task allocation problem, a seed is more valuable if he can complete more additional tasks by influencing users or by himself.

According to the analysis above, the key challenge is how to identify the seeds that can maximize the expected number

of completed tasks. Compared to the traditional task allocation mode in MCS, we consider not only the actual number of completed tasks but also the propagation ability to tasks that they cannot complete. With the object and challenges mentioned above, the main contributions of this work can be concluded as follows:

- (1) We formulate a social-aware task allocation problem (SATA), in which the platform allocates the tasks considering the social influence of users. If the selected user can not complete the task, he propagates the task information to his friends and expects to get sensing data from friends. The SATA problem is aimed at selecting k users as seeds to maximize the expected number of completed tasks
- (2) We prove that the problem is an NP-hard and sub-modular problem. Then, we propose a native greedy-based selection algorithm (NGS) to solve the problem. The NGS algorithm selects the user with the largest margin gain in each round
- (3) Considering the low efficiency of NGS, we further propose the fast greedy selection algorithm (FGS). FGS divides the seed selection process into two stages. In the first stage, FGS gives the priority to the user who can complete the maximum number of remaining tasks. For the second stage, in which no one can complete any task, FGS selects the user with maximum propagation performance from candidate users
- (4) We conduct extensive experiments using two real-world datasets, which contain users’ social relationship and mobility traces. The experimental results show that the performance of the FGS algorithm is slightly worse than that of the NGS algorithm. However, the FGS algorithm runs more efficiently than the NGS algorithm

The remainder of this paper is organized as follows. In Section 2, the related works about task allocation are presented. In Section 3, the problem of SATA is described in detail. Then, we analyze the problem and introduce our proposed algorithms in Section 4. In Section 5, we evaluate the performance of proposed algorithms and baseline algorithms. Finally, the conclusions are presented in Section 6.

2. Related Work

Task allocation has become an important problem in MCS and drawn considerable attention from researchers. There have been numerous studies on task allocation, which can be divided into the following two types.

2.1. Single Task Allocation. In the early state of MCS, several systems have been designed for single task allocation. For example, in [15, 16], Reddy et al. considered the location, time constraints, and habits of users and proposed a coverage-based framework to select proper users to maximize spatial coverage. Zhang et al. proposed a task allocation framework, which is aimed at maximizing the coverage

quality under budget constraint [17]. Wang et al. [18] predicted the mobility of vehicles and selected participants to minimize the overall recruitment cost. Another work [19] defined a new coverage metric, namely, “t-sweep k-coverage” and proposed two methods to select the smallest set of candidate users to satisfy the predefined requirements.

2.2. Multiple Task Allocation. With the development of crowdsensing, there may be many tasks to be allocated in the platform. Researchers made efforts to study the allocation for multitasks. Though the optimization and constraints may be similar to single-task allocation, several factors should be considered to deal with multiple task allocation.

Firstly, from the perspective of spatiotemporal aspects, different tasks vary in spatial and temporal granularity and other properties. Based on this consideration, Wang et al. formulated a multitask allocation problem, which considered heterogeneous spatial and temporal granularity. Users can complete multiple tasks without changing their locations. The authors proposed a two-stage allocation algorithm to solve this problem [11]. Li et al. [20] proposed an online task allocation algorithm for dynamic heterogeneous tasks to minimize the number of users while guaranteeing a certain level of coverage.

Secondly, from the perspective of task properties, our previous work [21] considered the type of sensing tasks and heterogeneity of users’ sensing devices and proposed a Particle Swarm Optimization (PSO) algorithm to select eligible users to maximize the number of completed tasks. Inversely, a recent work [12] considered that different tasks can share the same data property under some special circumstances and proposed a triple-layer task allocation framework, which considers not only the temporal and spatial correction but also the data property of sensing data. Jiang et al. [22] considered the skills that tasks required and allocated tasks to a natural user group, in which members can cooperate to complete the complex task. If the users in an assigned group do not possess all required skills, they can cooperate with other contextual groups. The authors selected a principal group with the maximum self-crowdsourcing value and allocated the tasks to the group. Then, the authors proposed two greedy-based approaches to select an assistant group according to the circumstance of principal group. Wang et al. [23] combined the content information with context information of tasks and unified all the factors together to measure the preference score of a user to a task. Then, the user with the maximum matching probability is selected to complete tasks.

Thirdly, from the perspective of global optimization, the optimal total quality of tasks does not always guarantee the optimal quality for every task. To solve this problem, Wang et al. [24] considered the task-specific sensing quality thresholds and proposed a descent greedy approach to select a set of users to complete tasks. Considering the total time consumption for multiple tasks depends on the time consumption of the last completed task. The authors in the study [25] considered the time-sensitive tasks and proposed a cooperation scheme to minimize the maximum completion time of tasks. Ni et al. [26] investigated the dependency-aware task allocation problem with the consideration of skills, moving distances, and deadlines. To allocate these tasks efficiently, the

authors grouped these tasks into associative task sets and proposed two algorithms to solve this problem. Song et al. [27] considered that the existing task matching strategies may cause a misaligned task coverage problem, which indicates that some popular tasks can find enough users but the less popular task cannot assigned successfully. To solve this problem, the authors proposed the *cTaskMat* framework. It learns users’ task preferences and their attitudes on task attributes. Then, it migrates certain qualified users to less popular tasks for increasing task coverage and reducing the sensing cost.

These works assumed that there are enough candidate users and the platform can push the task information to the eligible users. However, for those new applications, there have not been many users registered in the platform, resulting in the poor completion ratio of sensing tasks. Thus, a new task allocation scheme is urgently needed to solve this dilemma.

To solve this problem, Wang et al. [13] proposed a task allocation framework, which propagates the task information leveraging the social network to maximize the wide-area coverage. Lu and Zhu [14] proposed a hybrid user recruitment framework, which divides the recruitment process into two phases. In the offline phase, it recruits users to propagate task information in communication and social networks. In the online phase, it incentivizes users to move to uncovered subareas and fulfil tasks based on subarea clustering. However, in these two studies, the object is to maximize the coverage of the whole area, which can be divided into several fine-grained subareas. The selected seeds and their influenced users will cover all subareas that they pass even those subareas already covered, which may cause data redundancy. In this paper, seeds propagate the task information to their friends only when the task is not completed. Thus, it is promising to reduce the data redundancy. Peng et al. [28] selected eligible users for crowdsourcing tasks based on social relation cognition. Different from this work, we considered the propagation of location-based crowd sensing tasks, which is common in the real life and agent to be explored.

3. System Model and Problem Formulation

In this section, we present the overview of the crowdsensing system with limited users and then formulate the problem of SATA, in which platform selects mobile users to complete or propagate the information of tasks. Additionally, for the ease of presentation, we list the notations frequently used in this paper in Table 1.

3.1. System Model. In this paper, we advocate a mobile crowd system as shown in Figure 1, which consists of three characters: service requesters, platform, and mobile users. Their roles are defined as follows:

- (1) Service requesters: when the service requesters want to know some useful information (e.g., the driver wants to know the traffic condition of the target road where he is going to), the requesters specify the requests as a set of keywords, including the location and time requests of the task and the quality of sensing data. Then, service requesters upload the sensing

TABLE 1: The summary of notation.

Parameters	Meanings
$G = (U, E)$	Social network
u, v, w	Users in U
(u, v)	Edge from user u to user v
$N(u)$	Neighbor set of user u
S	The selected seed set
k	Number of selected seeds
L	The location set
loc_j	The location of task t_j
tim_j	The required time of task t_j
R	The range of tasks to recruit user
$sim(u, v)$	The <i>Jaccard</i> similarity of user u and user v
$pos(u, j)$	The geographical performance of user u in the location loc_j at time tim_j
$pos_j(u, v)$	The possibility of user v complete task t_j under influence of user u
$\Phi(S)$	Expected number of completed tasks of user set S

requests to the platform and expect to get the sensing data from the platform.

- (2) Platform: the major function of the platform is allocating the sensing task to the eligible mobile users according to their spatiotemporal characteristics and other capacities that tasks required.

Considering the task sets to be allocated in the platform, denoted as $T = \{t_1, t_2, \dots, t_n\}$, where t_j is the j^{th} task. Each task can be characterized by a tuple with three elements: $\{loc_j, tim_j, R\}$, $1 \leq j \leq n$, where loc_j is the location of task t_j and tim_j is the required time of task t_j . To improve the completion rate of tasks, if the distance between one user and the task t_j is less than R at time tim_j , the platform regards that this user can complete the task t_j .

- (3) Mobile users: mobile users are the main force to complete tasks. After receiving the task information from the platform, mobile users collect sensing data and upload it to the platform. In this paper, we define mobile users from the social aspect. Let $G = (U, E)$ denote the social network, in which U represents the set of users $U = \{u_1, u_2, \dots, u_m\}$ and E represents the relationship among users. For any two users u and v in U , we regard they are connected if $(u, v) \in E$ and can share the task information with each other with a certain probability.

According to the main roles in the system model, the social-aware task allocation problem (SATA) in this paper can be described as follows: the service requesters send the service requests to the platform and expect to get the service from the platform. To provide timely service, the platform

launches the task information and selects eligible users to collect sensing data. However, due to the ‘‘cold effect,’’ there are only a few users available to perform the tasks. Under this circumstance, the platform selects a part of users as seeds to complete tasks or propagate the information with the assistance of social networks under the specific rule to ask their friends for help. The detail of this rule is presented in the next section. The object of the platform is to maximize the expected number of completed tasks. It is worthy to note that if there are multiple seeds propagated the information to the common friends in the social network, we choose the influence path with maximum probability.

3.2. Propagation Model. As mentioned above, the SATA problem is similar to the location aware influence maximization problem in the social network, which is aimed at selecting a subset from candidate users with a cardinality of k to maximize the influence spread in the target location under a certain propagation model. The indicator of influence spread for selected seeds is usually defined as the number of influenced users. To simulate the propagation process among users, we employ the independent cascade (IC) model, which is widely used in information propagation. In the IC model, every edge in the social network is associated with a weight to measure the propagation possibility from one node to his neighbor. The process of a traditional IC model can be described as follows:

- (1) In the initial stage, all nodes in the social network are in the inactive state
- (2) If a node u is selected as seed, then, it will get a chance to influence its neighbor nodes, which has not been selected as seeds
- (3) The process continues until selected seed set reaches to the required cardinality

According to the process of the IC model, the task information propagation in our problem can be described as follows: if the seed cannot complete the task, he should propagate the task information to his friends with a certain probability and expect friends can complete the sensing task. To measure the probability of a nonseed can complete tasks under the influence of seeds, we define two major factors that affect task completion.

- (1) Geographical performance: when a user can not complete this task himself, he should propagate the task to his friends who might be in the target location. We adopt the empirical statistical model to measure the geographical performance of users. The possibility that a user visits the location loc_j at the time tim_j is modeled by the frequency of visiting in the location loc_j at the time tim_j . Mathematically, it can be computed as follows:

$$pos(u, j) = \frac{n_u(loc_j, tim_j)}{\sum_{loc_j \in L} n_u(loc_j, tim_j)}, \quad (1)$$

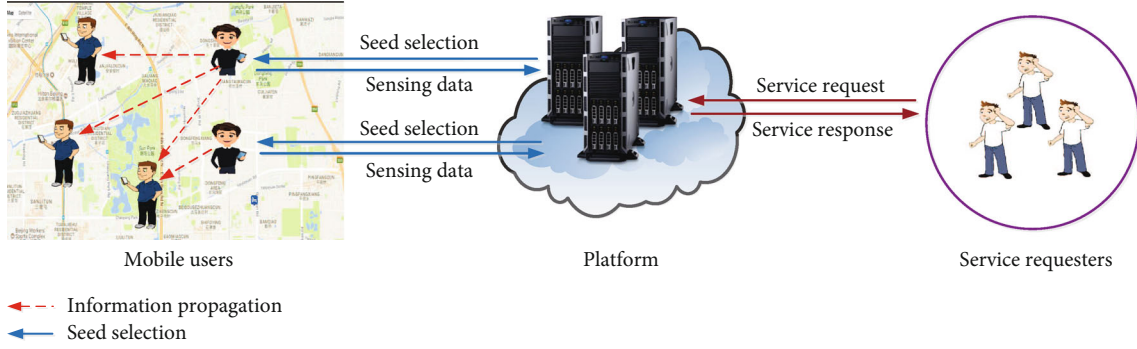


FIGURE 1: The framework of the MCS system.

where L is all the location set and $n_u(\text{loc}_j, \text{tim}_j)$ denotes the check-in times of user u in the location loc_j at the time tim_j in the historical check-in records.

- (2) Social similarity: intuitively, users prefer to share information with close friends. So the social relationship between two users should be considered in the process of propagation. In this paper, we adopt *Jaccard* coefficient to measure the social similarity of two users

$$\text{sim}(u, v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}, \quad (2)$$

where $N(u)$ is the neighbor set of user u . Larger similarity indicates that two users have a closer social relationship and can share task information with higher possibility.

By considering the factors of geographical performance and social similarity, the probability that nonseed v can complete the task t_j after influenced by a seed u can be calculated as follows:

$$\text{pos}_j(u, v) = \text{pos}(v, j) * \text{sim}(u, v). \quad (3)$$

Since selected seeds may have some common friends. During the information propagation, these common friends are influenced by multiple seeds simultaneously. It is worthy to note that the possibility that these common friends complete tasks is not simply computed by the utility summation of selected seeds. Alternatively, we regard a nonseed to complete the task with the probability that influenced by the seed who has maximum social similarity with him.

3.3. Problem Formulation. According to the definitions above, we can formulate the problem: given a social network and a set of tasks, the platform tries to allocate these tasks with the assistance of social networks. Specially, the platform selects k users as seeds. These seeds may only satisfy the spatiotemporal requirement of part of tasks. For those tasks that they cannot complete, they propagate the task information to their friends according to the rule defined in the last section.

Thus, the possibility of selected users to complete the task t_j can be computed from two perspectives: if u in the location of t_j at the required time, he can complete the task. Otherwise, u propagates the task information to his friends. Assume that friends are willing to complete the task after receiving the propagation from his friends if they are in the target location. So the completion possibility of a task can be computed by Equation (3).

For the selected seed set, their utility to a task t_j can be computed as follows: if it exists a seed that can complete the task t_j ; the complete possibility is equal to 1. Otherwise, the selected seeds propagate the task information to their friends. For the latter case, it is inappropriate to compute the utility summation of nonseeds influenced by the selected users. Instead, the completion possibility of a task can be measured by the possibility that at least one friend can complete under the influence of selected seeds. Thus, the formula is listed as follows:

$$p(S, t_j) = \begin{cases} 1, & \exists u \in S, \text{pos}(u, j) = 1, \\ 1 - \prod_{w \in N(u)} (1 - \text{pos}_j(u, w)), & \text{others.} \end{cases} \quad (4)$$

This task allocation problem is aimed at maximizing the expected number of completed tasks under a limited number of seeds and propagating task information with the assistance of social networks. Based on the above definitions, the problem can be formulated as follows:

$$\begin{aligned} \max : & \sum_{j=1}^n p(S, t_j) \\ \text{s.t.} & |S| = k. \end{aligned} \quad (5)$$

4. Theoretical Analysis and Algorithm Design

In this section, we first prove that the SATA problem is an NP-hard problem. Then, we further prove that it is also a submodular problem. For the ease of presentation, we use the $\Phi(S)$ as the expected number of completed tasks of user set S in the following sections.

4.1. NP-Hard

Theorem 1. *The SATA problem is an NP-hard problem.*

Proof. The SATA problem is aimed at selecting k seeds to maximize the expected number of completed tasks, which can be transferred into the maximum coverage problem. The traditional maximum coverage problem can be described as follows: given a collection set $S = \{S_1, S_2, \dots, S_m\}$ and a number k , the object is to find a subset $S' \subset S$, such that $|S'| < k$ and the number of covered elements in S is maximized. In our problem, we can treat the union set of tasks T and users U as the set S , i.e., $S = T \cup U$. Because every user can cover himself, selecting k user to cover the set S is equivalent to covering the set T . So our problem is a special case of maximum coverage problem, which is a well-known NP-hard problem. Thus, SATA is also an NP-hard problem.

4.2. Submodular Problem

Theorem 2. *Giving the seed set S , the $\Phi(S)$ is a submodular function. More specifically, for two seed sets A and B , $A \subset B$, a new user $u \in U/B$, the following equation holds:*

$$\Phi(A \cup \{u\}) - \Phi(A) \geq \Phi(B \cup \{u\}) - \Phi(B). \quad (6)$$

Proof. We prove the theorem from two properties, monotonic property and submodular property. For the sake of simplicity, we discuss the circumstance that $|B| - |A| = 1$ and $B - A = \{u'\}$.

- (1) Monotonic property: since $A \subset B$, so the sensing task completed by the set A is also completed by the set B . For those tasks that have not been completed by the set A , user u' can complete tasks or propagate these tasks to his friends. According to the definition of propagation performance and Equation (4), the following equation holds:

$$\Phi(B) - \Phi(A) = \sum_{j=1}^n P\left(\left\{u'\right\}, t_j\right) \geq 0. \quad (7)$$

Therefore, $\Phi(S)$ is a nondecreasing function.

- (2) Submodular property: Given a new user u , for those tasks that u can complete, two user sets have the same increment. So the following equation holds:

$$(\Phi(B \cup \{u\}) - \Phi(B)) - (\Phi(A \cup \{u\}) - \Phi(A)) = 0. \quad (8)$$

For those tasks that user u cannot complete, u propagates the task information to his friends. To analyze this situation elaborately, we divide this situation into three cases:

- (1) Case 1: user u is independent to set A and set B ; i.e., there are no common friends between u and two user sets. When adding the new user u into the two sets,

both two union sets only increase the performance of u . So Equation (8) holds.

- (2) Case 2: user u has common friends u' with set B but no common friends with set A . For the sake of simplification, we discuss there is only one task t_j to be propagated. Assume that u and u' have a common friend v . v completes the t_j with the probability of p_1 under the influence of u , and with the probability of p_2 under the influence of u' . The performance that u' influences other friends keeps unchanged. However, the performance that it influences user v should be recomputed; then we have:

$$\begin{aligned} & (\Phi(B \cup \{u\}) - \Phi(B)) - (\Phi(A \cup \{u\}) - \Phi(A)) \\ &= \Phi(B) - \Phi(u') + \Phi(u, u') - \Phi(B) - \Phi(u) \\ &= \Phi(u, u') - \Phi(u') - \Phi(u) \\ &= \max\{p_1, p_2\} - p_2 - p_1 < 0. \end{aligned} \quad (9)$$

- (3) Case 3: user u has common friends with sets A and B . In this case, the possibility of common friends being influenced changes equally. Equation (8) holds.

By combining Equations (7), (8), and (9), we can conclude that $\Phi(S)$ is a submodular function.

4.3. Algorithm Design. According to the analysis above, we can extend the native greedy selection algorithm to solve the SATA problem, which can return at least $(1 - 1/e)$ -approximation ratio of the optimal result and is widely used to solve the influence maximization problem [29]. Algorithm 1 shows the details of the native greedy selection algorithm. It selects the seed with the maximum margin gain in terms of the expected number of completed tasks in each round. Then, the completed tasks should be deleted least reallocated in the next round. This process continues until the number of the selected seeds reaches to k .

However, the NGS algorithm has some drawbacks. It should compute the performance of all candidate users to all tasks. So the running time of the NGS algorithm increases with the number of tasks. Different from the influence maximization problem, which simply computes the influence summation of selected users, according to the definition of the SATA problem, we aim to maximize the expected number of completed tasks, the most important part of which is the number of tasks that are actually completed. If a task can be completed by a seed, we need not to propagate the task information. Thus, a user who can complete more tasks should get higher priority than those who complete fewer tasks and propagate more tasks to friends. The advantage of this strategy is that after deleting the completed tasks from the task list, the number of tasks to be allocated in the next round will decrease, so it can reduce the running time to traverse task list.

```

Input:  $G$ : social network,  $k$ : seed set size,
          $T$ : task set,  $U$ : candidate seeds.
Output: a set of  $k$  seeds.
1:  $S \leftarrow \emptyset$ ;
2: while  $|S| < k$ 
3:   for  $u \in U/S$  do
4:      $u \leftarrow \arg \max \{\Phi(S \cup \{u\}) - \Phi(S)\}$ ;
5:      $S \leftarrow S \cup \{u\}$ ;
6:   end for
7:   delete the completed tasks from  $T$ ;
8: end while
9: return  $S$ ;

```

ALGORITHM 1: Native greedy selection algorithm (NGS).

According to the analysis above, we further propose the FGS algorithm which considers the number of tasks that seeds can actually complete. In the FGS algorithm, we divide the selection process into two stages: in the first stage, the candidate seeds can complete part of sensing tasks. We select the seed which can complete the maximum number of tasks. In the second stage, the rest candidate seeds cannot complete any remaining tasks. In this situation, we change the object to the propagation utility of seeds to the rest of tasks and select the seed with maximum propagation margin gain. Because FGS deletes more completed tasks in the first stage, there is only part of tasks to be allocated in the second stage, so it is more efficient. The details of FGS can be shown in Algorithm 2.

4.4. Complexity Analysis. In this section, we analyze the time complexity of the proposed algorithms. According to Algorithm 1 and Algorithm 2, we should traverse the remaining users in each round and compute their utility to the remaining tasks. This process continues until the cardinality of the selected seed set reaches to k . So the time complexity of NGS and FGS are $O(kmn)$. It is worthy to note that FGS has the same time complexity with NGS; however, FGS deletes more completed tasks in each round and reduces the running time in the next round. Thus, FGS is faster than NGS.

5. Evaluation

In this section, we evaluate the performance of our proposed algorithms. We firstly present the detailed description of datasets and some basic experiment settings. Secondly, we introduce the baseline algorithms for evaluation. Finally, the detailed results of the proposed algorithms and baseline algorithms are presented and analyzed.

5.1. Datasets. We adopt two widely used real-world datasets, Brightkite and Gowalla datasets [30], to evaluate the performance of the proposed algorithms and baseline algorithms. These two datasets contain both the mobility trace of users and the relationships among users.

The Brightkite dataset is a service provider where users shared their locations by check-in, and the friendship net-

```

Input:  $G$ : social network,  $k$ : seed set size,
          $T$ : task set,  $U$ : candidate seeds.
Output: a set of  $k$  seeds.
1:  $S \leftarrow \emptyset$ , flag = False
2: while  $|S| < k$  and flag == False do
3:   for  $u \in U/S$  do
4:      $u \leftarrow \arg \max \{complete\_num\}$ 
5:     if  $\max \{complete\_num\} == 0$  then
6:       flag = True;
7:       break;
8:     end if
9:      $S \leftarrow S \cup \{u\}$ ;
10:  end for
11:  delete the completed tasks from  $T$ ;
12: end while
13: if  $|S| < k$  then
14:    $u \leftarrow \arg \max \{\Phi(S \cup \{u\}) - \Phi(S)\}$ ;
15:    $S \leftarrow S \cup \{u\}$ ;
16: end if
17: return  $S$ ;

```

ALGORITHM 2: Fast greedy selection algorithm (FGS).

work can be collected using the public API. The dataset consists of 58228 nodes and 214078 edges. There are 4491143 check-in records during Apr. 2008-Oct. 2010. In our simulation experiments, we select the check-in records during Apr. 2008-Mar. 2009 to train the geographical performance of users and then randomly select the check-in locations on Apr.1, 2009, as the locations of sensing tasks. The required time of tasks is also randomly set in this day. The part of tasks can be shown in Figure 2(a).

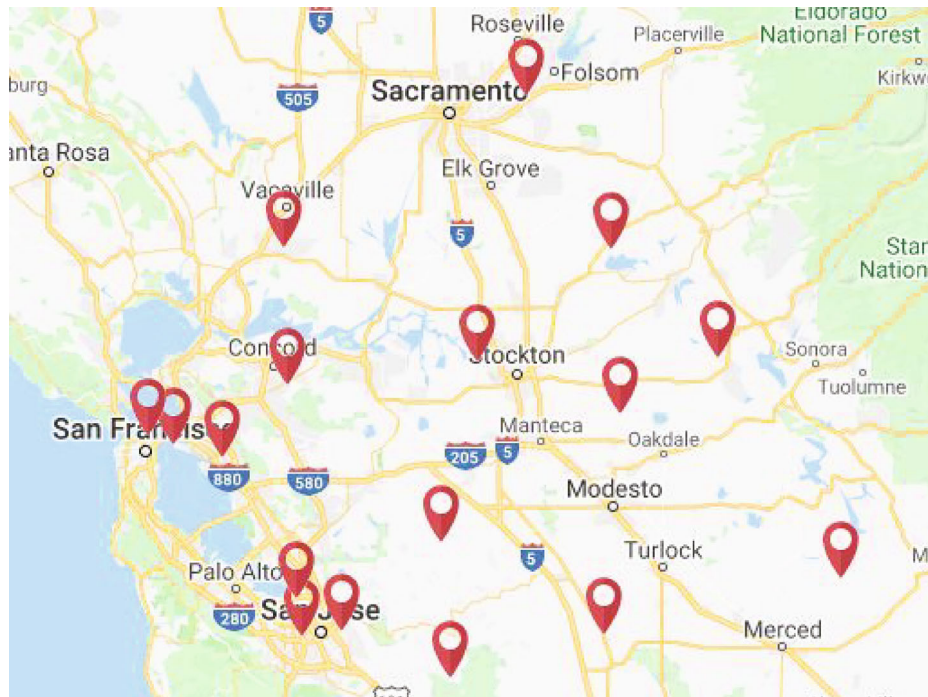
Similar to the Brightkite, Gowalla is also a location-based network that users share their locations by checking in. The friendship network was collected using their API. There are 196591 nodes and 950327 edges. The dataset collects 6442890 check-in records during Feb. 2009-Oct. 2010. In this paper, we select the check-in records during Feb. 2009-Jan. 2010 to train the geographical performance of users and then randomly select locations on Feb.1, 2010, as the locations of sensing tasks. The required time of tasks is also randomly set in this day. The part of tasks can be shown in Figure 2(b).

5.2. Baseline Algorithms. To compare the performance of the proposed algorithms, several baseline methods are designed as follows:

- (1) Propagation-based greedy selection (PGS): this method selects the seed with the largest propagation performance iteratively until the number of seeds reaches the limitation.
- (2) Degree-based greedy selection (DGS): this method simply selects the seeds with the largest degree, which is aimed at propagating the task information to more friends.
- (3) Random selection (RS): this method randomly selects k users as seeds to complete and propagate task



(a) Part of tasks in Brightkite



(b) Part of tasks in Gowalla

FIGURE 2: Part of tasks in two datasets.

information. To reduce the error caused by randomness, we run the RS algorithm 30 times and compute the average expected number of completed tasks as the final result.

In the following simulations, all the experiments are conducted on a PC with 3.10 GHz CPU and 16 GB memory. We compare two indicators of algorithms: the expected number of completed tasks and running time. In terms of efficiency

of algorithms, because the DGS selects the users with top k degree while the degree of users keeps unchanged in the whole seed selection process, so target seeds can be easily selected. For RS, it randomly selects k users as seeds without any heuristic strategies. Compared to the NGS, PGS, and FGS, these two algorithms can quickly identify the target users. It does not make sense to discuss the running time of these two algorithms. So we only compare the running time of the NGS, PGS, and FGS.

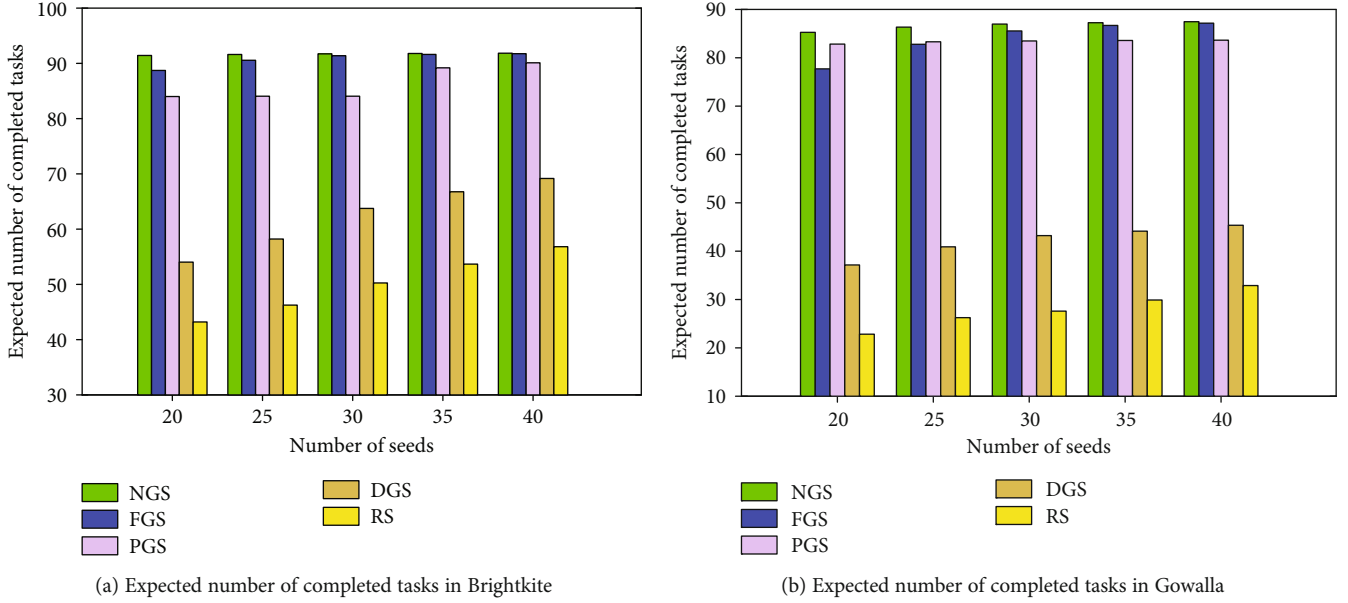


FIGURE 3: Performance comparison on the number of seeds.

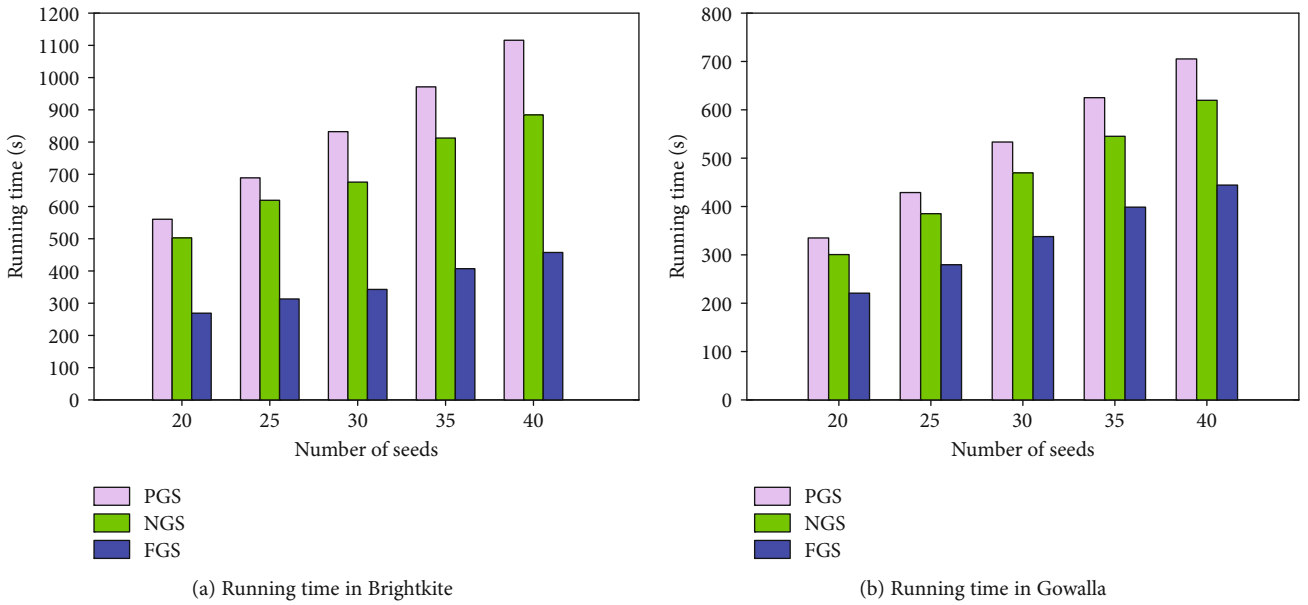
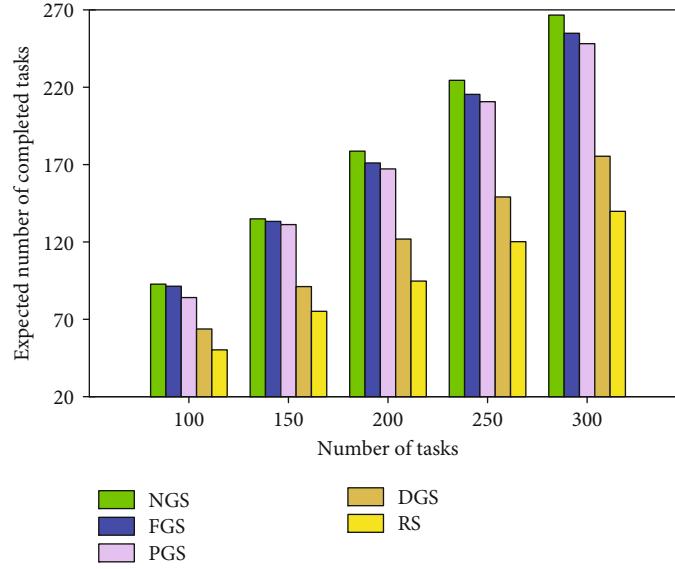


FIGURE 4: Running time comparison on the number of seeds.

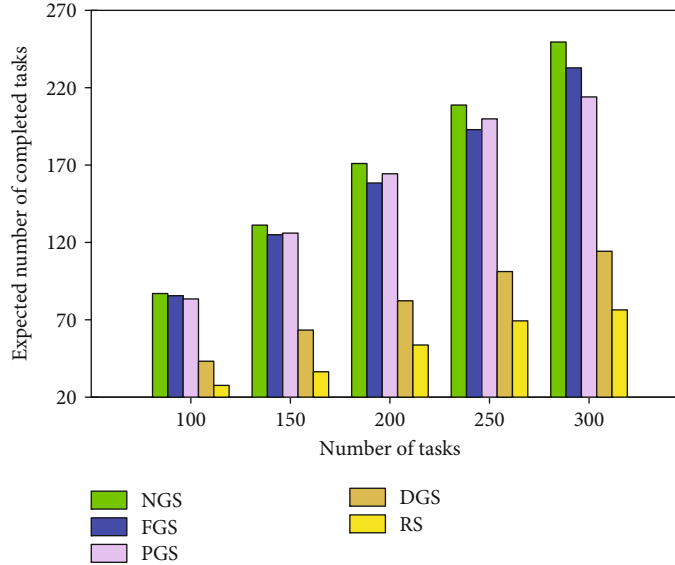
5.3. *Impact of Number of Seeds.* The main objective of the SATA problem is to maximize the expected number of completed tasks. To measure the impact of the seed number on the expected number of completed tasks, the related parameters are set as follows: for the Brightkite dataset, we set $m = 100$, $n = 100$, and $R = 10$ and change the number of seeds from 20 to 40 with an increment of 5. For the Gowalla dataset, we set $R = 5$, while other parameters are the same with those of the Brightkite dataset. Figures 3 and 4 show the results of our proposed algorithms and baseline algorithms.

From Figure 3, we can find that with the increment of seed number, the expected number of completed tasks also

increases. This is reasonable because more seeds can complete or propagate more tasks. NGS outperforms other algorithms while the increment of NGS is smaller than those of others. Because our task allocation problem is a submodular problem and has the diminishing return property, which matches the theoretical result in Theorem 2, FGS firstly selects the seed user who can actually complete the maximum number of tasks. However, these seeds may have limited capacity to propagate the tasks that they cannot actually complete. So the performance is a little worse than NGS. PGS just focuses on the seed with the largest propagation performance and ignores the number of actual completed tasks.



(a) Expected number of completed tasks in Brightkite



(b) Expected number of completed tasks in Gowalla

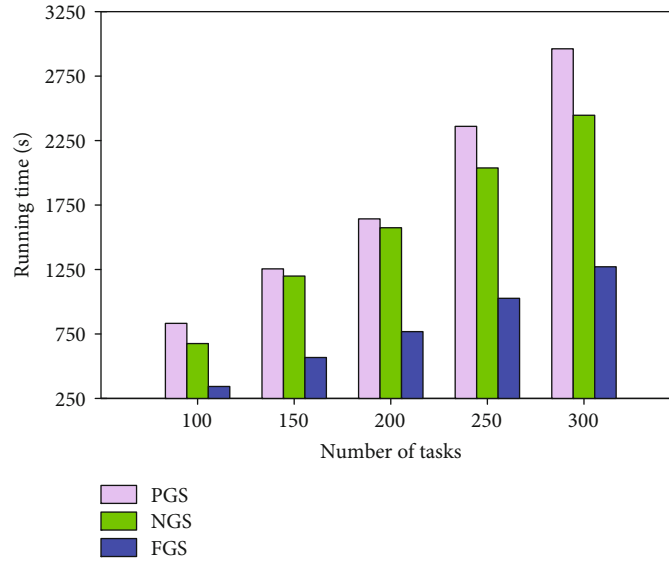
FIGURE 5: Performance comparison on the number of tasks.

So it gets worse results. For DGS, it firstly selects a user with the maximum degree, which is aimed at propagating the task information to more friends and ignores the number of tasks that seeds can actually complete. So it gets worse results than FGS and NGS algorithm. RS selects seeds randomly and does not consider any heuristic strategy in the process of seed selection, so it gets the worst results in terms of the expected number of completed task.

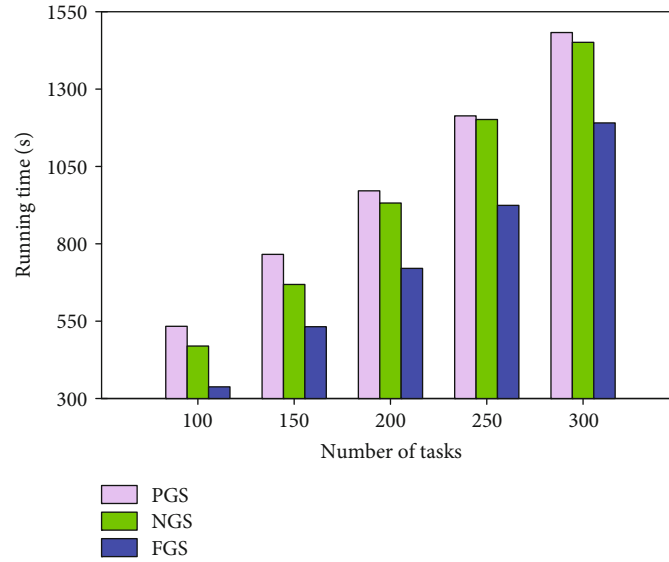
From Figure 4, we can see that the running time of three algorithms increases with the increment of seed number. Because selecting more seeds needs more rounds to traverse the candidate users, PGS consumes the most time among the three algorithms. The reason is that the seeds selected by the PGS have the maximum propagation performance. However, they may only complete a small part of tasks in

each round, leaving most of tasks to be allocated in the next round. Thus, it needs more time in every round. In contrast, FGS firstly selects the user who can complete most tasks; with the increment of seeds, more and more tasks can be completed. The number of tasks to be allocated is greatly decreased in the next round. So it is faster than the NGS and PGS. In summary, for the impact of the number of selected seeds, it is worthy to note that the expected number of completed tasks in FGS is slightly worse than that in NGS; however, the time consumption of FGS is far less than those of NGS and PGS.

5.4. Impact of Number of Tasks. In this section, we conduct experiments to illustrate the impact of the number of tasks on the performance of algorithms. The related parameters



(a) Running time in Brightkite



(b) Running time in Gowalla

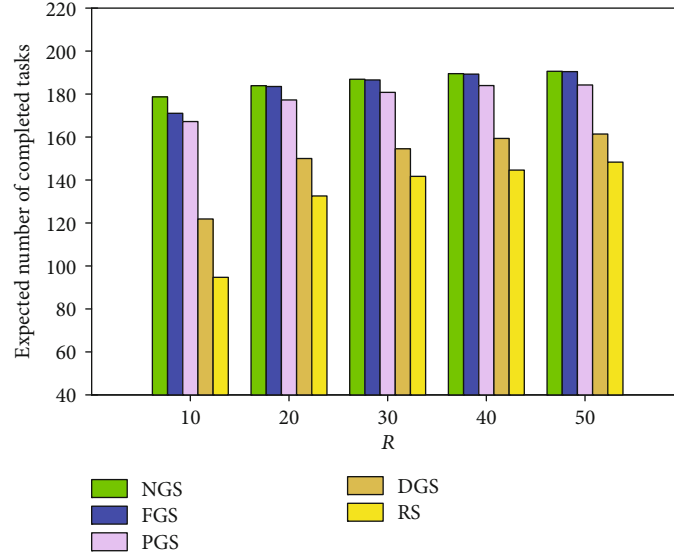
FIGURE 6: Running time comparison on the number of tasks.

are set as follows: for the Brightkite dataset, we set $m = 100$, $R = 10$, and $k = 30$ and change the number of tasks from 100 to 300 with an increment of 50. For the Gowalla dataset, we set $R = 5$, while other parameters are the same with those of the Brightkite dataset. The results of our proposed algorithms and baseline algorithms are shown in Figures 5 and 6.

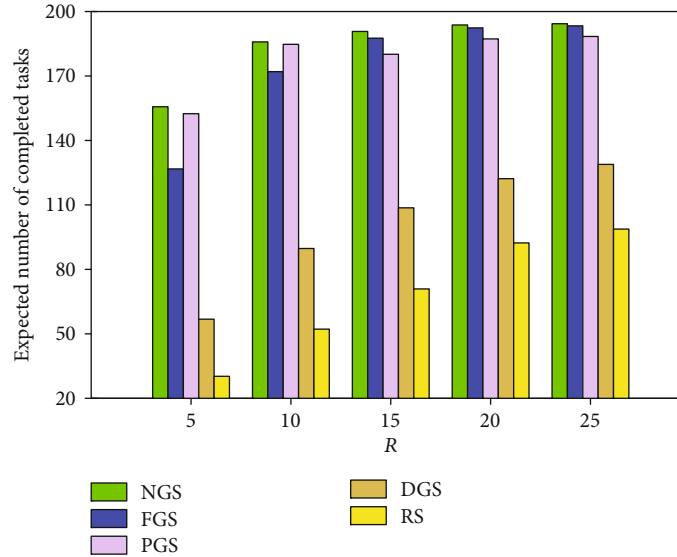
As shown in Figure 5, the expected number of completed tasks increases with the number of tasks. It indicates that seeds can complete and propagate more tasks even if the number of seeds is limited. NGS gets better results compared with other algorithms, and the difference among them becomes more and more significant with the increase of tasks. FGS focuses on the number of actually completed tasks. Under the limited candidate user resource, the number

of tasks that seeds can complete is also limited; the actual completion ratio decreases with the increment of the number of total tasks. In this situation, the seeds selected by FGS have the weaker capacity to propagate the tasks that they cannot complete. However, the NGS considers both the number of tasks that users can complete and the capacity to propagate the task information that they cannot complete. So it gets better results. PGS gets worse results than NGS and FGS in the Brightkite dataset. For the Gowalla dataset, PGS gets better results than FGS when the task number is 200 and 250. DGS selects the user with the maximum degree as seed to propagate the task, so the growth rate decreases with the increase of task number.

According to Figure 6, the running time of algorithms increases with the number of tasks. Because more tasks need



(a) Expected number of completed tasks in Brightkite



(b) Expected number of completed tasks in Gowalla

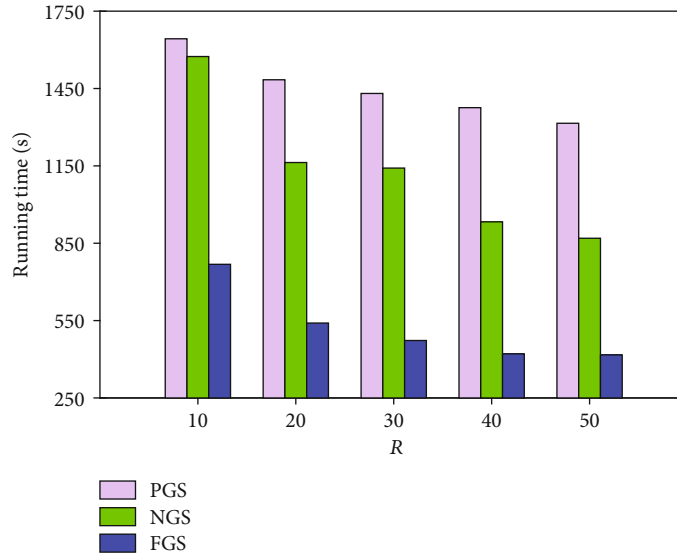
FIGURE 7: Performance comparison on R .

more time for seeds to propagate task information, PGS consumes the most time among three algorithms, because PGS needs to propagate more task information due to the low completion ratio of selected users in the previous rounds. In the Brightkite dataset, FGS costs nearly half running time compared with NGS. In the Gowalla dataset, though FGS costs less running time than NGS, the advantage in running time becomes less obvious with the increment of the number of tasks.

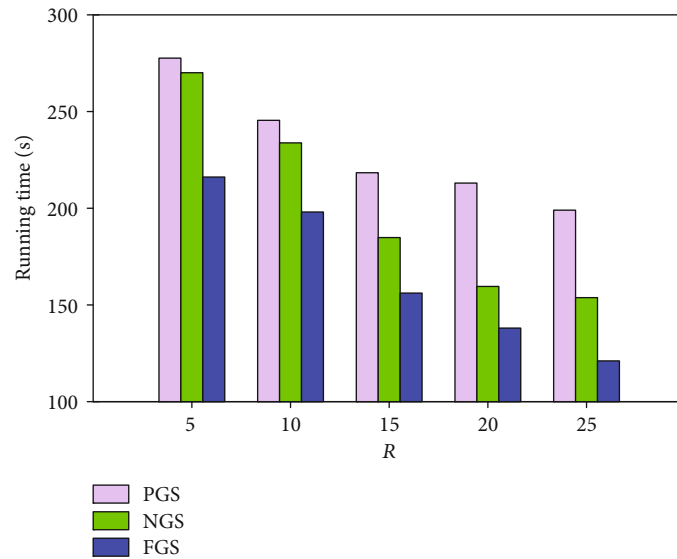
5.5. Impact of R . In this line of simulations, we conduct experiments to illustrate the impact of parameter R to the performance of algorithms, where R is the seed recruiting range for every task. The related parameters are set as follows: for the Brightkite dataset, we set $m = 100$, $n = 200$, and $k = 30$ and change R from 10 to 50 with an increment of 10. For the

Gowalla dataset, we set $k = 10$ and change R from 5 to 25 with an increment of 5. The results of the expected number of completed tasks and running time are shown in Figures 7 and 8.

From Figure 7, we can see that with the increment of R , the expected number of completed tasks increases. The advantage of NGS becomes more and more obvious. The reason can be listed as follows: on the one hand, seeds can complete more tasks within the larger range, even though the number of seeds is limited. On the other hand, seeds can propagate the task information to more friends if the geographical range is enlarged. So the expected number of completed tasks increases with the increment of R . However, for the Gowalla dataset, PGS gets better results compared to FGS when $R = 5$ and $R = 10$. Because PGS firstly selects the user with the maximum propagation



(a) Running time in Brightkite



(b) Running time in Gowalla

FIGURE 8: Running time comparison on R .

performance, if the selected user just completed a small part of tasks, it may get lower result in terms of the expected number.

According to Figure 8, the running time of three algorithms decreases with the increment of R . This can be explained that enlarging the range of users to complete tasks, the number of actually completed tasks increases, seeds selected by three algorithms can complete more task, and there is no need to propagate the task information to friends; thus, the time to propagate the task information decreases. Furthermore, the efficiency of FGS is prominent with the increment of R . Compared to the NGS and PGS, the seeds firstly selected by FGS can complete more tasks. The number of tasks to be allocated decreases significantly in the next round, so the running time decreases with the increment of recruit range.

6. Conclusion

In this paper, we study the social-aware task allocation problem (SATA) in mobile crowd sensing, which is aimed at selecting a part of users as seeds to complete or propagate tasks. Firstly, we illustrate the task information propagation under the IC model with the consideration of geographical performance and social similarity factors. Secondly, we prove that the SATA problem is an NP-hard and submodular problem. To solve this problem, we devise two greedy-based algorithms (NGS and FGS) to select seeds. NGS selects seeds with the maximum margin gain in terms of the expected number of completed tasks. FGS selects seeds which can actually complete the maximum number of tasks. The experimental results show that FGS gets slightly worse results than NGS; however, it greatly reduces the running time.

Data Availability

The source data used in the study is available in the following website: <http://snap.stanford.edu/data/>.

Conflicts of Interest

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

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (Nos. U1705262, 61772136, and 61672159).

References

- [1] B. Guo, Z. Wang, Z. Yu et al., "Mobile crowd sensing and computing: the the review of an emerging human-powered sensing paradigm," *ACM Computing Surveys*, vol. 48, no. 1, pp. 1–31, 2015.
- [2] W. Guo, W. Zhu, Z. Yu, J. Wang, and B. Guo, "A survey of task allocation: contrastive perspectives from wireless sensor networks and mobile crowdsensing," *IEEE Access*, vol. 7, pp. 78406–78420, 2019.
- [3] J. Wang, L. Wang, Y. Wang, D. Zhang, and L. Kong, "Task allocation in mobile crowd sensing: state-of-the-art and future opportunities," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3747–3757, 2018.
- [4] J. Chen and J. Yang, "Maximizing coverage quality with budget constrained in mobile crowd-sensing network for environmental monitoring applications," *Sensors*, vol. 19, no. 10, article 2399, 2019.
- [5] S. Morishita, S. Maenaka, D. Nagata et al., "Sakurasensor: quasi-real time cherry-lined roads detection through participatory video sensing by cars," in *UbiComp '15: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 695–705, New York, NY, USA, September 2015.
- [6] T. Ludwig, C. Reuter, and V. Pipek, "What you see is what I need: mobile reporting practices in emergencies," *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*, O. Bertelsen, L. Ciolfi, M. Grasso, and G. Papadopoulos, Eds., , pp. 181–206, Springer, London, 2013.
- [7] D. Zhang, L. Wang, H. Xiong, and B. Guo, "4w1h in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 42–48, 2014.
- [8] B. Guo, Y. Liu, L. Wang, V. O. K. Li, J. C. K. Lam, and Z. Yu, "Task allocation in spatial crowdsourcing: current state and future directions," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1749–1764, 2018.
- [9] W. Guo, J. Li, G. Chen, Y. Niu, and C. Chen, "A pso-optimized real-time fault-tolerant task allocation algorithm in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 12, pp. 3236–3249, 2014.
- [10] Y. Liu, B. Guo, Y. Wang, W. Wu, Z. Yu, and D. Zhang, "Taskme: multi-task allocation in mobile crowd sensing," in *UbiComp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 403–414, New York, NY, USA, September 2016.
- [11] L. Wang, Z. Yu, Q. Han, B. Guo, and H. Xiong, "Multi-objective optimization based allocation of heterogeneous spatial crowdsourcing tasks," *IEEE Transactions on Mobile Computing*, vol. 17, no. 7, pp. 1637–1650, 2018.
- [12] E. Wang, Y. Yang, J. Wu, K. Lou, D. Luan, and H. Wang, "User recruitment system for efficient photo collection in mobile crowdsensing," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 1, pp. 1–12, 2019.
- [13] J. Wang, F. Wang, Y. Wang, D. Zhang, L. Wang, and Z. Qiu, "Social-network-assisted worker recruitment in mobile crowd sensing," *IEEE Transactions on Mobile Computing*, vol. 18, no. 7, pp. 1661–1673, 2018.
- [14] A.-q. Lu and J.-h. Zhu, "Worker recruitment with cost and time constraints in mobile crowd sensing," *Future Generation Computer Systems*, vol. 112, pp. 819–831, 2020.
- [15] S. Reddy, K. Shilton, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, "Using context annotated mobility profiles to recruit data collectors in participatory sensing," in *Location and Context Awareness. LoCA 2009. Lecture Notes in Computer Science, vol 5561*, T. Choudhury, A. Quigley, T. Strang, and K. Suginuma, Eds., pp. 52–69, Springer, Berlin, Heidelberg, 2009.
- [16] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment framework for participatory sensing data collections," in *Pervasive Computing. Pervasive 2010. Lecture Notes in Computer Science, vol 6030*, P. Floréen, A. Krüger, and M. Spasojevic, Eds., pp. 138–155, Springer, Berlin, Heidelberg, 2010.
- [17] M. Zhang, P. Yang, C. Tian et al., "Quality-aware sensing coverage in budget-constrained mobile crowdsensing networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7698–7707, 2015.
- [18] X. Wang, W. Wu, and D. Qi, "Mobility-aware participant recruitment for vehicle-based mobile crowdsensing," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4415–4426, 2017.
- [19] Z. Yu, J. Zhou, W. Guo, L. Guo, and Z. Yu, "Participant selection for t-sweep k-coverage crowd sensing tasks," *World Wide Web*, vol. 21, no. 3, pp. 741–758, 2018.
- [20] H. Li, T. Li, and Y. Wang, "Dynamic participant recruitment of mobile crowd sensing for heterogeneous sensing tasks," in *2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems*, pp. 136–144, Dallas, TX, USA, October 2015.
- [21] W. Zhu, W. Guo, Z. Yu, and H. Xiong, "Multi-task allocation to heterogeneous participants in mobile crowd sensing," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 7218061, 10 pages, 2018.
- [22] J. Jiang, B. An, Y. Jiang, C. Zhang, Z. Bu, and J. Cao, "Group-oriented task allocation for crowdsourcing in social networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–16, 2019.
- [23] Z. Wang, J. Zhao, J. Hu et al., "Towards personalized task-oriented worker recruitment in mobile crowdsensing," *IEEE Transactions on Mobile Computing*, 2020.
- [24] J. Wang, Y. Wang, D. Zhang et al., "Multi-task allocation in mobile crowd sensing with individual task quality assurance," *IEEE Transactions on Mobile Computing*, vol. 17, no. 9, pp. 2101–2113, 2018.
- [25] H. Wang, D. Zhao, H. Ma, and L. Ding, "Min-max planning of time-sensitive and heterogeneous tasks in mobile crowd sensing," in *2018 IEEE Global Communications Conference*

- (*GLOBECOM*), pp. 1–7, Abu Dhabi, United Arab Emirates, December 2018.
- [26] W. Ni, P. Cheng, L. Chen, and X. Lin, “Task allocation in dependency-aware spatial crowdsourcing,” in *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, pp. 985–996, Dallas, TX, USA, April 2020.
- [27] S. Song, Z. Liu, Z. Li, T. Xing, and D. Fang, “Coverage-oriented task assignment for mobile crowdsensing,” *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7407–7418, 2020.
- [28] Z. Peng, X. Gui, J. An, R. Gui, and Y. Ji, “Tdsr: A task-distributing system of crowdsourcing based on social relation cognition,” *Mobile Information Systems*, vol. 2019, Article ID 7413460, 12 pages, 2019.
- [29] Y. Yang, Y. Xu, E. Wang, K. Lou, and D. Luan, “Exploring influence maximization in online and offline double-layer propagation scheme,” *Information Sciences*, vol. 450, pp. 182–199, 2018.
- [30] E. Cho, S. A. Myers, and J. Leskovec, “Friendship and mobility: user movement in location-based social networks,” in *KDD '11: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1082–1090, New York, NY, USA, August 2011.

Research Article

A Novel Attack-and-Defense Signaling Game for Optimal Deceptive Defense Strategy Choice

Yongjin Hu,¹ Han Zhang,^{1,2} Yuanbo Guo,¹ Tao Li,¹ and Jun Ma ^{1,3}

¹Information Engineering University, Zhengzhou 450001, China

²Zhengzhou University, Zhengzhou 450001, China

³School of Telecommunications Engineering, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Jun Ma; sijunhan@163.com

Received 9 April 2020; Revised 2 June 2020; Accepted 14 September 2020; Published 12 October 2020

Academic Editor: Huaqun Wang

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

Increasingly, more administrators (defenders) are using defense strategies with deception such as honeypots to improve the IoT network security in response to attacks. Using game theory, the signaling game is leveraged to describe the confrontation between attacks and defenses. However, the traditional approach focuses only on the defender; the analysis from the attacker side is ignored. Moreover, insufficient analysis has been conducted on the optimal defense strategy with deception when the model is established with the signaling game. In our work, the signaling game model is extended to a novel two-way signaling game model to describe the game from the perspectives of both the defender and the attacker. First, the improved model is formally defined, and an algorithm is proposed for identifying the refined Bayesian equilibrium. Then, according to the calculated benefits, optimal strategies choice for both the attacker and the defender in the game are analyzed. Last, a simulation is conducted to evaluate the performance of the proposed model and to demonstrate that the defense strategy with deception is optimal for the defender.

1. Introduction

IoT networks and devices are highly vulnerable to sophisticated cyber-attacks. Despite the widespread deployment of security monitoring tools, which include firewalls and intrusion detection systems (IDSs), attackers can infiltrate target IoT devices by leveraging multiple attack vectors [1].

Recently, honeypot-enabled deceptive security mechanisms were introduced as an emerging proactive cyber defense strategy for confusing or misleading attackers and showed significant advantages over traditional security techniques [2]. For attackers, deceptive behaviors of defenders increase the uncertainty of the target to be compromised [3]. Attackers must spend additional resources (e.g., time and money) to deal with the uncertainty via reconnaissance and to develop situational awareness. In addition, deceptive behaviors prevent attackers from launching efficient custom attacks. For example, by collecting an attacker's information

when he is compromising a target device that is disguised by honeypots, the defender can use the learned knowledge to enhance the IoT network security [4]. As a result, deception by providing seemingly convincing yet misleading information to deceive attackers has become a major defense mechanism. With the wide utilization of deception, the security status of organizations has been substantially improved. When attackers are following the seven phases of the cyber kill chain [5] in launching an attack, deception approaches can be performed effectively in disrupting each stage of the cyber kill chain, as illustrated in Figure 1.

The contributions of the paper are the following.

(1) A two-way signaling game model based on the signaling game is formally defined to describe the confrontation from the perspectives of both the defender and the attacker. (2) With the two-way signaling game model, an algorithm is defined to identify the refined Bayesian equilibrium in the game. (3) With the deception strategy introduced, the

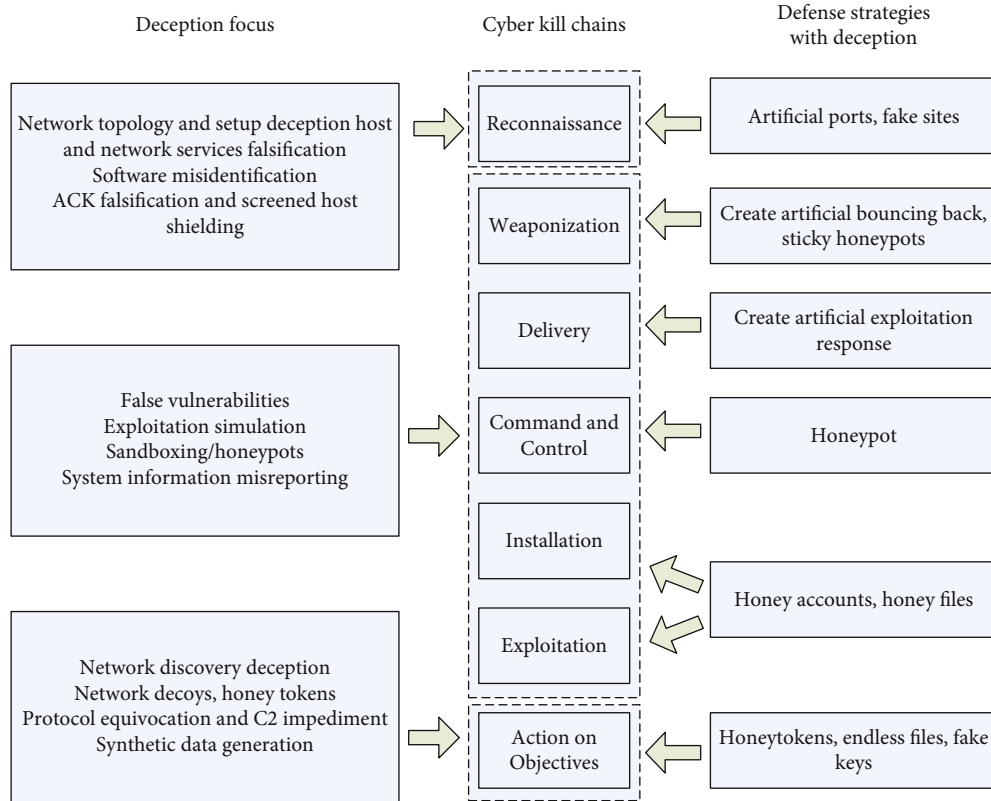


FIGURE 1: Deception focus on each stage of the cyber kill chain.

optimal strategies choice for both the attacker and the defender in the game is analyzed.

2. Related Works

In previous work [6], due to a lack of clarity regarding the concept of deception, deploying honeypots to detect an attacker and to obtain information on the attacker's intentions is the primary deception mode for the defender to use. For instance, Rowe et al. [7] showed how to decrease the number of attacks to which a network is subjected by utilizing fake honeypots, namely, by disguising normal systems as honeypots. Garg and Grosu [8] used a honeynet system to characterize deception, where defenders may have the choice to conceal a regular host as a honeypot (or inversely) in response to the attackers' probe. Seamus et al. [9] created a honeypot that simulates a ZigBee gateway to assess the presence of the ZigBee attack intelligence on a SSH attack vector in Wireless Sensor Networks (WSNs).

Until recent years, as deception became a powerful tool for protecting IoT networks and devices against attackers [10], game theory was introduced into the field of cybersecurity to model the interaction between defender and attacker and to identify the optimal defense strategies for both players. Cohen [11] comprehensively discussed deception as a technique for protecting information systems and concluded that deception has a positive effect for the defenders and a negative effect for the attackers. Carroll and Grosu [12] modeled the way deception affects the attack-defense interactions based on a game in which the players (defenders and

attackers) have incomplete knowledge of each other. Pawlick and Zhu [13] extended the signaling game by assuming that the adversary can obtain evidence of the true state of the system, and they concluded that the effectiveness of deceptive defenders sometimes increases if an adversary develops the ability to detect deception. Duan et al. [14] proposed an energy-aware trust derivation scheme using the game theoretic approach to manage overhead while maintaining adequate security of WSNs. Fugate and Ferguson [15] discussed techniques for combining artificial intelligence algorithms with game theory models to estimate hidden states of the attacker using feedback through payoffs to learn how to optimally defend the system using cyber deception. Additional works are listed in Table 1.

As discussed above, in contrast to the previous focus on the analysis of the defender, our work will describe the process from not only the perspective of the defender but also that of the attacker.

3. An Improved Signaling Game Model

3.1. Analysis of the Novel Attack-And-Defense Signaling Game. According to [22–24], the information that is released by the defender actively or the information that is leaked via defensive behavior passively is an important decision-making basis for the attacker. Such information is referred to as the signal that is sent by the defender, and the defense signal can affect the behavior of the attacker by changing the benefits to both the attacker and the defender. Furthermore, we believe that the information that is released by the

TABLE 1: Research on modeling deception defense by game theory.

Author	Focus of the study
Çeker et al. [16]	Modeled with a similar approach that uses game theory and provides the option of disguising a real system as a honeypot (or vice versa) to mitigate denial of service (DoS) attacks
Hichem et al. [17]	Proposed a game theoretic technique to activate anomaly detection technique only when a new attack's signature is expected to occur
Aaron et al. [18]	To increase the uncertainty of adversarial reconnaissance and introduced a novel game theoretic model of deceptive interactions between a defender and a cyber-attacker into responses to network scans or reconnaissance
Somdip [19]	Proposed a methodology in which game theory can be used to model the activity of stakeholders in the networks to detect anomalies such as collusion by using a supervised machine learning algorithm and algorithmic game theory
Pawlick and Zhu [20]	Investigated a model of signaling games in which the receiver can detect deception with a specified probability
Kun et al. [21]	Employed Nash equilibrium in the noncooperative game model and analyzes its efficiency in vehicular ad hoc networks

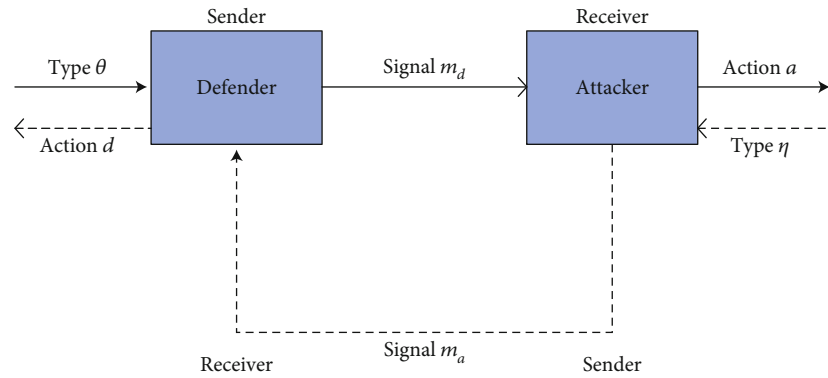


FIGURE 2: In a two-way signaling game model, the forward direction is defined as the defender sending a signal m_d to the attacker, who will infer the type of the defender θ and choose the action a ; the reverse direction is defined as the attacker sending a signal m_a to the defender, who will infer the type η of the attacker and choose the action d .

attacker and observed by the defender will also affect the defense decision and change the final attack-and-defense benefits. We construct an attack-and-defense behavior interaction model with incomplete information. According to signaling game theory, we analyze the dynamic game process and the signal mechanism from the perspectives of both attack and defense, and we investigate the influence of defense signals on the game equilibrium and strategy choice for both the attacker and the defender. We describe this process as a novel attack-and-defense signaling game that is defined as a two-way signaling game model, as illustrated in Figure 2.

The defender is defined as the leader of the signaling game, and the attacker is the follower when analyzing the forward signal transmission. The roles of the attacker and the defender will be exchanged when analyzing the reverse signal transmission. By constructing the attack-and-defense game process in both the forward and reverse directions, the influences of two examples on the defense strategy are analyzed: (1) in the forward phase, ① a defender mixes a defensive strategy with a (or no) deception strategy to deter, deceive, and induce the attacker and sends a defensive signal; ② the attacker forms an initial belief regarding the defender type by collecting reconnaissance information in advance and public information from the defender. The attack strategy is selected according to the calculation of the Bayesian posterior

probability for the defender type; and ③ the defender selects the optimal defense strategy for implementing security defense. (2) In the reverse phase, ① the attacker sends an attack signal while attacking; ② the defender forms a belief regarding the attacker. Under the action of the attack signal, the defender calculates the Bayesian posterior probability of the attacker type and corrects the defense strategy accordingly; and ③ the attacker corrects the current optimal attack strategy.

For convenience, we analyze the forward signaling game process and the reverse process separately; however, logically, these two processes are conducted simultaneously. Therefore, the strategy choice that is made by the defender is simultaneously affected by these two processes.

3.2. Formal Definition of the Two-Way Signaling Game Model

Definition 1. The two-way signaling game model for defense strategy selection with deception G_{DS} can be represented by a nine-tuple $G_{DS} = (N, \Theta, M, S, P_A, P_A', P_D, P_D', U)$, in which

① $N = (N_D, N_A)$ denotes the player set for a two-play game, where N_D denotes the set for the defender and N_A the set for the attacker.

② $\Theta = (\Theta_D, \Theta_A)$ denotes the type set for the defender and the attacker. The type of defender $\Theta_D = (\theta_i | i = 1, 2, \dots, n)$

is the private information, which determined by the defensive action that is taken; the type of attacker $\Theta_A = (\eta_j | j = 1, 2, \dots, n)$ is the private information of the attacker, which is determined by the attack action that is taken.

③ $M = (M_D, M_A)$ denotes the signal set for the defender and the attacker. $M_D = (m_d | d = 1, 2, \dots) M_D \neq \emptyset$ denotes that the defender selects and releases the signal according to the set signal release mechanism. For ease of representation, the signal name is consistent with the defender type name. The defense signal and the defender type are not necessarily consistent due to the objective of deceiving and inducing the attacker. Similarly, $M_A = (m_a | a = 1, 2, \dots) M_A \neq \emptyset$ denotes the attack signal that is sent by the attacker, and the signal name is the same as the attacker type name.

④ $S = (D, A)$ denotes the strategy set for the defender and the attacker, where $D = \{d_g | g = 1, 2, \dots\}$ and $A = \{a_h | h = 1, 2, \dots\}$ denote the defense strategy and the attack strategy, respectively.

⑤ P_A is the belief set of the attacker on the type of defender, where $P_A = (p_A(\theta_1), p_A(\theta_2), \dots, p_A(\theta_n)) = (\gamma_1, \dots, \gamma_n)$.

⑥ P_A' is the posterior probability set of the attacker on the type of defender, where, $P_A' = P_A'(\theta_i | m_d) = (\mu_1, \dots, \mu_n)$ denotes the posterior probability of the type of defender, which follows the Bayesian rule, after the attacker observes the defensive signal m_d .

⑦ P_D is the belief set of the defender on the type of attacker, where $P_D = (p_D(\eta_1), p_D(\eta_2), \dots, p_D(\eta_n)) = (\sigma_1, \dots, \sigma_n)$.

⑧ P_D' is the posterior probability set of the defender on the type of attacker, where $P_D' = P_D'(\eta_i | m_a) = (\delta_1, \dots, \delta_n)$ denotes the posterior probability of the type of attacker, which follows the Bayesian rule, after the defender observes the defensive signal m_a .

⑨ $U = (U_D, U_A)$ denotes the expected utility set of the defender and the attacker, whose value is determined by the strategies that are chosen by all players. The corresponding utility functions will be discussed in the next section.

3.3. Refined Bayesian Equilibrium Solution and the Optimal Defense Strategy Choice. According to Definition 1, this section extends the refined Bayesian equilibrium to the two-way signaling game model based on the definition of the refined Bayesian equilibrium [25] and proposes a refined Bayesian equilibrium algorithm for the two-way signaling game. Instances in the forward direction and in the reverse direction for the two-way signaling game model were constructed to show the details.

Definition 2. The equilibrium in a two-way signaling game model for defense strategy choice with deception is a refined Bayesian equilibrium if the following requirements are satisfied:

$$(I) a^*(m) \in \operatorname{argmax}_a \sum_{\theta} P_A'(\theta | m) U_2(m, a, \theta).$$

$$(II) m^*(\theta) \in \operatorname{argmax}_m U_1(m, a^*(m), \theta).$$

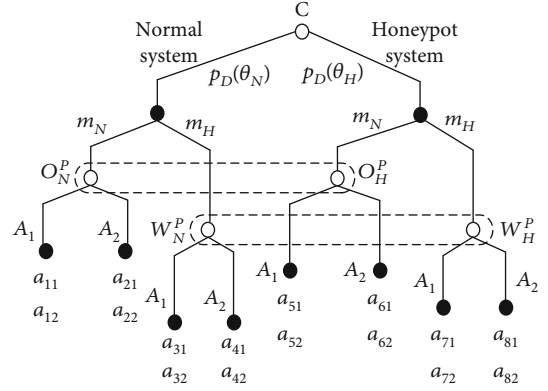


FIGURE 3: In the signaling game tree $G_{DS}(F)$ for the forward direction, $m^* = m_d^*(d^*, \theta) = m_d^*(\theta)$ indicates that defender θ is sending signal m_d^* and chooses strategy $d^*(m_d^*)$ according to the signal, which is denoted as $m_d^*(\theta)$; $a^* = a^*(a_h, m_d) = a^*(m_d)$ indicates that the attacker responds with action $a^*(a_h, m_d)$, which is denoted as $a^*(m_d)$; $P' = P_A'(\theta | m_d) = P_A'$ indicates that the attacker calculated $P_A'(\theta | m_d)$ as the posterior probability for the type of the defender, which is denoted as P_A' ; and the existence of a refined Bayesian equilibrium is abbreviated as $EQ = (m_d^*(\theta), a^*(m_d), P_A')$.

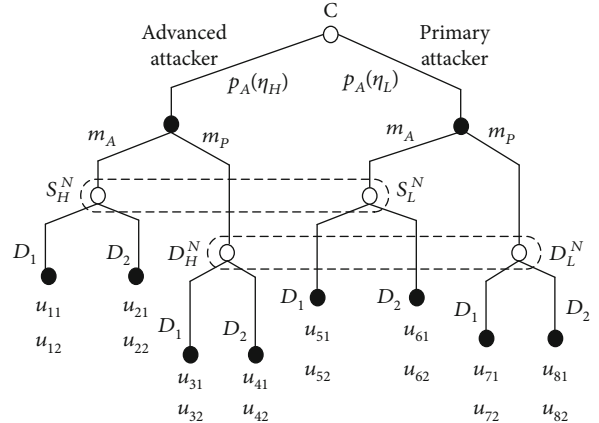


FIGURE 4: In the signaling game tree $G_{DS}(R)$ for the reverse direction, nature assigns type η_H with probability $p_A(\eta_H)$ and type η_L with $p_A(\eta_L)$. The attacker can send either signal m_A (signaling that the attacker is of type η_H) or m_P (signaling that the defender is of type η_L). The defender will revise her judgement on the type of the attacker by selecting $\{S_H^N, S_L^N\}$ if observing signal m_A and $\{D_H^N, D_L^N\}$ if observing signal m_P as the posterior probability for the type of the attacker $\{\eta_H, \eta_L\}$. u_{ij} denotes eight outcomes, where each outcome results in the corresponding payoff.

(III) $P'(\theta | m)$ is the posterior probability that is calculated by the signal receiver according to the Bayesian rule based on the prior probability $P(\theta)$, signal m , and the signal sender's optimal strategy $m^*(\theta)$.

In (I), $a^*(m)$ denotes the optimal action that is adopted by the signal receiver after obtaining the posterior probability $P'(\theta | m)$ of the type to which the signal sender belongs; U_2


```

Input: Model $G_{DS}$ , Signal direction parameter  $w$ 
Output: Optimal strategy for the defender
BEGIN
if ( $w=1$ )//forward-direction signalling game
{Initialize ( $\theta \in \Theta_D = (\theta_i | i = 1, 2, \dots, n)$ );
 //Initialize the type of the defender
 Initialize ( $M_D = (m_d | d = 1, 2, \dots), P = P_A = (p(\theta_1), \dots, p(\theta_n))$ );
 //Initialize the signal of the defender and the belief of the defender regarding the attacker
 }
If ( $w=0$ )//reverse-direction signalling game
{Initialize ( $\theta \in \Theta_A = (\eta_j | j = 1, 2, \dots, n)$ );
 //Initialize the type of the attacker
 Initialize ( $M_A = (m_a | a = 1, 2, \dots)$ 
  $P = P_D = (p_D(\eta_1), p_D(\eta_2), \dots, p_D(\eta_n))$ );
 //Initialize the signal of the attacker and the belief of the attacker regarding the defender
 }
Initialize ( $S = (D, A), D = \{d_1, \dots, d_g\}, A = \{a_1, \dots, a_h\}$ );
//Initialize the strategies for both players
while ( $a_h \in A \& \& m_j \in M \& \& d_g \in D$ )//Calculate the utility
{ $U_A(m_j, d_g, a_h, \theta_i) = \sum_{g,h} C_{sc}(d_g, a_h) - C_a$ ;
  $U_D(m_j, d_g, a_h, \theta_i) = \sum_{g,h} C_{sc}(d_g, a_h) - C_d - C_{ds}$ ;
 }
//Obtain the refined Bayesian Equilibrium
 $a^*(m) \in \arg \max_{a \in A} \sum P'(\theta | m) U_A(m^*(\theta), a, \theta)$ ;
 $m^*(\theta) \in \arg \max_{m \in M} U_D(m, a^*(m), d^*(m^*), \theta)$ ;
//Calculate the optimal strategy for attack and defence
Bayesian ( $P'_A(\theta)$ );
//Calculate the posterior probability and apply the Bayesian rule
for the defender
Create ( $m^*(\theta), a^*(m), P'_A(\theta)$ ); //Construct the refined Bayesian equilibrium
Sort ( $m^*(\theta)$ ); //descending
Output ( $m^*(\theta)$ ); //output the optimal strategy for the defender
End

```

ALGORITHM 1: Optimal strategy choice algorithm description based on a two-way signaling game model.

(m, a, θ) denotes the utility function of the signal receiver, which is the expected utility function of attacker $U_D(m_j, d_g, a_h, \theta_i)$ in the forward direction and the expected utility function of the defender $U_D(m_j, d_g, a_h, \theta_i)$ in the reverse direction; and $\theta \in \Theta = (\Theta_D, \Theta_A)$ denotes the type set for the defender and the attacker, where $\theta \in \Theta_D = (\theta_i | i = 1, 2, \dots, n)$ in the forward direction and $\theta \in \Theta_A = (\eta_j | j = 1, 2, \dots, n)$ in the reverse direction.

In (II), $m^*(\theta)$ denotes the optimal strategy that is selected by the signal sender after predicting the optimal action $a^*(m)$ of the signal receiver; $U_1(m, a^*(m), \theta)$ denotes the utility function of the signal sender, which is $U_D(m_j, d_g, a_h, \theta_i)$ in the forward direction, and $U_A(m_j, d_g, a_h, \theta_i)$ in the reverse direction.

In (III), $P'(\theta | m)$ indicates the posterior probability calculated by signal receiver according to the signal sent by the signal sender followed by the Bayesian rule, which is P_A' in the forward direction and P_B' in the reverse direction.

3.4. Method of Refined Bayesian Equilibrium in the Two-Way Signaling Game Model. The steps are as follows:

- (1) Construct the posterior inference $P(\theta | m)$ of various information sets on the signaling game tree
- (2) Calculate the optimal strategy for the signal receiver according to the posterior inference

When observing the signal $m \in M$, the signal receiver will choose optimal strategy $a^*(m)$ according to $P(\theta | m)$ for the type θ of the sender to maximize the expected utility U_2 , namely, the signal receiver will identify his optimal strategy $a^*(m)$ by calculating $\max \sum P(\theta | m) U_2(m(\theta), a, \theta)$.

- (3) Calculate the optimal strategy for the signal sender according to the posterior inference

The signal sender foresees that the signal receiver will select the optimal strategy based on observations of the signal that is released by him and chooses the strategy that maximizes the expected utility U_1 , namely, the signal sender identifies his optimal strategy $m^*(\theta)$ based on the posterior inference by calculating $\max U_1(m, a^*(m), \theta)$.

- (4) Calculate the refined Bayesian equilibrium

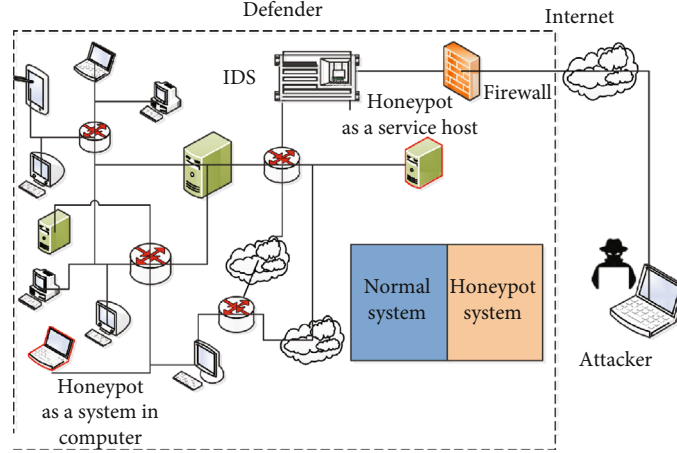


FIGURE 5: Game scenario with deception, which considers two decision makers, namely, a defender and an attacker. The defender deploys a honeypot in the IoT network as either a system or a service host. In the specified scenario, the forward and reverse transmissions occur simultaneously. The sequences of moves, type sets, and action sets follow the modeling elements that were discussed in the previous section. The incomplete information comes from the attacker’s uncertainty regarding the type of the system.

Calculate $P'(\theta)$ via the Bayesian rule according to $a^*(m)$ from (2), $m^*(\theta)$ from (3), and the belief P . If $P'(\theta)$ and $P(\theta|m)$ are not in conflict, then the refined Bayesian equilibrium solution is $EQ = (m^*(\theta), a^*(m), P'(\theta|m))$.

The following two instances of the forward direction and the reverse direction of the signaling game demonstrate the process above. The defender type is denoted as $\Theta_D = (\theta_N, \theta_H) = (Normal\ sys, Honeypot\ sys)$, and the signal corresponds to the defender type, namely, $M_D = (m_N, m_H) = (NomlSysSig, HonSysSig)$. In addition, the defensive strategy set is $d = \{d_g | g = 1, 2, \dots\}$, and the utility function is $U_D(m_j, d_g, a_h, \theta_i)$; the attacker type is denoted as $\Theta_A = (\eta_H, \eta_L) = (AdvAttacker, PrimAttacker)$, with the attack strategy $A = \{a_h | h = 1, 2, \dots\}$, and the utility function is $U_A(m_j, d_g, a_h, \theta_i)$.

3.5. Refined Bayesian Equilibrium Solution Method for the Forward Signaling Game. A game of incomplete information can be transformed into a game of imperfect information by adding a hypothetical player, namely, nature (denoted by C here), and by conditioning the payoffs on Nature’s unknown moves. The nature player moves first by randomly choosing the defender type with the prior probability distribution over all defender types. In the forward direction, nature assigns type θ_N with probability $p_D(\theta_N)$ and type θ_H with $p_D(\theta_H)$. Once the defender has learned her type, she decides what signal or message to send to the attacker. The signal provides indirect information for the attacker about the defender type. In our example, the defender can send either signal m_N (signaling that the defender type is θ_N) or m_H (signaling that the defender type is θ_H). The defender can send signal m_H , even in the case that her real type is θ_N , or send signal m_N , even in the case that her real type is θ_H . The attacker revises his judgement on the defender type and takes action $\{O_N^p, O_H^p\}$ if observing signal m_N and action $\{W_N^p, W_H^p\}$ if observing the signal m_H , as the posterior probability for the defender

TABLE 2: Attack strategy description.

No.	Basic attack option	Attack strategy	
		A_1	A_2
a_1	Remote buffer overflow	✓	✓
a_2	homepage attack	✓	
a_3	LPC to LSASS process	✓	✓
a_4	Apache chunk overflow		✓
a_5	Steal account and crack it	✓	
a_6	Oracle TNS listener		✓

TABLE 3: Defense strategy description.

No.	Basic defense option	Defense strategy	
		D_1	D_2
d_1	Honey file	✓	✓
d_2	Honey account	✓	✓
d_3	Using honeycomb	✓	
d_4	Uninstall delete Trojan		✓
d_5	Limit access to MDSYS		✓
d_6	Web app honeypot	✓	

type $\{\theta_N, \theta_H\}$. In the game tree, a_{ij} indicates eight outcomes, which results in a corresponding payoff. The forward signaling game tree $G_{DS}(F)$ is presented in Figure 3.

3.6. Refined Bayesian Equilibrium Solution Method for the Reverse Signaling Game. Nature moves first by randomly choosing the attacker type with the prior probability distribution over the attacker types. The reverse signaling game tree $G_{DS}(R)$ is presented in detail in Figure 4.

TABLE 4: Notation descriptions.

Notation	Description
$U_A(m_j, d_g, a_h, \theta_i)$	Expected utility function of the attacker
$U_D(m_j, d_g, a_h, \theta_i)$	Expected utility function of the defender
C_a : cost of attack	Cost of the attacker using various attack measures
C_d : cost of defense	Cost of the defender using various defense measures
$C_{sc}(d_g, a_h)$: cost of system compromised	System loss cost function with the defensive strategy d_g and the attack strategy a_h as parameters, which indicates the loss to the defender's system when it is compromised, namely, the benefit to the attacker of successfully compromising the system
C_{ds} : cost of deception signal	Cost of a signal using deception, namely, the cost that is incurred by the defender in sending a spoofing signal that does not match its type to deceive the attacker

According to the definition, $m^* = m_a^*(a^*, \eta) = m_a^*(\eta)$ indicates that attacker η sends signal m_a^* and chooses strategy $a^*(m_a^*)$ according to the signal, which is denoted as $m_a^*(\eta)$; $a^* = a^*(d_g, m_a) = a^*(m_a)$ indicates the defender's responding action $a^*(d_g, m_a)$, which is denoted as $a^*(m_a)$; $P' = P'_D(\eta | m_a) = P'_D$ indicates that the defender calculated $P'_D(\eta | m_a)$ as the posterior probability for the attacker type, which is denoted as P'_D ; and the existence of a refined Bayesian equilibrium is denoted as $EQ = (m_a^*(\eta), a^*(m_a), P'_D)$. Based on the two examples above and the algorithm in [26], the optimal strategy selection algorithm for the two-way signaling game model is presented as Algorithm 1.

4. Simulation Results and Analysis

4.1. Simulation Environment. To evaluate the proposed attack-and-defense signaling game model and the algorithm for optimal strategy selection, we construct the simulation environment illustrated in Figure 5.

4.2. Calculating the Utility. According to Richard [27], common vulnerability [28] and the database of attack-and-defense behaviors from MIT [29], attack strategies that are composed of basic options are listed in Table 2.

Common defense strategies with deception that are composed of basic operations are described in Table 3.

For selecting the optimal strategy more scientifically and intuitively, the most basic approach is to quantify the utilities of the strategies that are selected by the defender and the attacker. In this paper, we utilize the scheme that was proposed by Zhang and Li [30] to calculate the expected utility functions of the defenders and the attackers as follows:

$$U_A(m_j, d_g, a_h, \theta) = \sum_{g,h} C_{sc}(d_g, a_h) - C_a, \quad (1)$$

$$U_D(m_j, d_g, a_h, \theta_i) = \sum_{g,h} C_{sc}(d_g, a_h) - C_d - C_{ds}. \quad (2)$$

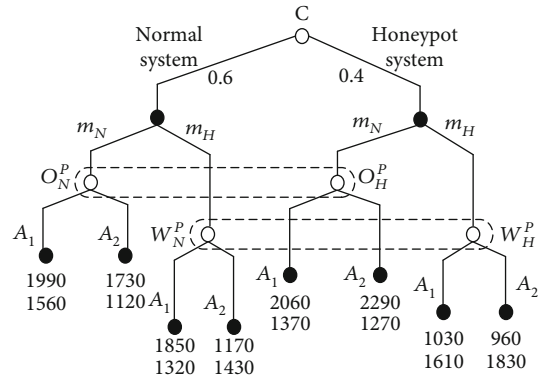


FIGURE 6: Forward signaling game tree with the calculated utilities.

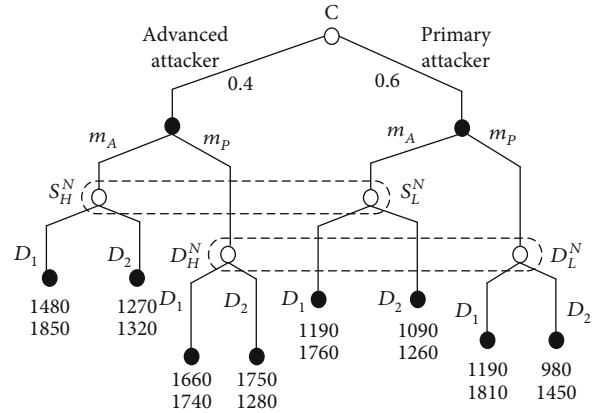


FIGURE 7: Reverse signaling game tree with the calculated utilities.

The notations that are used in equations (1) and (2) are described in Table 4.

For the defender type $\{\theta_N, \theta_H\}$, the defense strategy is assumed to be $D_1\{d_1, d_2\}$ or $D_2\{d_5, d_6\}$, and for the attacker type $\{\eta_H, \eta_L\}$, the attack strategy is $A_1 = \{a_1, a_2\}$ or $A_2 = \{a_4, a_6\}$. Based on historical data and experience,

$C_a = \{C_{A_1}, C_{A_2}\} = \{590, 320\}$, $C_d = \{C_{D_1}, C_{D_2}\} = \{360, 285\}$, and $C_{ds} = \{C_{D_1}, C_{D_2}\} = \{20, 10\}$.

TABLE 5: Possible equilibria in the forward direction.

Condition	Equilibrium	Type of equilibrium
$O_N^p > P_A'(\theta_N m_N), W_N^p > P_A'(\theta_H m_H)$	EQ = $[(m_H, m_H) \rightarrow (A_1, A_1), O_N^p = 0.7, W_N^p = 0.5]$	Pooling equilibrium
$O_N^p > P_A'(\theta_N m_N), W_N^p < P_A'(\theta_H m_H)$	EQ = $[(m_H, m_H) \rightarrow (A_1, A_2), O_N^p = 0.7, W_N^p = 0.5]$	Pooling equilibrium
$O_N^p < P_A'(\theta_N m_N), W_N^p > P_A'(\theta_H m_H)$	EQ = $[(m_H, m_N) \rightarrow (A_2, A_1), O_N^p = 1, W_N^p = 0]$	Separating equilibrium
$O_N^p < P_A'(\theta_N m_N), W_N^p < P_A'(\theta_H m_H)$	EQ = $[(m_N, m_N) \rightarrow (A_2, A_2), O_N^p = 0.7, W_N^p = 0.5]$	Separating equilibrium

TABLE 6: Possible equilibria in the reverse direction.

Condition	Equilibrium	Type of equilibrium
$S_H^N > P_D'(\eta_N m_A), D_H^N > P_D'(\eta_L m_p)$	EQ = $[(m_A, m_A) \rightarrow (D_1, D_1), S_H^N = 0.4, D_H^N = 0.7]$	Pooling equilibrium
$S_H^N > P_D'(\eta_N m_A), D_H^N < P_D'(\eta_L m_p)$	EQ = $[(m_p, m_A) \rightarrow (D_1, D_2), S_H^N = 0.4, D_H^N = 0.7]$	Pooling equilibrium
$S_H^N < P_D'(\eta_N m_A), D_H^N > P_D'(\eta_L m_p)$	EQ = $[(m_A, m_p) \rightarrow (D_2, D_1), S_H^N = 0, D_H^N = 1]$	Separating equilibrium
$S_H^N < P_D'(\eta_N m_A), D_H^N < P_D'(\eta_L m_p)$	EQ = $[(m_p, m_p) \rightarrow (D_2, D_2), S_H^N = 0.4, D_H^N = 0.7]$	Pooling equilibrium

TABLE 7: Comparison of approaches in terms of static or dynamic type, complete information or incomplete information, and signal direction.

Author	Type of game	Number of players	Signal direction
Wang et al. [31]	Complete information static	N	Single
Lin et al. [32]	Incomplete information static	3	Single
Zhang et al. [33]	Incomplete information dynamic	3	Single
Zhu et al. [34]	Incomplete information dynamic	N	Single
Our work	Incomplete information dynamic	N	Two-way

To calculate the utility of the forward-direction signaling game, we set

$$\begin{aligned}
p_D(\theta_N) &= 0.6, \\
p_D(\theta_H) &= 0.4, \\
O_N^p + O_H^p &= 1, \\
W_N^p + W_H^p &= 1.
\end{aligned} \tag{3}$$

All the utilities that are specified in Figures 6 and 7 were calculated via equations (1) and (2).

The posterior inferences can be constructed on various sets of information. Via Algorithm 1, we obtain possible equilibria in the forward direction, as presented in Table 5.

To calculate the utility of the reverse-direction signaling game, we set $p_A(\eta_H) = 0.4$, $p_A(\eta_L) = 0.6$, $S_H^N + S_L^N = 1$, and $D_H^N + D_L^N = 1$.

The posterior inferences that can be constructed on the two sets of information are $P_D'(\eta_H | m_A) = 0.46$ and $P_D'(\eta_L | m_p) = 0.65$. Via Algorithm 1, we obtain the possible equilibria in the reverse direction, which are presented in Table 6.

The algorithm proposed, and the game simulated in the paper is compared with other approaches in Table 7. We have analyzed both directions of signal transmission in a dynamic incomplete information game, which is more in line

with the actual attack-and-defense scenario, and the results can guide the defense decision much more precisely.

4.3. Result Analysis. By implementing the simulation above, we obtain the following results:

- (1) In the forward-signaling game model, if $(P_A'(\theta_N | m_d), P_A'(\theta_H | m_d))$ and (S_H^N, D_H^N) do not conflict, the refined Bayesian equilibrium is a pooling equilibrium. Hence, the defender chooses a honey system and releases the honey system signal, which deceives the attacker, thereby influencing the attacker's judgment on the defender type and on the choice of attack strategy. Thus, the defender uses the signal to demonstrate a capability that exceeds the actual capability, thereby reducing the likelihood of suffering a loss
- (2) In the reverse-signaling game model, the attacker moves first. He can be of type η_H and send signal m_A (presenting himself as an advanced attacker) or m_p (pretending to be the primary attacker). He can also be of type η_L and send the signal m_p (presenting himself as the primary attacker) or m_A (pretending to be an advanced attacker). According to Table 6, the refined Bayesian equilibrium is realized when the advanced attacker pretends to be the primary attacker and the defender chooses strategy D_1 with the deception technique. The advanced attacker

deliberately presents weak attack capabilities so that the defender will reduce the level of defense. However, the choice of the deception defense strategy by the defender can be used to increase the defense utility

- (3) From the perspective of utility for both the defender and the attacker in a two-way signaling game, regardless of whether the attacker's ability is low or high, the choice of the deception defense strategy would increase the payoff of the defender compared with the normal system without deception. The defense strategy with deception is the optimal strategy for the defender. Therefore, the defender would choose the deceptive strategy, namely, the normal system would be disguised as a honeypot

5. Conclusions

We model the confrontation between a defender and an attacker by utilizing signaling game theory. Additionally, we propose the concept of a two-way signaling game and propose an algorithm for identifying optimal defense strategies. Finally, we conduct an extensive simulation analysis to evaluate the performance of the proposed approaches by fortifying the attack-and-defense confrontation in a two-way signal releasing mechanism and calculating the utilities for both sides.

This paper mainly proposes a proactive defense mechanism that utilizes signal selection and release methods and does not consider other defense mechanisms. There are several limitations in our methods, one is that the expected utility functions used in equations (1) and (2) could not be extended to multistage games, and another is that the example shown in the simulation part did not consider the synchronous affect between the attacker and the defender during the game, both of which will be studied in the future work. However, the proposed two-way signaling game model is of substantial importance for subsequent research in the IoT network security. For example, with the method proposed, the defender of the IoT network could infer the optimal strategy of the attacker and take action such as improving the protection level in advance to defense attacks. In the future, we will integrate the analysis via mathematical description, implement the attack-and-defense model for multiple stage games, and explore the security defense decision-making method in IoT networks.

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 is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 61602515).

References

- [1] M. A. Khan and K. Salah, "IoT security: review, blockchain solutions, and open challenges," *Future Generation Computer Systems*, vol. 82, pp. 395–411, 2018.
- [2] Y. M. P. Pa, S. Suzuki, K. Yoshioka, and T. Matsumoto, "IoTPOT: analysing the rise of IoT compromises," in *9th USENIX Workshop on Offensive Technologies (WOOT 15)*, pp. 9–10, Washington, D.C, 2015.
- [3] M. H. Almeshekeh and E. H. Spafford, *Cyber security deception*, Springer International Publishing, 2016.
- [4] Y. M. P. Pa, S. Suzuki, K. Yoshioka, T. Matsumoto, T. Kasama, and C. Rossow, "IoTPOT: a novel honeypot for revealing current IoT threats," *Journal of Information Processing*, vol. 24, no. 3, pp. 522–533, 2016.
- [5] G. Briskin, D. Fayette, N. Evancich, V. Rajabian-Schwartz, A. Macera, and J. Li, "Design considerations for building cyber deception systems," in *Cyber Deception*, S. Jajodia, V. Subrahmanian, V. Swarup, and C. Wang, Eds., pp. 69–95, Springer, Cham, 2016.
- [6] L. Y. Shi, J. N. Zhao, Q. Li, W. Xiao, and L. Xin, "Signaling game analysis and simulation on network decoy defense strategies," *Journal of System Simulation*, vol. 28, no. 2, pp. 348–353, 2016.
- [7] N. C. Rowe, E. J. Custy, and B. T. Duong, "Defending cyberspace with fake honeypots," *Journal of Computers*, vol. 2, no. 2, pp. 25–36, 2007.
- [8] N. Garg and D. Grosu, "Deception in honeynets: a game-theoretic analysis," in *2007 IEEE SMC Information Assurance and Security Workshop*, pp. 107–113, West Point, NY, USA, 2007.
- [9] S. Dowling, M. Schukat, and H. Melvin, "A ZigBee honeypot to assess IoT cyberattack behaviour," in *2017 28th Irish Signals and Systems Conference (ISSC)*, pp. 1–6, Killarney, Ireland, 2017.
- [10] H. Šemić and S. Mrdovic, "IoT honeypot: a multi-component solution for handling manual and Mirai-based attacks," in *2017 25th Telecommunication Forum (TELFOR)*, pp. 1–4, Belgrade, Serbia, 2017.
- [11] F. Cohen, "A note on the role of deception in information protection," *Computers and Security*, vol. 17, no. 6, pp. 483–506, 1998.
- [12] T. E. Carroll and D. Grosu, "A game theoretic investigation of deception in network security," *Security and Communication Networks*, vol. 4, no. 10, pp. 1162–1172, 2011.
- [13] J. Pawlick and Q. Zhu, "Deception by design: evidence-based signaling games for network defense," 2015, <https://arxiv.org/abs/1503.05458>.
- [14] J. Duan, D. Gao, D. Yang, C. H. Foh, and H. H. Chen, "An energy-aware trust derivation scheme with game theoretic approach in wireless sensor networks for IoT applications," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 58–69, 2014.
- [15] S. Fugate and K. Ferguson-Walter, "Artificial intelligence and game theory models for defending critical networks with cyber deception," *AI Magazine*, vol. 40, no. 1, pp. 49–62, 2019.

- [16] H. Çeker, J. Zhuang, S. Upadhyaya, Q. D. Ia, and B. H. Soong, "Deception-Based Game Theoretical Approach to Mitigate DoS Attacks," in *Decision and Game Theory for Security. GameSec 2016*, Q. Zhu, T. Alpcan, E. Panaousis, M. Tambe, and W. Casey, Eds., vol. 9996 of Lecture Notes in Computer Science, Springer, Cham, 2016.
- [17] H. Sedjelmaci, S. M. Senouci, and T. Taleb, "An accurate security game for low-resource IoT devices," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9381–9393, 2017.
- [18] A. Schlenker, O. Thakoor, H. Xu et al., "Deceiving cyber adversaries: a game theoretic approach," in *The 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, vol. 9996, pp. 18–38, Stockholm, Sweden, 2016.
- [19] S. Dey, "Securing majority attack in blockchain using machine learning and algorithmic game theory: a proof of work," in *2018 10th Computer Science and Electronic Engineering (CEECS)*, pp. 7–10, Colchester, UK, 2018.
- [20] J. Pawlick and Q. Y. Zhu, "Quantitative models of imperfect deception in network security using signaling games with evidence [IEEE CNS 17 Poster]," in *2017 IEEE Conference on Communications and Network Security (CNS)*, pp. 394–395, Las Vegas, NV, USA, 2017.
- [21] K. Hua, X. Liu, Z. Chen, and M. Liu, "A Game Theory Based Approach for Power Efficient Vehicular Ad Hoc Networks," *Wireless Communications and Mobile Computing*, vol. 2017, 9 pages, 2017.
- [22] J. M. Huang and H. W. Zhang, "A method for selecting defense strategies based on stochastic evolutionary game model," *Acta Electronica Sinica*, vol. 46, no. 9, pp. 2222–2228, 2018.
- [23] X. Chen, A. Li, X. Zeng, W. Guo, and G. Huang, "Runtime model based approach to IoT application development," *Frontiers of Computer Science*, vol. 9, no. 4, pp. 540–553, 2015.
- [24] S. Shen, L. Huang, H. Zhou, S. Yu, E. Fan, and Q. Cao, "Multistage signaling game-based optimal detection strategies for suppressing malware diffusion in fog-cloud-based IoT networks," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1043–1054, 2018.
- [25] W. Y. Zhang, *Game Theory and Information Economics*, Shanghai People's Publishing House, 2004.
- [26] W. Guo, J. Chen, G. Chen, and H. F. Zheng, "Trust dynamic task allocation algorithm with Nash equilibrium for heterogeneous wireless sensor network," *Security and Communication Networks*, vol. 8, no. 10, pp. 1865–1877, 2015.
- [27] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das, "Analysis and Results of the 1999 DARPA Off-Line Intrusion Detection Evaluation," in *Recent Advances in Intrusion Detection. RAID 2000*, H. Debar, L. Mé, and S. F. Wu, Eds., vol. 1907 of Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2000.
- [28] A. Feutrill, D. Ranathunga, Y. Yarom, and M. Roughan, "The effect of common vulnerability scoring system metrics on vulnerability exploit delay," in *2018 Sixth International Symposium on Computing and Networking (CANDAR)*, pp. 1–10, Takayama, Japan, 2018.
- [29] L. Gordon, M. Loeb, W. Lucyshyn, and R. Richardson, "2005 CSI/FBI computer crime and security survey," *Computer Security Journal*, vol. 21, no. 3, 2005.
- [30] H. W. Zhang and T. Li, "Optimal active defense based on multi-stage attack-defense signaling game," *Acta Electronica Sinica*, vol. 45, no. 2, pp. 431–439, 2017.
- [31] J. D. Wang, D. K. Yu, and H. W. Zhang, "Active defense strategy selection based on static bayesian game," *Journal of Xidian University*, vol. 43, no. 1, pp. 144–151, 2016.
- [32] W. Q. Lin, H. Wang, and J. H. Liu, "Research on active defense technology in network security based on non-cooperative dynamic game theory," *Journal of Computer Research and Development*, vol. 48, no. 11, pp. 306–316, 2014.
- [33] H. W. Zhang, D. K. Yu, J. H. Han, J.-D. Wang, and T. Li, "Defense policies selection method based on attack-defense signaling game model," *Journal of Communication*, vol. 37, no. 5, 2016.
- [34] J. M. Zhu, B. Song, and Q. F. Huang, "Evolution game model of defense for network security based on system dynamics," *Journal on Communications*, vol. 35, no. 1, pp. 54–61, 2015.

Research Article

A Host-Based Anomaly Detection Framework Using XGBoost and LSTM for IoT Devices

Xiali Wang^{1,2} and Xiang Lu^{1,2}

¹*Institute of Information Engineering, CAS, 100093, China*

²*School of Cyber Security, UCAS, 100049, China*

Correspondence should be addressed to Xiang Lu; luxiang@iie.ac.cn

Received 28 March 2020; Revised 16 June 2020; Accepted 9 July 2020; Published 5 October 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Xiali Wang and Xiang Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Internet of Things (IoT) is rapidly spreading in various application scenarios through its salient features in ubiquitous device connections, ranging from agriculture and industry to transportation and other fields. As the increasing spread of IoT applications, IoT security is gradually becoming one of the most significant issues to guard IoT devices against various cybersecurity threats. Usually, IoT devices are the main components responsible for sensing, computing, and transmitting; in this case, how to efficiently protect the IoT device itself away from cyber attacks, like malware, virus, and worm, becomes the vital point in IoT security. This paper presents a brand new architecture of intrusion detection system (IDS) for IoT devices, which is designed to identify device- or host-oriented attacks in a lightweight manner in consideration of limited computation resources on IoT devices. To this end, in this paper, we propose a stacking model to couple the Extreme Gradient Boosting (XGBoost) model and the Long Short-Term Memory (LSTM) model together for the abnormal state analysis on the IoT devices. More specifically, we adopt the system call sequence as the indicators of abnormal behaviors. The collected system call sequences are firstly processed by the famous n -gram model, which is a common method used for host-based intrusion detections. Then, the proposed stacking model is used to identify abnormal behaviors hidden in the system call sequences. To evaluate the performance of the proposed model, we establish a real-setting IP camera system and place several typical IoT attacks on the victim IP camera. Extensive experimental evaluations show that the stacking model has outperformed other existing anomaly detection solutions, and we are able to achieve a 0.983 AUC score in real-world data. Numerical testing demonstrates that the XGBoost-LSTM stacking model has excellent performance, stability, and the ability of generalization.

1. Introduction

As the commercial 5G rolls out, innumerable studies [1–3] predict that a bump from 5G technology will spur IoT innovations to be an integral part of our economy and lifestyle [4], by virtue of its technical advantages in greater transmission speed, lower latency, and a greater number of device connections [5]. In such an exciting 5G era, the full potential of IoT will burst out in the way of Industrial IoT (IIoT) [6], Internet of Medical Things (IoMT) [7], and so on, by planting various sensors into cars, IP cameras, actuators, cardiac pacemaker, and insulin pumps. Gartner, Inc. reported that about 8.4 billion IoT endpoints were deployed in 2017 and will reach 20.4 billion by

2020 [8]. By connecting objects, IoT is bringing tremendous changes to the world, from significant productivity improvement, accurate predictive analysis, and rapid response to the innovation of business models.

Yet, along with all kinds of benefits brought by IoT, it also represents a growing attack surface and exposes its operators, managers, and customers to high-profile cyber threats [9], which are frequently inducing shocking IoT security breach incidents. Kaspersky said that the number of IoT attacks in the first half of 2019 is 9 times more than that in the same period of 2018 [10]. In such attacks, IoT devices, especially those exposed directly on the Internet, like IP cameras, Wi-Fi routers, and smart home facilities, are always the

principal victims. According to the latest released report by Palo Alto Networks [11], 57% of IoT devices are vulnerable to medium- or high-severity attacks, making IoT the low-hanging fruit of attackers. For example, the most notorious malware targeting IoT devices, Mirai, was first found in August 2016, which launched historically distributed denial-of-service (DDoS) attacks globally in 2016 by compromising IoT devices with weak passwords towards a large-scale IoT botnet. Since Mirai's source code was published online [12], variants of Mirai with customized functionalities, as well as new exploits, keep evolving to be a continuing threat [13].

We can see that, in almost every IoT security incidents, the fragile IoT terminal devices are always the Achilles' heel suffering from unencrypted traffics, outdated software, and neglected security designs. Therefore, the security issues of IoT endpoints, in terms of exploits, password guessing, IoT worms, and so on, should be one of the top considerations in the IoT threat landscape. How to secure massively deployed IoT devices should be the primary question to be answered by cybersecurity researchers.

To protect IoT endpoint devices against all kinds of orchestrated attacks, researchers from both academia and industry designed a large number of mechanisms in recent years. In this paper, we propose an abnormal state detection framework to indicate ongoing exploit-compromised or malware-infected attack behaviors happening on IoT devices, thereby identifying those eccentric endpoints. The basic idea of our framework is on the assumption that an IoT device's behaviors are commonly predictive, repetitive, and periodic in statistics [14]. If attacks occur, the compromised IoT devices are prone to showing some unusual behaviors in different manners, like unexpected traffics and strange device actions. Based on this idea, we focus on abnormal behaviors of embedded operating systems to identify malicious IoT device attacks and leverage the statistical analysis of system calls to indicate abnormal system behaviors from device attacks. To this end, we firstly set up an abnormal system call dataset captured in different attack scenarios. Then, to achieve more accurate results of attack indications, we then develop a fusion model for system call sequence analysis by coupling a machine learning model (XGBoost) and a deep learning model (LSTM neural network) together using stacking theory. Our experimental results demonstrate that the system call trace is a good indicator to identify abnormal behaviors from normal behaviors, and the performance of our anomaly detection framework exceeds the other excellent machine learning models.

From the above, we describe our contributions in this paper as follows:

- (1) We choose system call as the indicators and deploy several types of attacks to collect normal behaviors and abnormal behaviors in IoT devices
- (2) We propose an anomaly detection framework and design a fusion model using the stacking method. As for the base models, we select the XGBoost algorithm and LSTM neural network. Moreover, the metamodel is the logistic regression model

- (3) We evaluate the performance of our anomaly detection framework, and the results show that it is valid and stable; moreover, it has good generalization

The following paper is organized as follows. Section 2 introduces related works regarding the intrusion detection of IoT devices. Section 3 illustrates the proposed anomaly-based intrusion detection framework, as well as methods designed for data processing and analysis. In Section 4, we detail the experiment setup and discuss the experimental results. Section 5 concludes our works in this paper.

2. Related Works

As a classical cybersecurity research area, anomaly detection-based intrusion detection system designs keep evolving with fruitful literatures. In general, according to the information used for analysis and employed techniques for the decision-making of deviation away from normal behavior, the intrusion detection can be divided into five groups [15], that is, statistical methods, rule-based methods, distance-based methods, profiling methods, and model-based approaches. In these years, IoT device-oriented IDS is becoming the new hot topic in the classical area. With the rapid advance of artificial intelligence (AI) technologies, we can see an obvious trend; more and more IoT device-oriented IDS mechanisms are established on various AI models to achieve efficient feature extraction and accurate behavior patterns. Although many mechanisms turn to AI models, there also exist many differences between them, which mainly originate from data types used for anomaly indication and models used for data analysis.

[16] developed a model for intrusion detection in IoT microservices by using clustering models like *K*-means and BIRCH, based on the traffic dataset captured from 4 different emulated IoT sites. Similar to captured network traffics in [16, 17] adopted a deep learning model, named the Dense Random Neural Network, to detect security breaches in a smart home application scenario. In [18], the authors proposed a host-based intrusion detection system on the basis of a machine learning approach. In their designs, the immunity-inspired algorithms are involved to distinguish whether the current behavior patterns are matching with desired ones; if not, it means that the attack is happening. Also, for the smart home IoT scenario, [19] implemented a supervised intrusion detection system. For the sake of performance evaluations of the implemented system, the authors deployed 12 attacks, which belong to 4 threat types, including denial of service (DoS) attack, man-in-the-middle (MITM)/spoofing attack, reconnaissance attack, and replay attack.

Despite the fact that AI-based IoT IDS schemes are the dominant techniques, there is also a lot of work to do towards efficient and accurate anomaly detection on IoT endpoints. In existing literatures, network traffic data is the commonly used indicators to identify device attacks, since it is easy to be collected. Therefore, they are actually the so-called network-based IDS, not the host-based IDS [20]. However, in the real-setting IoT IDS scenarios, the salient features of IoT devices on massive proprietary protocols, uninterrupted

running, and sometimes encrypted traffics will significantly deteriorate the network-based IDS. Thus, the host-based IDS framework, as proposed in this paper, is indispensable for IoT security.

3. System Framework and Scheme Design

In this section, we firstly introduce the outline of the proposed anomaly-based intrusion detection framework, including explanations of the main building blocks. Then, we illustrate critical approach design used in data processing and model analysis.

3.1. Methodology Outline. As presented above, this work is aimed at proposing an anomaly-based intrusion detection framework to identify ongoing attacks on IoT devices. To achieve this objective, we leverage system call sequences as the anomaly indicator to capture malicious system behaviors brought by vulnerability exploitations or malware infections. Figure 1 shows the main building blocks of the proposed anomaly-based IoT IDS framework, which is also the basic idea outline of this paper.

To make the framework implemented as the running intrusion detection module embedded in the IoT device, we have to finish several critical scheme designs, regarding the construction of training data set, the system call data preprocessing, the data analysis model design, and performance evaluations, as shown in Figure 1, all of which will be illustrated in details in the following subsections.

3.2. Training Dataset Constructions. Different from network-based IoT IDS, which mainly depends on the captured network traffic data for data analysis and modelling, in our framework design, we leverage system call sequences as the anomaly indicator for intrusion detection inside the IoT device, also called as host-based IDS. Therefore, how to collect system call data, especially in different attack scenarios, becomes the primary issue in our framework. To construct such a training dataset, we tried to look for IoT system call sequence dataset under attack conditions, but there is no existing one. In this case, we turn to embedded Linux for system call sequence data collection, which is the most popular operating system used in IoT devices [21], taking about 21% market shares globally. More specifically, we program shell scripts by leveraging the “strace” command [22] and the “crond” mechanism [23] for system call sequence collection when different attacks are loaded. Among these two, the “strace” command is usually used as the system call tracer, while the “crond” mechanism is designed to execute scheduled commands. With the shell script, we are able to capture traces of all processes (except for “strace” its own process) and stored them in the file while the attack is deployed.

In terms of typical IoT attacks, we choose a HiSilicon Hi3516CV300-based IP camera as the example to deploy attacks. For the dataset construction, we collect system call sequence data in 5 classes, as shown in Table 1. Class 0 data is captured in the normal state used to indicate expected system behaviors when the IP camera is running, whose main functions include scene sensing, video coding, video transmis-

sion, and even web access. Class 1 data is acquired when a vulnerability exploit is executed. The vulnerability is from CVE-2016-5195 [24], also known as “Dirty COW,” which allows local users to gain privileges by leveraging incorrect handling of a copy-on-write feature to write to a read-only memory mapping. Class 2 data is collected when a notorious malware, BASHLITE [25], is infecting devices. Class 3 data is from an emulated abnormal (malicious) operation on the IP camera device; that is, we add a user and set the password on the device. Class 4 data is from a memory leak attack, in which we keep creating threads to request connections to the embedded Real-Time Streaming Protocol (RTSP) stack and lead to memory leak and a fatal firmware error.

As the way in Table 1, we totally collect about 5.9 million system calls in different classes. We then use the dataset to verify the effect and efficiency of the framework proposed in this paper.

3.3. Feature Extraction and Selection. The objective of data preprocessing in this work is to extract features from seemingly clueless system call sequences for the following anomaly decision-making. Figure 2 shows an example of a system call sequence. To achieve efficient feature extractions, we take 2 key steps for dataset sorting.

The first step is to digitalize the dataset by employing the system call table to replace the system call function with a unique system call number, while ignoring impacts of parameters. For example, after the digitalization, the sequence shown in Figure 2 turns to the notation of a trace as follows:

$$\text{Trace} = (5, 114, 78, 174). \quad (1)$$

After the digitalization of dataset, in the second step for classification of the behavior from system call traces, we should extract features from the system call traces. Usually, data representation techniques can be used to convert the system call trace into a feature vector. We adopt the famous n -gram model [26] from the Natural Language Processing area, which is also a common data processing method in intrusion detection scheme design [27].

N -gram is a contiguous of n items from a given sample of text or speech; it is an algorithm based on the statistical language model. Its basic idea is to slide the contents of the text into n -size Windows according to bytes, forming a sequence of n -length byte fragments. Each byte fragment is called gram. The occurrence frequency of all grams is counted and filtered according to the threshold set in advance to form the key gram list, namely, the vector eigenspace of the text. Each gram in the list is an eigenvector dimension.

To explain this language model, let us consider the following trace:

$$\text{Trace} = (6, 6, 63, 6, 42, 120, 6, 195, 120, 6, 6, 114, 114). \quad (2)$$

We can set n to different values, such as 3, 5, and 7. If we set n as 5, we can get the following set of sequences:

$$\text{Trace} = (6, 6, 63, 6, 120), (6, 63, 6, 120, 6), \dots, (120, 6, 6, 114, 114). \quad (3)$$

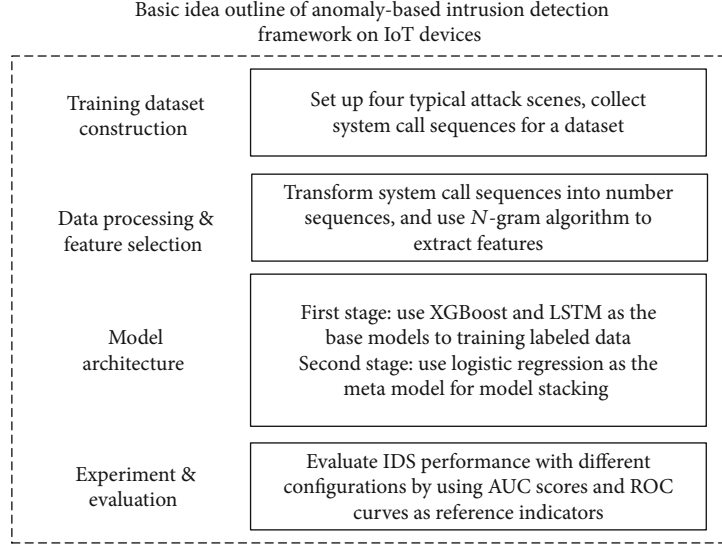


FIGURE 1: Methodology outline of this paper.

TABLE 1: Five classes of collected system call sequence dataset.

No.	Class name	State	Notes
1	Class 0	Normal state	Syscall sequence data in normal state
2	Class 1	Vulnerability exploiting	Syscall sequence data in CVE-2016-5195
3	Class 2	Malware infection	Syscall sequence data in BASHLITE malware
4	Class 3	Abnormal operation	Syscall sequence data in user add operation
5	Class 4	Memory leak	Syscall sequence data in RTSP memory leak

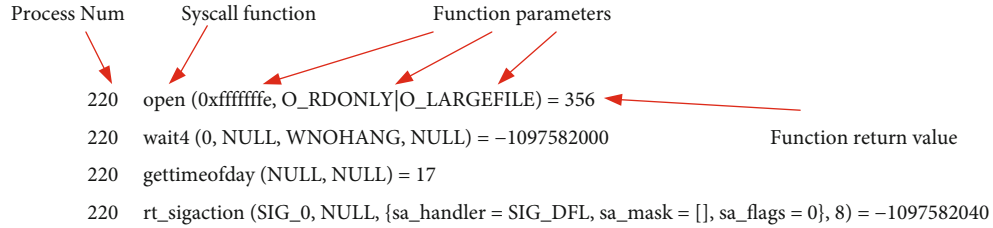


FIGURE 2: An example of a system call sequence.

By considering collected system call traces as a set of document and system calls as words, we can apply classical data representation and classification techniques used in the area of natural language processing (NLP) and information retrieval (IR).

Let us consider a system call trace S with L system calls, the n -gram is defined as tuple $(g_1, g_2, g_3, \dots, g_n)$ of n terms where n is generally a small positive integer. For a given gram, its weight $\omega(g)$ in trace S is given by

$$\omega(g) = \frac{f(g)}{|L|}, \quad (4)$$

where $f(g)$ is the frequency of g in S and $|L|$ is the length of one trace.

Through using the n -gram language model, the syscall number sequences are processed into many subsequence features and build a dictionary which key is the grams and the value is the frequency. These features from the system call traces can be used for normal or abnormal behavior classification.

After we extract the n -gram features and compute the frequency of the weight, it is easy to find that, as n increases, the feature extraction work will become a time-consuming task. So, we must select some important features as the model training features. We set a parameter m ; it represents the top m features that we pick based on frequency. We use the top m features to sketch the system call behaviors in IoT devices and construct models based on these features.

3.4. Extreme Gradient Boosting. In Sections 3.4 and 3.5, we will introduce the machine learning methods we used in

our anomaly detection system. In our anomaly detection framework, we chose the XGBoost (Extreme Gradient Boosting) [26] classifier as one base model of the anomaly detection framework.

XGBoost is a decision-tree-based ensemble machine learning algorithm that uses a gradient boosting algorithm to a known dataset and then classifies the data accordingly. The two reason to use XGBoost are also the two goals of a machine learning project: execution speed and model performance. XGBoost is really fast when compared to other implementations of gradient boosting. From Szilard Pafka's experiment [28], he proved that XGBoost is fast, memory efficient, and of high accuracy. So, we use XGBoost as a part of our anomaly detection architecture to obtain perfect performance.

Here are XGBoost's mathematical explanation. XGBoost is composed of multiple CARTs. CART is known that a basic decision tree can be established on the concept of entropy. The object of CART is Gini coefficient:

$$\text{Obj} : \min \text{Gini}_{\text{index}}(D, a) = \sum_{v=1}^V \frac{|D^v|}{|D|} \text{Gini}(D^v), \quad (5)$$

$$\text{Gini}(D) = \sum_{k=1}^K \sum_{k' \neq k} p_k p_{k'} = 1 - \sum_{k=1}^K p_k^2,$$

where a is one of the attributes we selected, V is the scale of a , v is one of the values of the attribute a , and K is the scale of labels. Intuitively, the Gini coefficient reflects the probability of two samples in the dataset that the labels are different. Moreover, this is the principle to construct one tree.

The goal of XGBoost is to fit the residual. Residual means the difference between the real value and predicted value. XGBoost is defined as an addition model:

$$F(X, w) = \sum_{k=0}^K \alpha_k h_k(X; w_k) = \sum_{k=0}^K f_k(X, w_k), \quad (6)$$

where X is the input data, $F(X, w)$ is the model that we finally get, h_k means one single tree, w is the parameter of the tree, and α_k is the weight of the tree. By minimising the loss function, we can get the optimal model $F(X, w)$. The loss function is defined as

$$F^* = \arg \min_F \sum_{k=0}^K L(Y_i, F(X_i; w_k)),$$

$$\text{loss} = \sum_i l(\hat{Y}_i, Y_i) + \sum_k \Omega(f_k), \quad (7)$$

$$\Omega(f_k) = \gamma N_{\text{leaf}} + \frac{1}{2} \lambda \|w_k\|^2,$$

$$l(\hat{Y}_i, Y_i) = (Y \wedge_i - Y_i)^2,$$

where N_{leaf} means the number of leaf nodes in the decision tree. This is a published measurement to restrain the

complexity of our model. Y_i is the real value and \hat{Y}_i is the predicted value. γ and λ are the parameters.

3.5. Long Short-Term Memory. In our anomaly detection architecture, we also use the LSTM (Long Short-Term Memory) neural network as the other base model in the model stacking algorithm. In this part, we introduce the LSTM neural network and explain why we use this model to make classifications in our anomaly detection architecture.

The LSTM model is an improved version of the RNN model, which overcomes back-propagation gradient dispersion or gradient explosion and is more suitable for processing sequence data with long-term dependence. So, we choose the LSTM model as the other base model to combine with the XGBoost model in order to get a valid and stable anomaly detection system for IoT devices.

Here is the basic theory of the LSTM model. Figure 3 shows a single LSTM cell. And we describe the gate mechanism which is the special characters of the LSTM model and explain the equations to compute the values of three gates and cell state.

Compared with the RNN neural network, the LSTM neural network adds the gate mechanism: input gate, forget gate, and output gate in each neuron cell. It is this mechanism that enables LSTM to solve the gradient disappearance and gradient explosion problems in the long sequence training. Each forget gate reads the output value x_t and h_{t-1} , splices the two values, and retains information through the control of the forget gate. LSTM internal operations mainly include sigmoid, tanh, addition, and multiplication, among which sigmoid and tanh are two different activation functions reserved for the next memory unit. The formula is

$$S(x) = \frac{1}{1 + e^{-x}}. \quad (8)$$

The tanh function is a hyperbolic tangent function that maps any value to the interval $[-1, 1]$. The formula is

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (9)$$

As shown in Figure 2, state preservation is very important in the LSTM network model. C_{t-1} and C_t in Figure 2 represent states that are used for information that needs to be retained over time. The specific training process of LSTM is as follows.

In the first step, LSTM uses the forget gate to determine the retention degree f_t of information. In formula (10), W_f is the weight matrix of the forget gate and b_f is the offset value of the forget gate. This step integrates the previous output value with the current input value and passes it to the next memory unit through the forget gate. If the forget gate outputs zero, the previous memory will be cleared. If the output is one, all the previous content is credited to the memory unit:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (10)$$

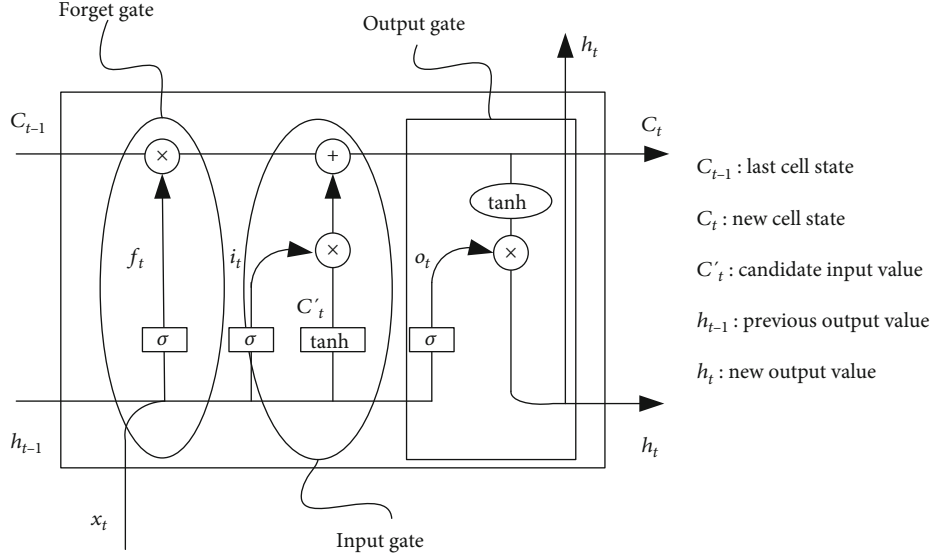


FIGURE 3: LSTM algorithm structure diagram.

In the second step, select memory. This step is used to determine the information held in the cell state. It is mainly divided into the generating temporary new state and updating old state. Assume that c'_t is a temporary new state and i_t is the input gate output to determine which data from the previous step needs to be updated. The old state c_{t-1} multiplies f_t to decide which information to forget and which information to keep. Finally, the new temporary state c'_t multiplies i_t to get the new state c_t . Among these phases,

$$\begin{aligned} c'_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i). \end{aligned} \quad (11)$$

Finally, output. LSTM firstly determines the output information through the sigmoid function and then maps the unit state to $[-1, 1]$ through the tanh function. The output value is obtained by multiplying the sigmoid threshold:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o). \quad (12)$$

In the above steps, W_c , W_i , and W_o are the weight matrixes. The matrix is randomly generated at the beginning of the iteration, and in each process of backward transmission, the information of the previous unit is retained and updated after obtaining the accurate matrix.

3.6. Anomaly Detection Architecture. In this part, we introduce the anomaly detection architecture which is applied in the model building stage of our anomaly detection framework.

So far, we have used the n -gram language model to segment the system call number sequences, selected the top max features in order to solve the problem of too many feature dimensions, and introduced the core principles and main advantages of XGBoost and LSTM neural network. Finally, we use these two models to construct an anomaly

detection architecture which is combined with feature selection of the tree model and sequence memory of the LSTM neural network.

In our anomaly detection architecture, we use the XGBoost and LSTM neural network to integrate as base models. The algorithm of model integration is stacking. It is an ensemble machine learning algorithm and uses a meta-learning algorithm to learn how to best the predictions from two or more base machine learning algorithms and base deep learning algorithms. The benefit of stacking is that it can harness the capabilities of a range of well-performing models on a classification task and make predictions that have better performance than any single model in the ensemble.

Let us simply introduce the stacking theory. The architecture of a stacking model involves two or more base models, often referred to as level 0 levels (base models), and a metamodel that combines the predictions of the base models, referred to as a level 1 model (metamodel). The metamodel is trained on the predictions made by base models on out-of-sample data. That is, data not used to train the base models is fed to the base models, predictions are made, and these predictions, along with the expected outputs, provide the input and output pairs of the training dataset used to fit the metamodel.

The main idea of constructing a predictive model by combining different models can be schematically illustrated in Figure 4.

Figure 4 describes the key information for model stacking. At first, initial training data (X) has m samples and n features. There are M different models that are trained on X (by some method of training, like crossvalidation). Each model provides predictions for the outcome (y) which are then cast into a second level training data X^{l2} which is now $m * M$. Namely, the M predictions become features for this second level data. A second level model (or models) can then be trained on this data to produce the final outcomes which will be used for predictions.

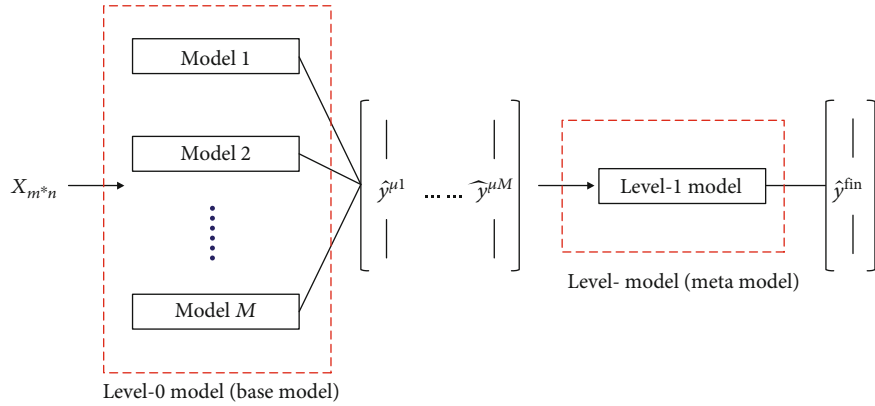


FIGURE 4: The main idea of constructing a predictive model by combining different models.

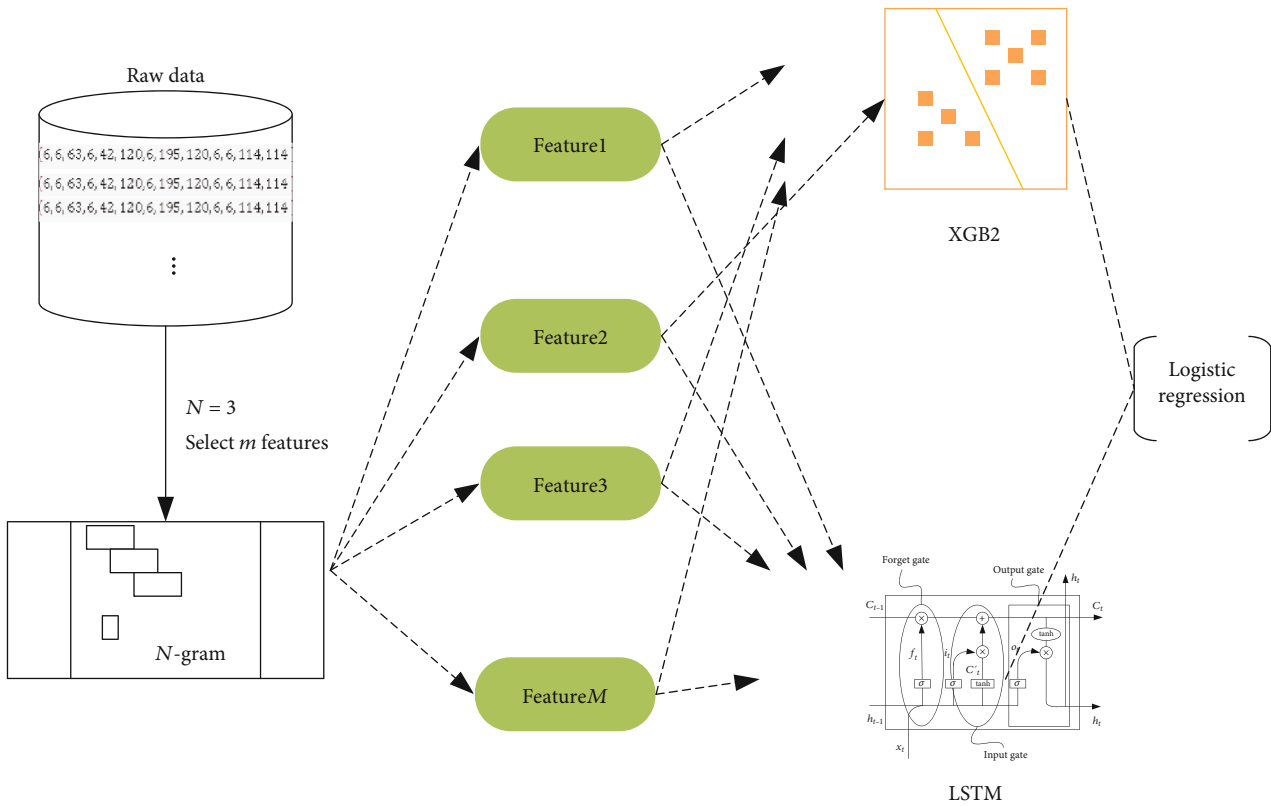


FIGURE 5: Model stacking with XGBoost and LSTM.

Based on the stacking theory, we design a procedure to classify abnormal behavior and normal behavior. The architecture of the procedure is shown in Figure 5, and a detailed description is as follows.

Step 1. (Data collection). We deploy several types of attacks in IoT devices and leverage the “strace” command and the “crond” mechanism for system call sequence collection.

Step 2. (Feature extraction and important feature selection). We use the n -gram model to extract features, and after that,

we select a set of features with top frequencies as the input for the model construction in our framework.

Step 3. (Base model training). Use the selected features; we labelled the dataset: “0” is used to indicate normal behaviors, and “1” is used to indicate abnormal behaviors. In model construction, we use stacking theory to build a fusion model. The base models are XGBoost and LSTM models.

Step 4. (Model stacking using logistic regression). This step uses the logistic regression model to train the predictions

from the above base models; the prediction from the logistic regression model will be considered the final predictions. We can use the prediction to predict the unknown system call trace to tell it normal or abnormal and test the performance of the anomaly detection framework.

Step 5. (Model evaluation and model analysis). We test the performance of the anomaly detection framework with some baselines. After that, we further analyze the model and discuss how some changes affect the performance of our anomaly detection framework.

4. Evaluation

4.1. Experimental Settings. In our experiment, as mentioned before, we used a HiSilicon Hi3516CV300-based IP camera as an example, to deploy the system call sequence data collection program, as well as 5 different device states illustrated in Table 1, for the real-setting data collection.

According to Section 3.2, we construct our training dataset for our anomaly detection framework. In the experiment of the model construction stage, we split the dataset to training data and test data. We use just training normal data in training parse with a normal label, and at test parse, we use all attack data and partial data to measure the performance evaluation. Table 2 summarizes the training data and test data statistics.

After the settings of the experimental environment, we also set the basic model parameters of XGBoost and LSTM. As Table 3 shows, booster is used to determine what model to adopt. We choose “gbtree” for the tree booster. Max_depth is the maximum tree depth for base learners. An objective specifies the learning task, and we use binary:logistic for our framework. Learning_rate makes the model robust by shrinking the weights on each step.

In the LSTM model, there are several parameters which effect the performance of the LSTM model. We choose the parameters as Table 4 shows. The drop rate can slow down with regularization methods. We set three hidden layers to explore additional hierarchical learning capacity. As for the epochs of the training parse, we set it between 50 and 6000. Batch size controls how often to update the weights of the network; we select this parameter from 32 to 256. Last but not least, we set the learning rate to 0.001.

4.2. The Experiment Indicators. As for the classification algorithm, the evaluation indexes are accuracy (ACC), precision-recall curve (PRC), characteristic curve (receiver operating characteristic, ROC), and AUC (Area Under the Curve) of the curve area. Compared with PRC, the advantage of ROC curves and AUC is that they can remain basically unchanged when the positive and negative sample distribution changes in the test data.

The ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. The ROC curve is created by plotting the true positive rate (TPR) against the false

TABLE 2: The training data and the test data statistics.

Dataset	Samples	Numbers of syscall	Class of traces
Training data	4233	4,729,365	Normal
Test data	1048	1,171,359	Attack

TABLE 3: The parameters of XGBoost classifier.

Parameters	Value
Booster	gbtree
Max_depth	4
Objective	Binary:logistic
Eval_metric	AUC
Learning_rate	0.01

TABLE 4: The parameters of LSTM.

Parameters	Value
Dropout rate	0.2
Hidden layers	3
Epoch	50~6000
Batch size	32~256
Learning_rate	0.001

TABLE 5: The comparison with other models.

Method	AUC
SVM	0.924
Isolation forest	0.931
XGBoost	0.966
LSTM	0.978
Fusion model	0.983

positive rate (FPR) at various threshold settings. Let us define the true positive rate and the false positive rate:

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{FPR} &= \frac{\text{FP}}{\text{FP} + \text{TN}}. \end{aligned} \quad (13)$$

In the above computational formula, we consider TP as true positive, FP as false negative, FN as false negative, and TN as true negative. We use the ROC curve and AUC as the evaluation metrics for our anomaly detection framework.

Moreover, in Section 4.4, we also analyze the performance of our approach by using the precision, recall, and F -measure in order to perform a comprehensive performance evaluation. Precision = $\text{TP}/(\text{TP} + \text{FP})$ shows the percentage of true anomalies among all anomalies detected, Recall = $\text{TP}/(\text{TP} + \text{FN})$ measures the percentage of anomalies in the datasets being detected, and F -measure = $2 * \text{Precision} * \text{Recall}/(\text{Precision} + \text{Recall})$ is the harmonic mean of the two.

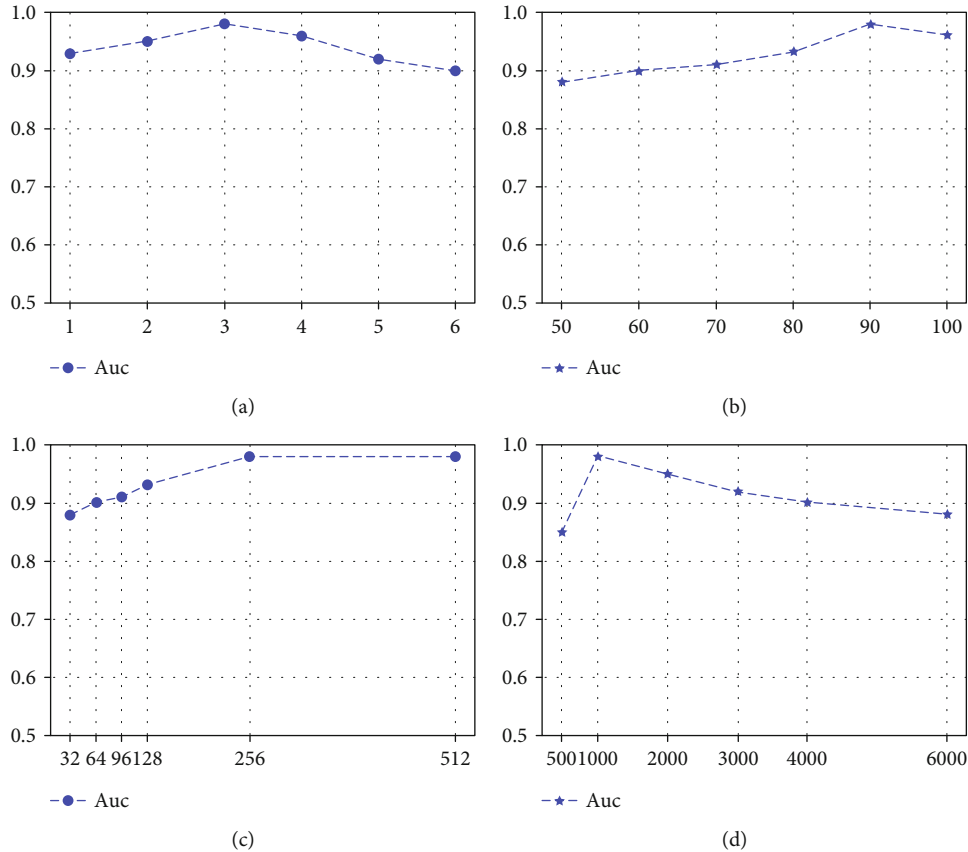


FIGURE 6: The performance with different parameters.

4.3. Comparison with Baselines. For baselines, the general solution for anomaly detection is to use a single machine learning model or approach. So, we select the most popular/-representative anomaly-based approaches which have been used in these research domains. We consider the XGBoost and LSTM as two baselines and also consider the SVM [29] and isolation forest [30] as two baselines. SVM are good at classification problems [31]. The isolation forest is an unsupervised learning algorithm for anomaly detection that works on the principle of isolating anomalies. For our framework, we set the numbers of n -gram in the feature extraction stage (n : between 2 and 7), the numbers of features selecting from the first layers (m : between 50 and 100), the number of batch size (b : between 32 and 512), and the number of training epochs (e : between 500 and 6000). By default, our anomaly detection framework use $n=3$, $m=80$, $b=256$, and $e=1000$ for evaluation and investigate the impacts of these parameters in our experiments.

Table 5 shows the performance of our anomaly detection framework and other popular approaches. The result is shown in Table 5; we can find that XGBoost and LSTM neural network can perform well on our dataset. Clearly, our anomaly framework has achieved the best overall performance by which the AUC score is 0.983.

4.4. Model Analysis. To have a clear understanding of our anomaly detection framework, we conduct a model ablation to observe the effectiveness of the crucial structure and

parameters for performance. By choosing the parameters and using the model stacking, we can achieve excellent performance. The comparative results are shown mainly in the following subsections.

4.5. Discussion of the Performance Impact of Different Parameters in Data Preprocessing Parse. We study the impact of different parameters, including n , the numbers of features in data preprocessing before training in our anomaly detection framework. Figure 6 shows an in-depth comparison using AUC. We change the values of one parameter per time while the other parameter values are default in each experiment. The parameters and their settings are as follows.

For the parameter of the N -gram language model, we take 2 as the initial configuration of the parameter of the n -gram method; then, we use from 3 to 7 to construct a comparative trial. As shown in Figure 1, we compare these n values on our model. It can be summarized that the best method is to use the 3-gram method; the following experiments will use a 3-gram language model to prepare the input data.

The parameter of the number of important features is selected by frequency. We give a range value of l between 50 and 100. We find that when we control the value around 80, the performance of our anomaly detection framework can achieve the best result. It can be seen from the experiment that, when l is less than a certain threshold, the more features, the better result.

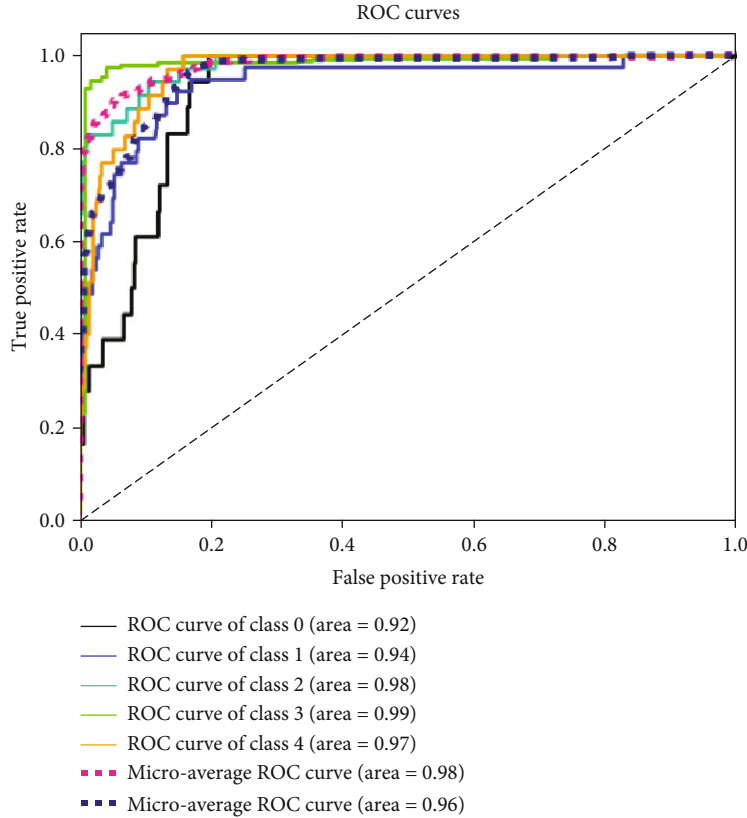


FIGURE 7: The ROC curves of different types of data classes.

In Figures 6(c) and 6(d), we design a systematic test for evaluating LSTM model configurations. We set different values of batch size and epochs in the LSTM classifier to evaluate the impact of these parameters. The result shows that when the batch size is 256, model performance peaked. And when the number of epochs is 1000, our framework can get best performance on the anomaly detection task.

As for the other parameters of our anomaly detection framework, we also tune them for better performance, for example, the k value in k -fold crossvalidation. In the third level of our anomaly detection framework, we use 10 to observe the results because a value of $k = 10$ is very common in the field of machine learning and is recommended by specialists and scholars.

In general, the performance of our anomaly detection framework is relatively stable that the results of experiment are relatively stable when adjusting the parameter values. Also, we get better performance by tuning the key parameters to improve our anomaly detection framework.

4.6. Discussion of the Performance Impact of Different Types of Attacks. To have a clear understanding of how our anomaly detection framework performs on each class of abnormal events, we use one class per time to our fusion model while the other events are null. Figure 7 shows the results of ROC curves of our anomaly detection model on different attacks. We find that some attacks are easy to detect from others, such as the class 2 and class 3 attacks, while some attacks have low probabilities to be detected using our anomaly detection

framework, such as the class 0 (that is normal state) and class 1 attacks.

What is more, from the overall attacks' anomaly detection task and the single attack anomaly detection task, we find that the detection including the overall attacks performs better than all detections on a single abnormal state.

4.7. Discussion of the Performance of Different Datasets. Our anomaly detection framework for IoT devices includes the dataset obtained by ourselves. In order to verify the validity of the dataset and our anomaly detection framework, we use the Australian Defence Force Academy-Linux Dataset (ADFA-LD) for a comparison. ADFA-LD is a similar dataset collecting system call traces on the Linux system for PCs. Figures 8 and 9, respectively, represent the detection performance of our framework on the two datasets. To be specific, Figure 8 shows the result of the dataset constructed by ourselves, and Figure 9 shows the result on the ADFA-LD dataset. From the comparison of performance, our framework can be proved to be effective and accurate. What is more, our framework has the ability of generalization so that it can be used in a similar dataset.

4.8. Discussion of the Performance of Different Metrics. A classifier is only as good as the metric used to evaluate it. A wrong metric to evaluate models will lead us to miss the expected performance of our anomaly detection framework. We compare four metrics which are mainly used for classification tasks, as shown in Table 6. The result shows that different

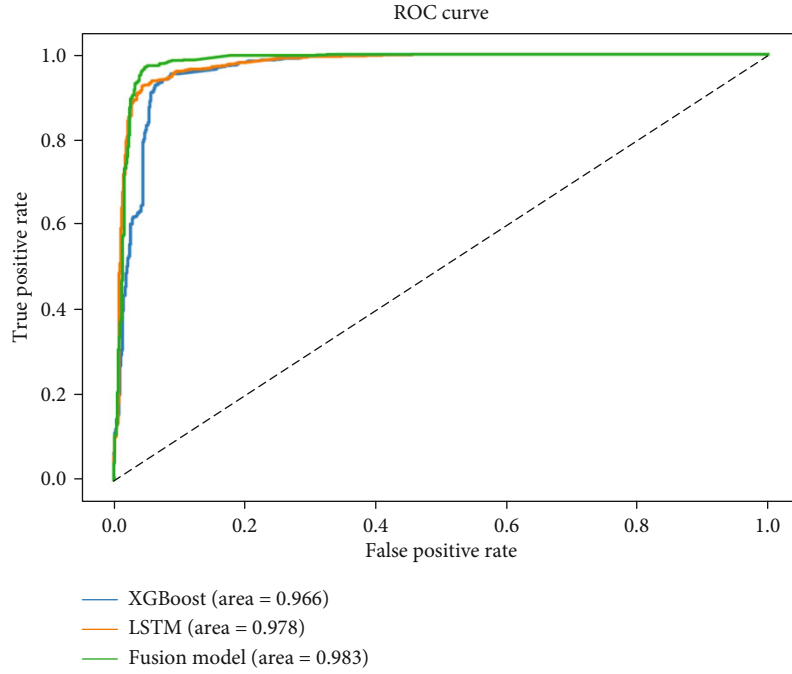


FIGURE 8: Performance on our datasets.

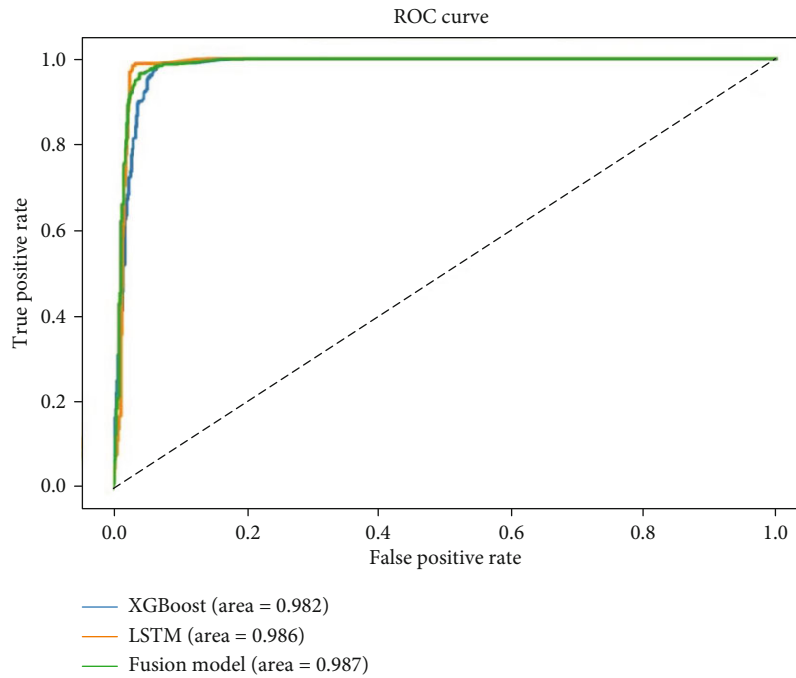


FIGURE 9: Performance on ADFA-LD datasets.

TABLE 6: Different metrics for anomaly detection framework.

Metrics	AUC	Precision	Recall	F-measure
Fusion model	98.3%	97.2%	98.1%	97.9%

evaluation metrics have little influence on performance. It also indicates the robustness of our anomaly detection framework.

5. Conclusion

In this paper, we propose a new anomaly-based intrusion detection framework for IoT devices. In the proposed

framework, we collect system call sequences as a dataset and use a machine learning method XGBoost and deep learning method LSTM to construct a model using stacking idea to detect the abnormal behavior from the normal behavior. The experiments prove that our anomaly detection framework is valid and has better classification performance. We propose a complete anomaly-based intrusion detection process for IoT devices, and it can be used in follow-up studies for scholars in this research field. In the future work, we can also improve the model from the following aspects: adjusting the parameters of this model and improving the diversity of the dataset.

Data Availability

The experiment data used to support the findings of this study were divided into two parts. One part of the data is constructed by authors. These data were designed for evaluation by system call anomaly-based IDS for IoT devices. Requests for access to these data should be made to Xia-Li Wang, xialiwang4@gmail.com. The other part of the data is the ADFA-LD dataset. It is designed for evaluation by system call base HIDS. The dataset can be contacted at <https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-IDS-Datasets/>.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- [1] S. Baghdassarian, *Report Highlight for Market Trends: 5G Opportunities in IoT for Communications Service Providers*, Gartner Research, 2019, <https://www.gartner.com/en/documents/3964479/report-highlight-for-market-trends-5g-opportunities-in-i>.
- [2] M. Rouse, *Internet of Things (IoT), IOT Agenda*, 2019.
- [3] J. Kaplan, *These 3 transformative technologies will have the biggest impact on 2020*, Digital Trends, 2019, <https://www.digitaltrends.com/cool-tech/2020-trends-to-watch-ai-5g-iot/>.
- [4] P. Collela, *5G and IoT: ushering in a new era*, 2020, <https://www.ericsson.com/en/about-us/company-facts/ericsson-worldwide/india/authored-articles/5g-and-iot-ushering-in-a-new-era>.
- [5] IoT Solutions World Congress, *Advantages of 5G and how will benefit IoT*, 2019, <https://www.iotworldcongress.com/advantages-of-5g-and-how-will-benefit-iot/>.
- [6] World Economic Forum, *Accelerating the impact of industrial IoT in small and medium-sized enterprises: a protocol for action*, 2020, <https://www.weforum.org/whitepapers/accelerating-the-impact-of-industrial-iot-in-small-and-medium-sized-enterprises-a-protocol-for-action/>.
- [7] A Whitepaper by Intersog, *What is the Internet of medical things, in the Journal of mHealth*, 2018.
- [8] Gartner Inc, *Gartner says 8.4 billion connected "things" will be in use in 2017*, 2016, <https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016>.
- [9] 5G Americas White Paper, *The future of IoT*, 2019.
- [10] P. Muncaster, *Over 100 million IoT attacks detected in 1H 2019*, 2019, <https://www.infosecurity-magazine.com/news/over-100-million-iot-attacks/>.
- [11] Unit 42, *2020 Unit 42 IoT Threat Report*, Palo Alto Networks, 2020, <https://start.paloaltonetworks.com/unit-42-iot-threat-report>.
- [12] Krebs On Security, *Source code for IoT botnet 'Mirai' Released*, 2016, <https://krebsonsecurity.com/2016/10/source-code-for-iot-botnet-mirai-released/>.
- [13] P. Paganini, *Mirai – the evolving IoT threat*, 2018, <https://resources.infosecinstitute.com/mirai-botnet-evolution-since-its-source-code-is-available-online/#gref>.
- [14] M. Kuniavsky, *User experience and predictive device behaviour in the Internet of Things*, 2015, <https://conferences.oreilly.com/experience-design-iot/public/schedule/detail/42519>.
- [15] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection," *Proceedings of the 2003 SIAM International Conference on Data Mining*, pp. 25–36, 2003.
- [16] M. Pahl and F. Aubet, "All eyes on you: distributed multi-dimensional IoT microservice anomaly detection," in *2018 14th International Conference on Network and Service Management (CNSM)*, pp. 72–80, Rome, Italy, November 2018.
- [17] E. Gelenbe and Y. Yin, "Deep learning with dense random neural networks," in *International Conference on Man-Machine Interactions*, pp. 3–18, Springer, Cham, 2017.
- [18] A. Briana, B. LiEsa, R. Rahmira, and A. Esterline, "Behavioral modelling intrusion detection system (BMIDS) using Internet of Things (IoT) behavior-based anomaly detection via immunity-inspired algorithms," in *2016 25th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–6, Waikoloa, HI, USA, August 2016.
- [19] E. Anthi, L. Williams, M. Slowinska, G. Theodorakopoulos, and P. Burnap, "A supervised intrusion detection system for smart home IoT devices," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9042–9053, 2019.
- [20] SANS, *Host- vs. Network-Based Intrusion Detection Systems*, SANS Penetration Testing, 2005, <https://cyber-defense.sans.org/resources/papers/gsec/host-vs-network-based-intrusion-detection-systems-102574>.
- [21] ASPENCORE, *2019 Embedded Markets Study*, 2019, https://www.embedded.com/wp-content/uploads/2019/11/EETimes_Embedded_2019_Embedded_Markets_Study.pdf.
- [22] Strace, linux syscall tracer <https://strace.io>.
- [23] Linux man page <https://linux.die.net/man/8/crond>.
- [24] Common vulnerabilities and exposures, CVE-2016-5195 <https://cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2016-5195>.
- [25] A. Marzano, D. Alexander, O. Fonseca et al., "The evolution of Bashlite and Mirai IoT botnets," in *2018 IEEE Symposium on Computers and Communications (ISCC)*, pp. 813–818, Natal, Brazil, June 2018.
- [26] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016.
- [27] A. Fabrizio, A. Luciano, and F. Angelo, "Exploiting N-gram location for intrusion detection," in *IEEE International Conference on Tools with Artificial Intelligence*, Vietri sul Mare, Italy, November 2015.

- [28] B. Candice, C. Anna, and M. M. Gonzalo, *A Comparative Analysis of XGBoost*, 2019.
- [29] J. A. Suvkens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [30] F. T. Liu, M. T. Kai, and Z. H. Zhou, "Isolation-based anomaly detection," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 6, no. 1, pp. 1–39, 2012.
- [31] K. P. Bennett and E. J. Bredensteiner, "Duality and geometry in SVM classifiers," *ICML*, vol. 2000, pp. 57–64, 2000.

Research Article

TLFW: A Three-Layer Framework in Wireless Rechargeable Sensor Network with a Mobile Base Station

Anwen Wang¹, Xianjia Meng¹, Lvju Wang¹, Xiang Ji¹, Hao Chen², Baoying Liu¹, Feng Chen¹, Yajuan Du³, and Guangcheng Yin⁴

¹School of Information Science and Technology, Northwest University, Xi'an, China

²China University of Labor Relations, Beijing, China

³Beijing Zunguan Technology Company (National Electronic Computer Quality Supervision and Inspection Center), China

⁴Inspur Software Group Co., Ltd., China

Correspondence should be addressed to Xianjia Meng; xianjiam@nwu.edu.cn

Received 23 January 2020; Revised 24 August 2020; Accepted 11 September 2020; Published 26 September 2020

Academic Editor: Huaqun Wang

Copyright © 2020 Anwen 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.

Wireless sensor networks as the base support for the Internet of things have been a large number of popularity and application. Such as intelligent agriculture, we have to use the sensor network to obtain the growing environment data of crops and others. However, the difficulty of power supply of wireless nodes has seriously hindered the application and development of Internet of things. In order to solve this problem, people use low-power sleep scheduling and other energy-saving methods on the nodes. Although these methods can prolong the working time of nodes, they will eventually become invalid because of the exhaustion of energy. The use of solar energy, wind energy, and wireless signals in the environment to obtain energy is another way to solve the energy problem of nodes. However, these methods are affected by weather, environment, and other factors, and they are unstable. Thus, the discontinuity work of the node is caused. In recent years, the development of wireless power transfer (WPT) has brought another solution to this problem. In this paper, a three-layer framework is proposed for mobile station data collection in rechargeable wireless sensor networks to keep the node running forever, named TLFW which includes the sensor layer, cluster head layer, and mobile station layer. And the framework can minimize the total energy consumption of the system. The simulation results show that the scheme can reduce the energy consumption of the entire system, compared with a Mobile Station in a Rechargeable Sensor Network (MSiRSN).

1. Introduction

Internet of things (IoTs) are applied everywhere now. Wireless sensor network as the base support for the Internet of things has been a large number of popularity and application. Such as intelligent agriculture, we have to use the sensor network to obtain the growth environmental data of crops and others. However, energy supplied is the bottleneck of wireless sensor network. In the past, the method is to extend the network life through energy saving. The latest method is to combine wireless power transmission to make the wireless sensor network immortal, but the energy consumption caused by the travel time ignored in this method cannot be ignored in the sensor network with high node density. To reduce the traveling time in a period, we propose a tiered system archi-

ture in this paper. In wireless sensor networks, finite battery capacity is a major limitation of untethered nodes. Sensor nodes will operate for a finite duration, only as long as the battery lasts. The difficulty of the power supply of wireless nodes has seriously hindered the application and development of Internet of things. In order to solve this problem, there are several solution techniques that have been proposed to maximize the lifetime of wireless sensor network, such as energy-aware routing protocols [1, 2], energy-efficient MAC protocols [3], redundant development of nodes [4], and power management strategies [5, 6]. All the above techniques can maximize the lifetime of network. But the lifetime still remains bound, and they will eventually become invalid because of the exhaustion of energy. The use of solar energy, wind energy, and wireless signals in the

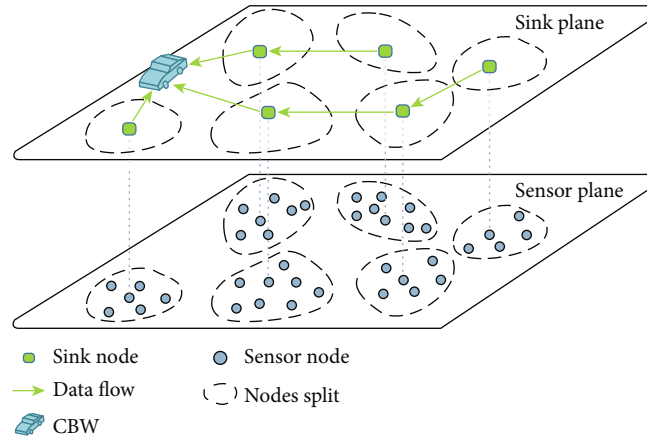


FIGURE 1: Layered system model.

environment [7, 8] to obtain energy is another way to solve the energy problem of nodes. However, these methods are affected by weather, environment, and other factors, and are unstable. Thus, the discontinuous work of the node is caused.

In recent years, the development of wireless power transfer (WPT) [9] has brought another solution to this problem. Wireless power transfer based on magnetic resonant coupling [10, 11] has been demonstrated to be a promising technology to address the problem in a wireless sensor network [12–15]. In the paper, a Mobile Station in a Rechargeable Sensor Network (MSiRSN) [12], the author shows how charging vehicle (WCV) can charge batteries of sensor nodes and how to carry the base station (MBS) to gather data. There is a home service station for the vehicle. The authors addressed the problem of colocating the MBS on the WCV in a WSN by studying an optimization problem with a focus on the traveling path problem of the WCV, the data flowing routing depending on where the WCV is in the network, stopping points, and charging schedule to minimize energy consumption of the entire system while ensuring none of the sensor nodes runs out of energy. In each charge period, WCV travels inside the network and charges every sensor node. In the above papers, the traveling time is a little proportion of total time consisting of traveling time, vacation time, and charging time. For example, the traveling time equals 1022 s, the vacation time equals 10.26 hours, and the charging time is 3.41 hours in the solution of simulation [13]. However, as the traveling time increases with the node density and the traveling time is a great part of the total time, traveling and charging every node in a period are improper.

To sum up, energy supplied is the bottleneck of wireless sensor network. In the past, the method is to extend the network life through energy saving. The latest method is to combine wireless power transmission to make the wireless sensor network immortal, but the energy consumption caused by the travel time ignored in this method cannot be ignored in the sensor network with high node density.

To reduce the traveling time in a period, we propose a tiered system architecture consisting of the sensor layer, clus-

ter head layer, and mobile station layer, as illustrated in Figure 1. The CBW (a Car as Mobile bastion with Wireless power transfer) travels all cluster heads at cluster head layer and selects the cluster in which CBW travels all sensor nodes at sensor layer in a subperiod. Several subperiods form a period in which the sensor nodes in every cluster are traveling once. Compared to traveling all nodes in [12, 13], the strategy reduces the proportion of the traveling time in total time, leading to reduction of the total energy consumption in the entire system, which includes power used by the CBW and the power consumed for wireless power transfer.

1.1. Summary and Contribution

- (i) In this paper, we design a three-layer framework for rechargeable wireless sensor network based on mobile base station, which reduces the energy consumption of mobile charging process
- (ii) A centralized clustering algorithm is proposed, which organizes sensors into m clusters, optimizes mobile charging strategy, and shortens mobile charging time
- (iii) An optimization method of joint charging plan is designed. The problem of energy supply for high-density wireless sensor network is solved

2. Related Work

The lifetime of wireless sensor networks is often limited by energy supplies. The problem of node energy supply is also a key problem in the application development of wireless sensor networks. To solve this problem, researchers have explored a wide variety of solutions.

One type of the methods is to save energy by optimizing the hardware and software [16] of the nodes. There are several solution techniques that have been proposed to allow nodes to work as long as possible in a limited amount of energy. Such as energy-aware routing protocols [1], energy-efficient MAC protocols [3], redundant development of nodes [4], and power management strategies [5]. But no

matter how energy-efficient, the battery will eventually be used up. Then, the network is invalid.

Another type of the methods is to automatically obtain energy by nodes from the natural environment, such as the wind and solar energy [17]. The energy-harvesting techniques referring to harnessing energy from the environment and converting energy to electrical energy make that a node can be powered perpetually possible, such as [7, 8]. Due to uncontrollability and unpredictability of the energy source that refers to the ambient source of energy to be harvested, the techniques cannot ensure that nodes run in every moment.

The third type of the methods is to obtain energy using ubiquitous radio signals [18]. However, this technology is still in its initial stage of research and can obtain very little energy. This is mainly caused by the far distance and the limited transmitting power of the electromagnetic wave. Recently, wireless power transfer based on magnetic resonant coupling [10] has been demonstrated to be a promising technology to address the problem in a wireless sensor network [12–14, 19, 20]. In MSiRSN, the authors showed how charging vehicle (WCV) can support wireless power transfer by bringing an energy source charge to proximity of sensor nodes and charging their batteries wirelessly. But the overall energy consumption is higher.

In this paper, we propose a tiered system architecture consisting of the sensor layer, cluster head layer, and mobile station layer to reduce the traveling time in a period. The CBW travels all cluster heads at cluster head layer and selects the cluster in which CBW travels all sensor nodes at sensor layer in a subperiod. Several subperiods form a period in which the sensor nodes in every cluster are traveling once. Compared to traveling all nodes in the above methods [12, 13], the strategy reduces the proportion of the traveling time in total time, leading to the reduction of the total energy consumption in the entire system, which includes power used by the CBW and the power consumed for wireless power transfer.

3. Overview

In order to make the nodes run forever, this paper proposes a tiered system framework for rechargeable mobile data collection wireless sensor network. The framework is divided into three layers, including sensor layer, cluster head layer, and mobile station layer. The application of this framework can reduce the traveling time in a period. In this three-layer framework, The CBW travels all cluster heads at cluster head layer and selects the cluster in which CBW travels all sensor nodes at sensor layer in a subperiod. Several subperiods form a period in which the sensor nodes in each cluster are travelled only once. Compared to other methods, this strategy reduces the proportion of the traveling time in total time, leading to reduction of the total energy consumption in the entire system, which includes power used by the CBW and the power consumed for wireless power transfer.

At the sensor layer, a centralized clustering algorithm is proposed for sensors to organize them into m clusters and the sensor nodes transmit data to the cluster head via a single

hop. In contrast to existing clustering methods which balance energy consumption, our scheme generates m cluster heads to minimize the total energy consumption. The single-hop data routing reduces energy consumption through that the sensor turns off the radio when there is no data generated by themselves to transmit.

At cluster head layer, cluster heads can cooperate with each other and the cluster head information is forwarded to CBW (a Car as Mobile bastion with Wireless power transfer) via multihop. The optimal flow routing is solved for a CBW moving trajectory to save energy.

At the mobile station layer, we study an optimization problem that joints charging schedule for cluster heads and sensors, and flow routing for cluster heads.

4. Background

Wireless power transfer based on magnetic resonant coupling [10] has been demonstrated to be a promising technology to address the problem in a wireless sensor network in [12, 13]. In MSiRSN [12], the authors showed how charging vehicle (WCV) can support wireless power transfer by bringing an energy source charge to proximity of sensor nodes and charging their batteries wirelessly, and carry the base station (MBS) to gather data. There is a home service station for the vehicle. The authors addressed the problem of colocating the MBS on the WCV in a WSN by studying an optimization problem with a focus on the traveling path problem [21] of the WCV, the data flowing routing depending on where the WCV is in the network, stopping points, and charging schedule to minimize energy consumption of the entire system while ensuring none of the sensor nodes runs out of energy. In each charge period, WCV travels inside the network and charges every sensor node.

5. Layered Network Model

In this section, we give an overview of entire framework. As depicted in Figure 1, it consists of three layers: sensor layer, cluster head layer, and mobile station layer.

We consider a set of sensor nodes N^* distributed over a two-dimensional area. Each sensor node has a battery with a capacity of E_{\max} and the initial energy of battery is a random value. E_{\min} is denoted as the minimum level of energy at a battery for it to be operational. Each sensor node i generates sensing data with a rate r_i (in b/s), $i \in N^*$. Within the sensor network, there is a mobile CBW to charge sensor nodes and gather the entire network information.

In the paper [12, 13], the authors proposed strategies to keep all nodes running forever using the wireless power transfer. In MSiRSN, the authors studied the problem of colocating the MBS on the WCV in a WSN to minimize energy consumption of the entire system. The WCV follows a periodic schedule to travel inside the network for charging every sensor node.

However, as the traveling path increases with the number of sensor nodes, the time of traveling all sensor nodes is a large proportion of a period with large sensor nodes in a

wireless network. So, traveling all sensor nodes in a period is an unwise strategy.

To address the issue, we introduce a three-layer model, consisting of the sensor layer, cluster head layer, and mobile station layer, as illustrated in Figure 1. The CBW travels all cluster heads N at cluster head layer and selects a cluster in which CBW travels all sensor nodes at sensor layer in a sub-period. Several subperiods form a period in which the sensor nodes in every cluster are traveling once. The schedule shows the cluster heads with higher energy consumption have higher charging frequency than normal sensor nodes.

6. Cluster Selection Algorithm

Since sensor nodes are energy-constrained, the network's lifetime is a major concern, especially for applications of WSNs in harsh environments. There are several solution techniques proposed, such as energy-aware routing protocols [1], energy-efficient MAC protocols [3], redundant development of nodes [4], and power management strategies [5]. To support scalability of large WSN, nodes are often grouped into disjoint and mostly nonoverlapping cluster. The most well-known hierarchical routing protocols are LEACH, HEED, TEEN, PEGASIS [22], etc. Other clustering algorithms in the literature vary in their objectives, such as load balancing [23], fault-tolerance [24], increased connectivity and reduced delay [25–27], and minimal cluster count [28]. However, the above clustering algorithms are all proposed to prolong the lifetime of the network. In this paper, a cluster selection algorithm is proposed to minimize the energy consumption of the whole network, and it can be combined with wireless power transmission technology to make the nodes run forever. Then, let the wireless network achieve immortality instead of maximizing the network life.

We propose an algorithm to minimize the total communication energy consumption $C_s = \sum_{i=1}^m \sum_{t \in S_i} C_{ti} \bullet r_t$ ($i \in N$) in sensor layer. In order to clearly describe the algorithm, denote \bar{N} as the number of elements in set N . The inputs of the algorithm are all node N^* locations, communication consumption model parameters β_1 , β_2 , and ω , and the number of cluster head m . The output of the algorithm is the m cluster in which there are several normal nodes and a cluster head to minimize the within-cluster sum of communication energy consumption. The algorithm has four steps:

Step 1. Randomly give initial cluster head set N and $\bar{N} = m$

Step 2. We assign each node to a cluster that the node's communication energy consumption to the cluster head is minimum. The strategy yields the least within-cluster sum of communication energy consumption

Step 3. We update the cluster head in a cluster through calculating the new mean to be the centroid of the sensor nodes in a cluster and setting the sensor nodes closest to the centroid as new cluster head in the cluster

Step 4. Alternate Step 2 and Step 3 until the centroids of all clusters do not change in range

7. Layer 1: Normal Sensor Nodes

7.1. Static Routing. We suppose every node's location is known in the network. By a clustering algorithm in Section 6, we can get m clusters and two type sensor nodes (normal nodes and cluster head nodes). To conserve the energy, we suppose the normal sensor nodes have no collaboration capability, only send data to the cluster head via a single hop, and do not forward packets coming from other sensor nodes.

The total data rate in cluster i (denoted as R_i) contains two parts: the first is data received from normal sensor nodes in cluster i and the second is the data generated by cluster head node i . Hence,

$$R_i = \sum_{t \in S_i} r_t + r_i, \quad (1)$$

where S_i is the set of the type 0 nodes in cluster i , r_i and r_t is the data generation rate of nodes i and t respectively. Given that cluster results, we have that R_i is constant.

7.2. Energy Consumption in a Cluster. Denote C_{ti} as the energy consumption rate for transmitting one unit of data flow from normal sensor node t to cluster head i . Then, C_{ti} (in Joule/bit) can be modeled as [29]: $C_{ti} = \beta_1 + \beta_2 D_{ti}^\omega$, where D_{ti} is the physical distance between node t and node i , β_1 and β_2 are constant terms, and ω is the path loss index and typically between $2 \leq \omega \leq 4$ [30].

$$D_{ti} = \sqrt{(x_t + x_i)^2 + (y_t + y_i)^2}, \quad (2)$$

where (x_t, y_t) and (x_i, y_i) are the coordinates of cluster head i and normal sensor node t . Given that all nodes are stationary and the cluster result is stationary, we have D_{ti} and C_{ti} that are all constants.

Denote α (in Joule/bit) as the energy consumption rate for sensing one unit of data. The power consumption of the CPU is not taken into account.

For a normal sensor node t , the total energy consumption rate c_t is as follows:

$$\alpha \bullet r_t + C_{ti} \bullet r_t = c_t, \quad (t \in S_i), \quad (3)$$

where S_i is the set of normal nodes within the cluster i . Denote C_i^* as the total energy consumption of all normal sensor nodes in a cluster i . Because the normal sensor nodes at a cluster only send data to the cluster head, there is no receiving energy consumption. Then, we have the following:

$$C_i^* = \sum_{t \in S_i} c_t = \sum_{t \in S_i} (\alpha \bullet r_t + C_{ti} \bullet r_t), \quad (i \in N). \quad (4)$$

Given a cluster solution, C_i^* is constant. C_s is denoted as the total energy consumption in the sensor layer, and then we have the following:

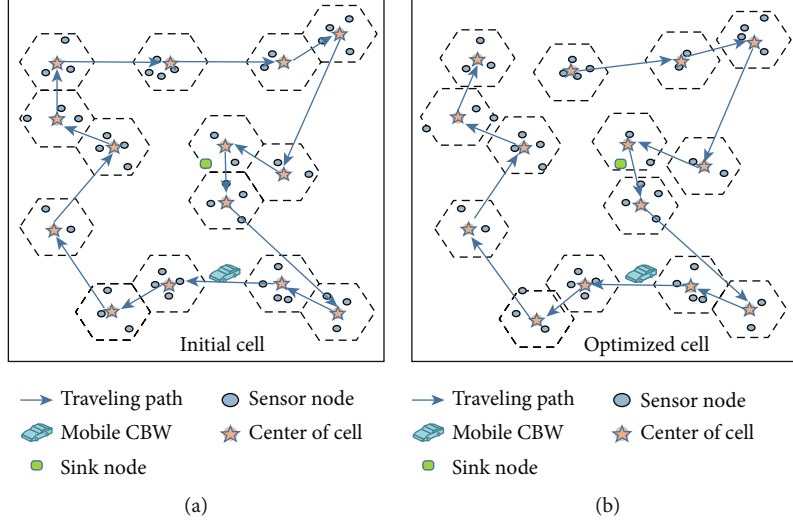


FIGURE 2: Cell charging model at a cluster.

$$C_s = \sum_{i=1}^m C_i^* = \sum_{i=1}^m \sum_{t \in S_i} (\alpha \cdot r_t + C_{ti} \cdot r_t), \quad (i \in N), \quad (5)$$

where m is the number of cluster.

7.3. Charging Model and Charging Behavior. In this section, we give a charging model and a charging behavior for normal sensor nodes in a cluster.

7.3.1. Charging Model. Based on the charging technology [31], the vehicle with a wireless power transfer can charge neighboring nodes as long as they are within its charging range. We denote $U_{tB}(p)$ as the power reception rate at normal sensor node t when the vehicle position is p . Denote the efficiency of wireless charging by $\mu(D_{tB}(p))$ when the node is in charge range. U_{\max} is denoted as the maximum output power for a node and D_δ is denoted as the charging range of wireless power transfer. We assume power reception rate is too low to make magnetic resonant coupling work properly at the node battery, when the distance between the node and the mobile CBW wireless charging model is as follows [13]:

$$U_{tB}(p) = \begin{cases} \mu(D_{tB}(p)) \cdot U_{\max}, & \text{if } D_{tB}(p) \leq D_\delta, \\ 0, & \text{if } D_{tB}(p) > D_\delta, \end{cases} \quad (6)$$

where $\mu(D_{tB}(p))$ is a decreasing function of $D_{tB}(p)$, and $0 \leq \mu(D_{tB}(p)) \leq 1$.

7.3.2. Cellular Structure and Energy Charging Behavior. We consider all normal sensor nodes in a cluster. We employ the partition strategy in [13]. The two-dimensional plane of a cluster is partitioned into hexagonal cells with side length of D , as illustrated in Figure 2. We optionally optimize the cell partition solution by the algorithm which the solution is illustrated in Figure 2. To charge normal sensor nodes in a cell, the mobile CBW only needs to visit the center of a cell. All normal sensor nodes within a hexagonal cell are within a distance of D from the cell center. Denote Q_i as the set of all

cell centers and $q^*(q \in Q_i)$ as a cell. So, the power reception rate of t in a cell with a center q is $U_{tB}(q) = \mu(D_{tB}(q)) \cdot U_{\max}$ ($q \in Q_i, t \in q^*$), where $D_{tB}(q)$ is the distance between t and the center q . Note that given the cell deployment and the t' position, the $D_{tB}(q)$ is constant. So we simply convert $D_{tB}(q)$ into D_t and $U_{tB}(q)$ into U_t .

We employ the so-called logical energy consumption rate $\gamma_t = c_t/U_t$ at a normal sensor node. Denote Γ_i^* as the logical energy consumption rate at a cluster.

$$\gamma_t = \frac{c_t}{U_t} = \frac{\alpha \cdot r_t + C_{ti} \cdot r_t}{U_t}, \quad (t \in S_i), \quad (7)$$

$$\Gamma_i^* = \sum_{t \in S_i} \gamma_t = \sum_{t \in S_i} \frac{c_t}{U_t} = \sum_{t \in S_i} \frac{\alpha \cdot r_t + C_{ti} \cdot r_t}{U_t}, \quad (i \in N). \quad (8)$$

7.4. Traveling Period in a Cluster. Denote \mathcal{P}_i as the traveling path and τ_i as the amount of time for each cycle. Then, τ_i includes three components:

- (i) The total traveling time along path \mathcal{P}_i is $\mathcal{D}_{\mathcal{P}_i}/V$, where $\mathcal{D}_{\mathcal{P}_i}$ is the distance along path \mathcal{P}_i and V is the traveling speed of the vehicle
- (ii) The total sojourn time along path \mathcal{P}_i , which is defined as the sum of all stopping times of the vehicle when it travels on \mathcal{P}_i
- (iii) The vacation time for the vehicle in a cluster i , τ_{vac_i} , which starts when the vehicle leaves the cluster i and ends when the vehicle travels the path \mathcal{P}_i for charge all normal nodes in the cluster i

Then, we have the following:

$$\tau_i = \frac{\mathcal{D}_{\mathcal{P}_i}}{V} = \sum_{p \in \mathcal{P}_i}^{\omega_i(p) > 0} \omega_i(p) + \tau_{\text{vac}_i}, \quad (p \in \mathcal{P}_i), \quad (9)$$

where $\omega_i(p)$ denotes the aggregate amount of time when the vehicle stays at point $p \in \mathcal{P}_i$, and τ_{vac_i} denotes the vehicle leaves the cluster i and is out of cluster i charging period.

Note the mobile CBW only visits the cell center. To minimize the traveling time in a cluster, the mobile CBW must move along the shortest Hamiltonian cycle that connects the cluster head and the centers of cells in which there is at least one normal sensor node. The shortest Hamiltonian cycle can be obtained by solving the well-known Traveling Salesman Problem (TSP) [32]. D_{P_i} is denoted as the solution of the TSP.

7.5. Energy Constraints for Normal Sensor Node. We offer two conditions of energy renewable and show that once they are met, the energy level at normal sensor node never falls below

$$E_{\max} - \left[\gamma_t \bullet \tau_{\text{vac}_i} + \sum_{p \in \mathcal{P}_i}^{\omega_i(p) > 0, D_{tB}(p) > D_\delta} \gamma_t \bullet \omega(p) + \gamma_t \bullet \frac{\mathcal{D}_{\mathcal{P}_i}}{V} \right] \geq E_{\min}, \quad (i \in N, t \in S_i), \quad (10)$$

$$\gamma_t \bullet \tau_{\text{vac}_i} + \sum_{p \in \mathcal{P}_i}^{\omega_i(p) > 0, D_{tB}(p) > D_\delta} \gamma_t \bullet \omega(p) + \gamma_t \bullet \frac{\mathcal{D}_{\mathcal{P}_i}}{V} \leq \sum_{p \in \mathcal{P}_i}^{\omega_i(p) > 0, D_{tB}(p) > D_\delta} U_{tB}(p) \bullet \omega(p), \quad (i \in N, t \in S_i). \quad (11)$$

We employ a cellular structure for normal sensor nodes (in Section 7.2). From Equations (10) and (11), we obtain the following:

$$E_{\max} - \left[\gamma_t \bullet \tau_{\text{vac}_i} + \sum_{q \in Q_i} \gamma_t \bullet \omega(p) + \gamma_t \bullet \frac{\mathcal{D}_{\mathcal{P}_i}}{V} \right] \geq E_{\min}, \quad (i \in N, t \in S_i), \quad (12)$$

$$\begin{aligned} & \gamma_t \bullet \tau_{\text{vac}_i} + \sum_{q \in Q_i} \gamma_t \bullet \omega(p) + \gamma_t \bullet \frac{\mathcal{D}_{\mathcal{P}_i}}{V} \\ & \leq \sum_{p \in \mathcal{P}_i}^{\omega_i(p) > 0, D_{tB}(p) > D_\delta} U_{tB}(p) \bullet \omega(p), \quad (i \in N, t \in S_i), \end{aligned} \quad (13)$$

where Q_i is the set of all cell centers in cluster i .

8. Layer 2: Cluster Head Nodes

8.1. Dynamic Routing. With a clustering algorithm in Section 6 to minimize the total energy of all sensor nodes, we can get some specific nodes and denote them as cluster heads. Different from normal nodes, the cluster head has collaboration capability.

8.2. Dynamic Flow Balance. Due to the mobility of the vehicle, data flow routing is dynamic with routing topology changing over time. Denote $f_{ij}(p)$ and $f_{iB}(p)$ as flow rates from cluster head i to cluster head j and to the base station when the vehicle is at location $p \in P$, respectively. Then, we

E_{\min} that means the normal sensor can run forever. First, we split energy consumption at normal sensor node t in cluster i into three parts:

- (i) Energy is consumed when the CBW does not select the cluster i to charge battery: $\gamma_t \bullet \tau_{\text{vac}_i}$
- (ii) Energy is consumed when the CBW makes stops at all centers of cell in which there is at least one normal sensor node: $\sum_{p \in \mathcal{P}_i}^{\omega_i(p) > 0, D_{tB}(p) > D_\delta} \gamma_t \bullet \omega(p)$
- (iii) Energy is consumed when the CBW is moving along \mathcal{P}_i that is Hamiltonian cycle that connects the cluster head and the centers of cells in which there is at least one normal sensor node, $\gamma_t \bullet (\mathcal{D}_{\mathcal{P}_i}/V)$

have the following flow balance constraint at each cluster head i .

$$\sum_{k \in N}^{k \neq i} f_{ki}(p) + R_i = \sum_{j \in N}^{j \neq i} f_{ij}(p) + f_{iB}(p), \quad (i \in N), \quad (14)$$

where N is the set of cluster heads gotten by cluster selection algorithm in Section 6, and R_i is determined by Equation (1).

8.3. Energy Consumption. Like the energy consumption model for normal sensor nodes, the communication energy consumption between two cluster nodes i and j can be modeled as follows:

$$C_{ij} = \beta_1 + \beta_2 D_{ij}^\omega, \quad (15)$$

where $D_{ij} = \sqrt{(x_i + x_j)^2 + (y_i + y_j)^2}$, and (x_i, y_i) and (x_j, y_j) are the coordinates of cluster heads i and j . $D_{iB}(p) = \sqrt{(x_i + x_B)^2 + (y_i + y_B)^2}$, and (x_j, y_j) and (x_B, y_B) are the coordinates of type 1 node i and vehicle B at $(p \in \mathcal{P})$. Given that the cluster result and all cluster heads are stationary, we have D_{ij} and C_{ij} that are all constants. However, $D_{iB}(p)$ and $C_{iB}(p)$ varied with vehicle position p . Denote γ (in Joule/bit) as the energy consumption rate for receiving one unit of data. Then, the total energy consumption rate for transmission, reception, and sense at cluster head i when the

vehicle is at $p \in \mathcal{P}$, denoted as $c_i(p)$, is as follows:

$$\begin{aligned} \alpha \cdot r_i + \rho \cdot \sum_{t \in S_i} r_t + \rho \cdot \sum_{k \in N}^{k \neq i} f_{ki}(p) + \sum_{k \in N}^{j \neq i} C_{ij} \cdot f_{ij}(p) \\ + C_{iB}(p) \cdot f_{iB}(p) = c_i(p), \quad (i \in \mathcal{P}), \end{aligned} \quad (16)$$

where $\alpha \cdot r_i$ is denoted as the sensing consumption, $\rho \cdot \sum_{t \in S_i} r_t$ is denoted as the consumption for receiving data from all normal sensor nodes at cluster i , $\rho \cdot \sum_{k \in N}^{k \neq i} f_{ki}(p)$ is denoted as the consumption for receiving data from other cluster heads, $\sum_{k \in N}^{j \neq i} C_{ij} \cdot f_{ij}(p)$ is denoted as the consumption for transmitting data to other clusters, and $C_{iB}(p) \cdot f_{iB}(p)$ is denoted as the consumption for transmitting data to the mobile CBW. Note that the cluster head consumption $c_i(p)$ dynamically changes with the position p .

8.4. Charging Model. Like the energy charging model for normal sensor nodes in Section 7.3, we use wireless power transfer [31] to charge the rechargeable battery of cluster heads. Different from normal sensor nodes charging schedule, the charging point for every cluster head is located in cluster head, taking into account of the distance between any two cluster heads is longer than the charging range of wireless power transfer D_δ , which means it is impossible to charge two cluster heads simultaneously.

$$U_{iB}(p) = \begin{cases} U_{\max}, & \text{if } p = i, \\ 0, & \text{if } p \neq i, \end{cases} \quad (17)$$

where p is the mobile CBW position, $p = i$ denotes the mobile CBW and cluster head i are at the same position, and $p \neq i$ denotes the mobile CBW and cluster head i are at two different positions. Equation (17) shows the mobile CBW just charges a cluster head while they are at the same position and do not charge battery when it is moving.

8.5. Traveling Period for Cluster Layer. Like Section 7.4, the charge time for cluster heads is $\tau = \mathcal{D}_{\mathcal{P}}/V = \sum_{p \in \mathcal{P}}^{\omega(p) > 0} \omega(p) + \tau_{\text{vac}}$, ($p \in \mathcal{P}$), where $\omega(p)$ denotes the aggregate amount of time the vehicle stays at point $p \in \mathcal{P}$ and p_{vac} denotes the location of the home service station. To minimize the traveling time of all cluster heads, the mobile CBW must move along the shortest Hamiltonian cycle that connects the server station and all cluster heads. Like traveling time at a cluster $D_{\mathcal{P}}$, $D_{\mathcal{P}}$ is denoted as the solution of the TSP.

Based on the distance between two type 1 nodes that is longer than charging range of wireless power transfer D_δ , the points where the vehicle stops for charge cluster heads are in cluster head position. Then, Equation (9) can be written as follows:

$$\tau = \frac{\mathcal{D}_{\mathcal{P}}}{V} = \sum_{p \in \mathcal{P}}^{\omega(p) > 0} \omega(p) + \tau_{\text{vac}}, \quad (p \in N_i). \quad (18)$$

8.6. Energy Consumption in a Subperiod. We offer two energy renewable conditions and show that once they are met, the energy level at clusters head will never fall below E_{\min} , which means the cluster head can run forever. First, we split energy consumption at normal sensor node t into three parts:

- (i) Energy is consumed when the CBW makes a stop at the service station: $c_i(p_{\text{vac}}) \cdot \tau_{\text{vac}}$
- (ii) Energy is consumed when the CBW makes stops at all cluster heads: $\sum_{p \in N, p \neq i} c_i(p) \cdot \omega(p)$
- (iii) Energy is consumed when the CBW is moving along \mathcal{P} that is Hamiltonian cycle that connects all cluster heads and the service station: $\int_{s \in [0, D_{\mathcal{P}}]}^{\omega(p(s))=0} (1/V) \cdot c_i(p(s)) ds$

$$E_{\max} - \left[c_i(p_{\text{vac}}) \cdot \tau_{\text{vac}} + \sum_{p \in N, p \neq i} c_i(p) \cdot \omega(p) + \int_{s \in [0, D_{\mathcal{P}}]}^{\omega(p(s))=0} \frac{1}{V} \cdot c_i(p(s)) ds \right] \geq E_{\min}, \quad (i \in N), \quad (19)$$

$$c_i(p_{\text{vac}}) \cdot \tau_{\text{vac}} + \sum_{p \in N, p \neq i} c_i(p) \cdot \omega(p) + \int_{s \in [0, D_{\mathcal{P}}]}^{\omega(p(s))=0} \frac{1}{V} \cdot c_i(p(s)) ds \leq \sum_{p \in \mathcal{P}}^{\omega(p) > 0, D_{iB}(p) \leq D_\delta} U_{iB}(p) \cdot \omega(p), \quad (i \in N). \quad (20)$$

9. Layer 3: Charging Schedule at CBW

We consider minimizing energy consumption of the entire system which includes normal sensor nodes and cluster heads. Firstly, we minimize the total transmission energy consumption of all normal nodes in a cluster through a cluster selection strategy and give an optional charge strategy including an approximative optional path and charge

time. Secondly, for cluster head layer, we formulate the problem including mobile CBW traveling path, dynamic flow routing, and charge time, and solve the problem by CPLEX solver [33].

9.1. Formulation for Normal Sensor Nodes in a Cluster. We develop a travel schedule for the mobile CBW and charging schedule among normal sensor nodes so that no normal node

never runs out of energy. For the objective function, we consider minimizing energy consumption in sensor layer. We have followed optimization problem (OPT-normal) (time constraints (9), energy consumption model ((7), (8)), and energy renewable constraints ((12), (13))):

$$\begin{aligned}
& \max \frac{\tau_{\text{vac}_i}}{\tau_i} \\
& \text{s.t. Timeconstraints} \\
& \text{Energyconsumptionmodel} \\
& \text{Energyrenewableconstraints} \\
& \tau_i, \tau_{\text{vac}_i}, \omega_i(p) \geq 0, (p \in \mathcal{P}_i).
\end{aligned} \tag{21}$$

To minimize the traveling time in a cluster, the mobile CBW must move along the shortest Hamiltonian cycle that connects the cluster head and the centers of cells in which there is at least one normal sensor node. So, \mathcal{P}_i is denoted as the solution of this TSP.

9.2. Formulation for Cluster Nodes. We develop a travel schedule for the mobile CBW, charging schedule, and data flow routing among cluster heads so that no cluster head never runs out of energy. For the objective function, we consider minimizing energy consumption in cluster head layer. We have followed optimization problem (OPT-cluster) (time constraints (18), energy consumption model (16), and energy renewable constraints ((19), (20))):

$$\begin{aligned}
& \max \frac{\tau_{\text{vac}}}{\tau} \\
& \text{s.t. Timeconstraints} \\
& \text{Energyconsumptionmodel} \\
& \text{Energyrenewableconstraints} \\
& \tau, \tau_{\text{vac}}, \omega(p) \geq 0, (p \in \mathcal{P}) \\
& f_{ij}(p), f_{iB}(p), r_i(p) \geq 0, (i, j \in N, i \neq j, p \in P).
\end{aligned} \tag{22}$$

To minimize the traveling time of all cluster heads, the mobile CBW must move along the shortest Hamiltonian cycle that connects the server station and all cluster heads. So, P is denoted as the solution of the TSP.

9.3. Joint Solution. We find solutions to D_{tps_i} , rate_i , and t_{\max_i} for a cluster i and τ_{vac} , τ , and τ_{tps} for cluster head by CPLEX [33]. Denote the t_{\max} as the $\min(t_{\max_i}, i \in N)$.

Denote h as the number of subperiod during which the mobile CBW charges all cluster heads once, in an entire period. Denote T as the total time of a period. Then, we have $h \cdot \tau = T$. Denote T_{vac} as the vacation time in the entire period. T_{vac} equals the total time of a period minus the sum of the traveling time of cluster heads and normal sensor nodes, and the charging time of all nodes in the wireless network. T_{vac} equals the total vacation time of cluster head traveling in h subperiod minus the sum of charging time and traveling

TABLE 1: Location and data rate R_i for each node in a 50-node network.

Node index	Location	R_i	Node index	Location	R_i
1	(0.547, 0.644)	0.1	26	(0.833, 0.115)	0.2
2	(0.662, 0.757)	0.7	27	(0.639, 0.658)	0.1
3	(0.037, 0.859)	0.4	28	(0.704, 0.930)	0.6
4	(0.723, 0.741)	1.0	29	(0.977, 0.306)	0.8
5	(0.529, 0.778)	0.9	30	(0.673, 0.386)	0.5
6	(0.316, 0.035)	0.4	31	(0.021, 0.745)	0.7
7	(0.190, 0.842)	0.8	32	(0.924, 0.072)	0.6
8	(0.288, 0.106)	0.8	33	(0.270, 0.829)	0.1
9	(0.040, 0.942)	0.2	34	(0.777, 0.573)	0.8
10	(0.264, 0.648)	0.4	35	(0.097, 0.512)	0.9
11	(0.446, 0.805)	0.5	36	(0.986, 0.290)	0.2
12	(0.890, 0.729)	0.5	37	(0.161, 0.636)	0.7
13	(0.370, 0.350)	0.1	38	(0.355, 0.767)	0.9
14	(0.006, 0.101)	0.7	39	(0.655, 0.574)	0.5
15	(0.393, 0.548)	0.1	40	(0.031, 0.052)	0.4
16	(0.629, 0.623)	0.1	41	(0.350, 0.150)	0.3
17	(0.084, 0.954)	0.5	42	(0.941, 0.724)	0.1
18	(0.756, 0.840)	0.2	43	(0.966, 0.430)	0.2
19	(0.966, 0.376)	0.7	44	(0.107, 0.191)	0.3
20	(0.931, 0.308)	0.6	45	(0.007, 0.337)	0.3
21	(0.944, 0.439)	0.1	46	(0.457, 0.287)	0.4
22	(0.626, 0.323)	0.4	47	(0.753, 0.383)	0.1
23	(0.537, 0.538)	0.2	48	(0.945, 0.909)	0.1
24	(0.118, 0.082)	0.3	49	(0.209, 0.758)	0.3
25	(0.929, 0.541)	0.2	50	(0.221, 0.588)	0.8

path time of normal sensor nodes at all clusters. Then, we have $h \cdot \tau - \sum_{i \in N} \Gamma_i^* \cdot T + D_{tps_i}/V = T_{\text{vac}}$.

To minimize the entire system consumption jointing sensor layer, cluster head layer, and mobile bastion station, we study the following problem (OPT-joint).

$$\begin{aligned}
& \max \frac{T_{\text{vac}}}{T} \\
& \text{s.t. } h \cdot \tau = T \\
& T \leq t_{\max} \\
& h \cdot T_{\text{vac}} - \sum_{i \in N} \Gamma_i^* \cdot T + \frac{D_{tps_i}}{V} = T_{\text{vac}} \\
& \text{rate}_i \cdot T + \leq T_{\text{vac}}, (i \in N) \\
& h > k.
\end{aligned} \tag{23}$$

Constraint $T \leq t_{\max}$ shows the entire period is no greater than the minimum value of maximum lifetime for a cluster in all clusters, which ensures every normal sensor node is not out of energy. $\text{rate}_i \cdot T + D_{tps_i}/V \leq T_{\text{vac}}$ ($i \in N$) shows the time that the mobile CBW is into a cluster and charges normal sensor nodes is no longer than the T_{vac} that ensures in a

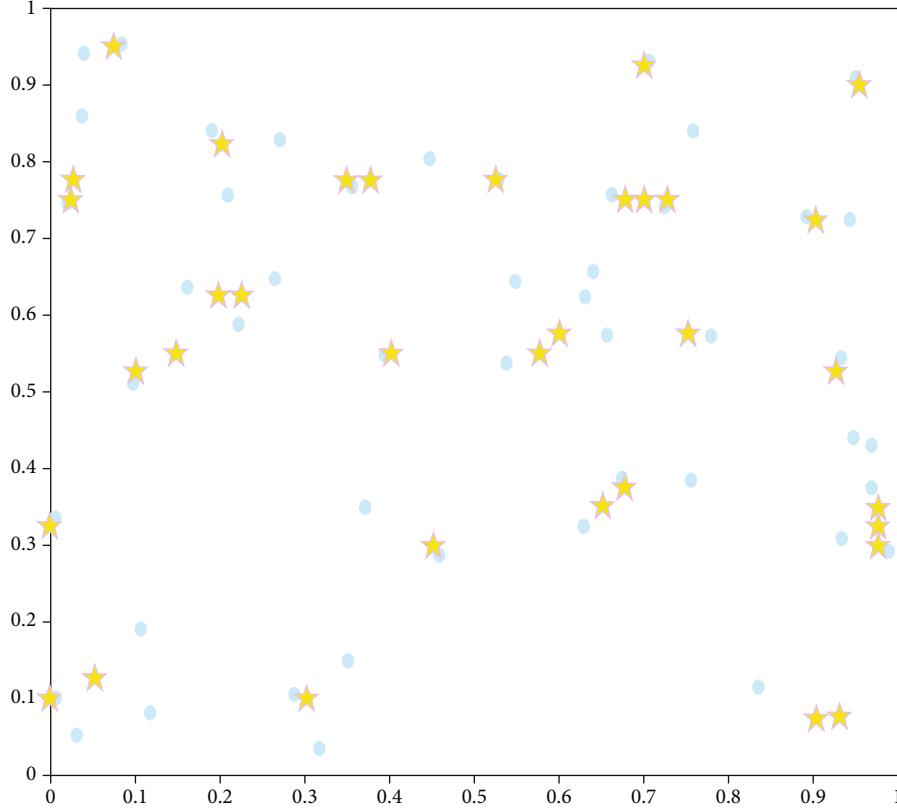


FIGURE 3: The solution in MSiRSN. The yellow star represents the stopping points, and the blue circle represents the sensor node.

subperiod; the mobile CBW can charge all normal sensor nodes. For the fractional objective function T_{vac}/T , we define $\eta_{\text{vac}} = T_{\text{vac}}/T$. Then, we can reformulate the above problem as follows:

$$\begin{aligned}
 & \max \eta_{\text{vac}} \\
 & s.t \ h \cdot \tau = T \\
 & \quad T \leq t_{\max} \\
 & h \cdot \eta_{\text{vac}} - \sum_{i \in N} \left(\Gamma_i^* \cdot T + \frac{D_{t_{ps}_i}/V}{T} \right) = \eta_{\text{vac}} \\
 & \text{rate}_i \cdot T + \leq \eta_{\text{vac}}, \quad (i \in N) \\
 & \quad h > k.
 \end{aligned} \tag{24}$$

The equation $h \cdot \eta_{\text{vac}} - \sum_{i \in N} (\Gamma_i^* \cdot T + (D_{t_{ps}_i}/V)/T) = \eta_{\text{vac}}$ shows the η_{vac} increases with T . So, we can maximize the η_{vac} via maximizing T with two constraints $T \leq t_{\max}$ and $\Gamma_i^* \cdot T + (D_{t_{ps}_i}/V)/T \leq \tau_{\text{vac}}/T$, where t_{\max} and τ_{vac}/T can be calculated in Section 9.1 and Section 9.2, respectively. So, the joint problem can be solved.

10. Performance Evaluations

In this section, we present some numerical results to demonstrate how our solution works to achieve wireless energy

transfer and evaluate the performance of the system compared to MSiRSN.

10.1. Simulation Settings. In this section, we evaluate the performance of the system and compared it with the strategy in MSiRSN. The network parameters are set like in MSiRSN. We assume sensor nodes are deployed over a 1×1 square area. The service station is at $(0.5, 0.5)$. The traveling speed of the mobile CBW is $V = 0.1$. The data rate r_t , $t \in N^*$, from each node is randomly generated within $[0.1, 1]$. Power consumption coefficients are $\beta_1 = 1, \beta_2 = 1, \rho = 1, \alpha = 0$. The path loss index is $\omega = 4$. Suppose that a sensor node uses a rechargeable battery with $E_{\max} = 10,000$ and $E_{\min} = 500$. For the charging efficiency function, $\mu(D_{tB}(p)) = -40D_{tB}(p)^2 - 4D_{tB}(p)^2 + 1$. Let $U_{\max} = 50$ and $D_{\delta} = 0.1$ for a maximum distance of effective charging. We consider a 50-node network. The normalized location of each node and its data rate are given in Table 1 in MSiRSN.

10.2. Solution with Strategy in MSiRSN. The simulation results in MSiRSN are given as follows. The traveling path in a period is $D_{P_{\text{OPT-lb}}} = 4.89$ and the traveling time is $D_{P_{\text{OPT-lb}}}/V = 48.9$. The cycle time is $\tau = 9414$, the vacation time is $\tau_{\text{vac}} = 6410$, and the objective value is 68%. The traveling path is shown in Figure 3.

10.3. Solution with Our Strategy. With our strategy, we can get a layer framework of the network, shown in Figure 4. We solve optimization problems (OPT-normal and OPT-

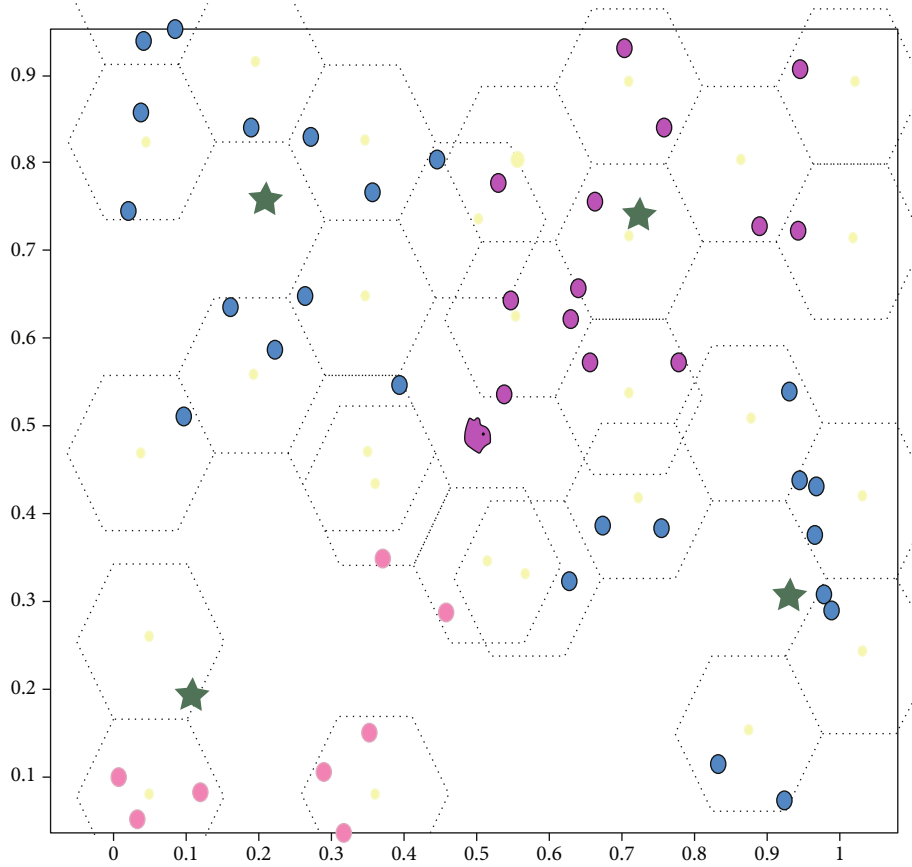


FIGURE 4: The layer framework of a network in Section 9. The green star represents the cluster head, the yellow ellipse represents the cell center that is the stopping point for charging normal nodes in the cell, and other four colored circles represent four types of normal nodes.

cluster) by CPLEX [33] and get the following solutions. Denote τ_c as the total charging time for cluster heads in a subperiod, τ_{c_i} as the total charging time for normal sensor nodes at cluster i .

$$\begin{aligned}
 D_{tps} &= 2.5425, \tau_{vac} = 209.1, \tau_c = 52.3, \tau = 286.8 \\
 D_{tps_1} &= 1.3196, \tau_{vac_1} = 10674, \tau_{c_1} = 30, \tau_1 = 10717 \\
 D_{tps_2} &= 1.239, \tau_{vac_2} = 8703, \tau_{c_2} = 56, \tau_2 = 8771.3 \\
 D_{tps_3} &= 1.5124, \tau_{vac_3} = 8416.7, \tau_{c_3} = 89, \tau_3 = 8520 \\
 D_{tps_4} &= 1.732, \tau_{vac_4} = 5946.57, \tau_{c_4} = 171, \tau_4 = 6134
 \end{aligned} \tag{25}$$

We set a period $T = \tau_4 = 6134$. We can get charging time for normal nodes $\sum_{\tau_{c_i}} 17 + 39 + 64 + 171 = 291$, charging time for cluster heads $\tau_c \cdot T/\tau = 1115$, traveling time for normal nodes $\sum D_{tps_i}/V = 84$, and traveling time for cluster heads $(D_{tps}/V) \cdot (T/\tau) = 543$ in this period. Then, we can get $\tau_{vac}/T = (6134 - (291 + 1115 + 84 + 543))/6134 = 0.71$. Our objective solution of 71% is greater than 68% in MSiRSN.

11. Conclusion

Wireless sensor network is the main part of IoTs. With the high developing time of IoTs, the difficulty of the power sup-

ply of wireless nodes has seriously hindered the application and development of IoTs. In this paper, we proposed a three-layer framework consisting of the sensor layer, cluster head layer, and mobile station layer in a rechargeable wireless sensor network. We studied the problem of charge schedule and traveling path of a mobile CBW and a cluster selection algorithm in order to minimize the energy consumption of entire system. The simulation result shows that the scheme can get a smaller energy consumption of the entire system, compared with MSiRSN.

Data Availability

The data used in this paper can be obtained directly in the sentences and tables of the paper or generated by combining them with the algorithm. The core steps and algorithms of data processing method are introduced in the paper in detail, too.

Disclosure

Anwen Wang and Xianjia Meng are co-first authors.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Key Research and Development Program of Shaanxi (Program No. 2019GY-012), the National Natural Science Foundation of China (Grant No. 61702416), the Scientific Research Program funded by Shaanxi Provincial Education Department (Program Nos. 18JK0773 and 17JK0775), and the Science Foundation of Northwest University (no. 15NW32 and no. 15NW31).

References

- [1] L. Ming, C. Jiannong, C. Guihai, and W. Xiaomin, "An energy-aware routing protocol in wireless sensor networks," *Sensors*, vol. 9, no. 1, p. 445, 2009.
- [2] C. Shao, H. Roh, T. Kim, and W. Lee, "Multisource wireless energy harvesting-based medium access control for rechargeable sensors," *IEEE Transactions on Consumer Electronics*, vol. 62, no. 2, pp. 119–127, 2016.
- [3] L. Tang, Y. Sun, O. Gurewitz, and D. B. Johnson, "PW-MAC: an energy-efficient predictive-wakeup MAC protocol for wireless sensor networks," in *2011 Proceedings IEEE INFOCOM*, pp. 1305–1313, Shanghai, China, 2011.
- [4] H. M. Ammari and S. K. Das, "Centralized and clustered k-coverage protocols for wireless sensor networks," *IEEE Transactions on Computers*, vol. 61, no. 1, pp. 118–133, 2011.
- [5] C. Alippi, G. Anastasi, M. D. Francesco, and M. Roveri, "Energy management in wireless sensor networks with energy-hungry sensors," *Instrumentation & Measurement Magazine IEEE*, vol. 12, no. 2, pp. 16–23, 2009.
- [6] Y. Shu, G. S. Kang, J. Chen, and Y. Sun, "Joint energy replenishment and operation scheduling in wireless rechargeable sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 99, pp. 125–134, 2017.
- [7] J. Jeong and D. Culler, "A practical theory of micro-solar power sensor networks," *Acm Transactions on Sensor Networks*, vol. 9, no. 1, pp. 1–36, 2012.
- [8] Y. K. Tan and S. K. Panda, "Optimized wind energy harvesting system using resistance emulator and active rectifier for wireless sensor nodes," *IEEE Transactions on Power Electronics*, vol. 26, no. 1, pp. 38–50, 2010.
- [9] J. Zhang, M. Wang, X. Shen, J. Fan, and B. Zhao, "Multi-hop energy sharing in rechargeable wireless sensor networks," *International Journal of Sensor Networks*, vol. 20, no. 4, pp. 230–242, 2016.
- [10] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 5834, pp. 83–86, 2007.
- [11] L. Fu, P. Cheng, Y. Gu, J. Chen, and T. He, "Optimal charging in wireless rechargeable sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 278–291, 2016.
- [12] L. Xie, Y. Shi, Y. T. Hou, W. Lou, and H. D. Sherali, "On traveling path and related problems for a mobile station in a rechargeable sensor network," in *Fourteenth ACM International Symposium on Mobile Ad Hoc NETWORKING and Computing*, pp. 109–118, New York, NY, USA, 2013.
- [13] Y. Shi, L. Xie, Y. T. Hou, and H. D. Sherali, "On renewable sensor networks with wireless energy transfer," in *INFOCOM, 2011 Proceedings IEEE*, pp. 1350–1358, Shanghai, China, 2012.
- [14] L. Fu, L. He, P. Cheng, Y. Gu, J. Pan, and J. Chen, "ESync: energy synchronized mobile charging in rechargeable wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7415–7431, 2016.
- [15] B. H. Liu, N. T. Nguyen, V. T. Pham, and Y. X. Lin, "Novel methods for energy charging and data collection in wireless rechargeable sensor networks," *International Journal of Communication Systems*, vol. 30, article e3050, 2015.
- [16] X. Ding, J. Han, and L. Shi, "The optimization based dynamic and cyclic working strategies for rechargeable wireless sensor networks with multiple base stations and wireless energy transfer devices," *Sensors*, vol. 15, no. 3, pp. 6270–6305, 2015.
- [17] H. Chen, X. Li, and F. Zhao, "A reinforcement learning-based sleep scheduling algorithm for desired area coverage in solar-powered wireless sensor networks," *IEEE Sensors Journal*, vol. 16, no. 8, pp. 2763–2774, 2016.
- [18] J. R. Jiang and J. H. Liao, "Efficient wireless charger deployment for wireless rechargeable sensor networks," *Energies*, vol. 9, no. 9, p. 696, 2016.
- [19] X. Li, Q. Tang, and C. Sun, "Energy efficient dispatch strategy for the dual-functional mobile sink in wireless rechargeable sensor networks," *Wireless Networks*, vol. 24, pp. 671–681, 2018.
- [20] P. Zhong, Y. T. Li, W. R. Liu, G. H. Duan, Y. W. Chen, and N. Xiong, "Joint mobile data collection and wireless energy transfer in wireless rechargeable sensor networks," *Sensors*, vol. 17, no. 8, pp. 1–23, 2017.
- [21] C. Lin, J. Zhou, C. Guo, H. Song, G. Wu, and M. S. Obaidat, "TSCA: a temporal-spatial real-time charging scheduling algorithm for on-demand architecture in wireless rechargeable sensor networks," *IEEE Transactions on Mobile Computing*, vol. 17, no. 99, pp. 211–224, 2017.
- [22] A. C. Ferreira, L. B. Oliveira, E. Habib, C. W. Hao, and A. A. Loureiro, "On the security of cluster-based communication protocols for wireless sensor networks," in *International Conference on Networking*, pp. 449–458, Springer, Berlin, Heidelberg, 2005.
- [23] S. Banerjee and S. Khuller, "A clustering scheme for hierarchical control in multi-hop wireless networks," in *INFOCOM 2001. Twentieth Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, pp. 1028–1037, Anchorage, AK, USA, 2000.
- [24] G. Gupta and M. Younis, "Fault-tolerant clustering of wireless sensor networks," in *Wireless Communications and Networking, 2003. WCNC 2003*, vol. 3, pp. 1579–1584, New Orleans, LA, USA, 2003.
- [25] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, pp. 1713–1723, San Francisco, CA, USA, 2003.
- [26] L. Liang, Y. Song, H. Zhang, H. Ma, and A. V. Vasilakos, "Phy-sarum optimization: a biology-inspired algorithm for the Steiner tree problem in networks," *IEEE Transactions on Computers*, vol. 64, no. 3, pp. 819–832, 2015.
- [27] Y. Song, L. Liang, H. Ma, and A. V. Vasilakos, "A biology-based algorithm to minimal exposure problem of wireless sensor networks," *IEEE Transactions Network and Service Management*, vol. 11, no. 3, pp. 417–430, 2014.
- [28] E. I. Oyman and C. Ersoy, "Multiple sink network design problem in large scale wireless sensor networks," in *IEEE*

- International Conference on Communications*, vol. 6, pp. 3663–3667, Paris, France, 2004.
- [29] W. B. Heinzelman, *Application-Specific Protocol Architectures for Wireless Networks*, Massachusetts Institute of Technology, 2000.
- [30] T. Rappaport, *Wireless Communications: Principles and Practice*, Publishing House of Electronics Industry, 2013.
- [31] A. Kurs, R. Moffatt, and M. Soljagic, “Simultaneous mid-range power transfer to multiple devices,” *Applied Physics Letters*, vol. 96, no. 4, p. 34, 2010.
- [32] D. L. Applegate, R. E. Bixby, V. Chvatal, and W. J. Cook, *The Traveling Salesman Problem: A Computational Study*, Princeton University Press, 2006.
- [33] IBM, “IBM ILOG CPLEX optimizer,” <http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/>.

Research Article

BEHT: Blockchain-Based Efficient Highway Toll Paradigm for Opportunistic Autonomous Vehicle Platoon

Zuobin Ying^{1,2}, Longyang Yi², and Maode Ma¹

¹School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore 639798

²School of Computer Science & Technology, Anhui University, China 230601

Correspondence should be addressed to Zuobin Ying; james.ying@ntu.edu.sg

Received 25 March 2020; Revised 17 April 2020; Accepted 26 August 2020; Published 24 September 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Zuobin Ying 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.

Autonomous vehicle platoon is a promising paradigm towards traffic congestion problems in the intelligent transportation system. However, under certain circumstances, the advantage of the platoon cannot be fully developed. In this paper, we focus on the highway Electronic Toll Collection (ETC) charging problem. We try to let the opportunistic platoon pass the ETC as a whole. There are three main issues in this scenario. Firstly, the opportunistic platoon is temporarily composed; vehicles do not trust each other. Secondly, single vehicle may try to escape from the ETC charging by following the platoon. Finally, platoon members may collude with each other and try to underreport the number of vehicles in the platoon so as to evade payment. To solve these challenges, we propose a blockchain-based efficient highway toll paradigm for the opportunistic platoon. The driving history, credential information of every registered vehicle, is recorded and verified from the blockchain. A roadside unit (RSU) is adopted to distinguish the single vehicle from the platoon and in charge of lane allocation. Additionally, an aggregate signature is introduced to accelerate the authentication procedure in the RSU. We analyse the potential security threats in this scenario. The experimental result indicates that our scheme is efficient and practical.

1. Introduction

Vehicle platooning is one of the innovations in the automotive industry that is aimed at improving the safety, efficiency, mileage, and time of travel of vehicles while relieving traffic congestion, decreasing pollution, and reducing stress for passengers. Reliable vehicle platooning relies on the cooperation of multiple advanced technologies, including low-latency communication (e.g., 5G), proven autopilot system (e.g., level 5 full automation), multidriving model seamless switching (e.g., cooperative adaptive cruise control and self-govern autonomous driving), and, last but not least, flexible platoon management system [1]. Since autonomous vehicles are completely dominated by artificial intelligence, the credibility of the data will determine the safety of the entire platoon. However, data always suffer from various types of attacks, such as spoofing, data tampering, and compromised data integrity. Moreover, there also exist some network attacks, for example, Sybil attack and Distributed Deny of Service (DDoS). Therefore, how to guarantee the data as well as the

network security in the autonomous vehicle platoon (AVP) becomes a considerable issue. In this paper, a blockchain-based approach is presented to deal with an interesting application issue in the *opportunistic platoon* scenario. We describe the issue and the relevant security and management problems to be solved first and then state our contributions.

The *opportunistic platoon* belongs to the dynamic platoon type [2]. There are two kinds of the dynamic platoon, namely, the *real-time platoon* and the *opportunistic platoon*. The difference between these two is that in the *real-time platoon* individual vehicles send a request to join a preexisting platoon, while there is no preexisting platoon in the *opportunistic platoon* scenario. Individual vehicles have to discover the vehicles with similar features (e.g., destination, vehicle type, and route) first and then try to formulate a platoon. Functionally speaking, they all accomplish the objective of platooning. However, from a security aspect, there was a huge difference between the two. Firstly, the *real-time platoon* is often launched by a company such as a supermarket or logistics company. The original vehicles can be regarded

TABLE 1: Security feature comparison.

	PL characteristic	PL variation	Original PM characteristic	New PM characteristic
<i>Real-time platoon</i>	Predesignated/fully trusted	Unchanged	Fully trusted	Semitrusted
<i>Opportunistic platoon</i>	Snap election/semitrusted	Changeable	—	Semitrusted

as fully trusted. Yet in the *opportunistic platoon*, vehicles do not even recognize each other and suffer from lack of trust. Secondly, compared with the *real-time platoon*, the *opportunistic platoon* is more changeable since vehicles may have different destinations, and the platoon leader has to manipulate more authentication tasks to realize vehicle joining or leaving. Finally, the platoon leader (PL) has to take on more work than the platoon member (PM). Moreover, according to aerodynamics, the platoon leader will sustain more air resistance. Thus, in the *opportunistic platoon*, no vehicle wants to be the leader. The security feature comparison is given in Table 1.

In this paper, we focus on a practical scenario in the highway. Nowadays, the highway tollgate still needs to decelerate before passing through the Electronic Toll Collection (ETC). After the ETC detects the vehicle, it raises the fence and releases the vehicle. The follow-up vehicle needs to be kept at a certain distance from the preceding vehicle in order to allow the ETC system to detect it. Thus, although the vehicle could enjoy the benefit from platooning, they would inevitably be separated and decelerate before passing the ETC. After that, vehicles have to reform a platoon. Obviously, this cumbersome operation will reduce people's enthusiasm for the platoon. Therefore, we considered the following question. *How to let an opportunistic platoon passing through the ETC as a whole?* For example, we can designate the PL to pay the ETC charge for the entire platoon. Then, other PMs could pass through the ETC without waiting or slowing down the speed. However, this is a complicated problem which includes not only theoretical problems but also the practical ones. We list the questions below:

- (1) Since in the *opportunistic platoon* vehicles have no relationship with each other, if PMs try to escape from the part that they should pay, how to guarantee the rights of the PL?
- (2) There exist both single vehicle and platoon in the highway. If a single vehicle tries to escape from paying the ETC charge by following the platoon, what should be done to prevent this situation?
- (3) Generally speaking, the PL would not state that there are more vehicles than it actually exists in a platoon. Yet the PL may state less than the actual number of vehicles to cheat the ETC system. Since in our proposed scheme vehicles could keep a very high speed (e.g., 100 km/h) when passing the ETC, it is hard for the ETC to detect the accurate number of vehicles

Motivated by solving the abovementioned problems, we proposed the BEHT system: a blockchain-based efficient highway toll paradigm for opportunistic autonomous vehi-

cle platoon. Our main contributions could be summarized as follows:

- (1) The mutual mistrust issues in the *opportunistic platoon* are solved through blockchain. When the platoon passes through the ETC, the PL will pay for the entire platoon. Afterward, the PM could not repudiate to pay the part it should take. The smart contract will help to supervise the payment transfer from PM to PL
- (2) A lane allocation mechanism is proposed to distinguish single vehicle with platoon so as to prevent single vehicle from escaping ETC charging by following the platoon
- (3) We implement the aggregate signature into the *opportunistic platoon*. The PL needs to announce the number of vehicles in the platoon to the roadside unit (RSU) first, and all the members should sign on the announcement. Although the PL could underreport the number of vehicles, it is still possible to trace the actual number of vehicles in the platoon through the blockchain record

The rest of this paper is organized as follows. In Section 2, we present the state-of-the-art platooning management methods in general. Section 3 gives the relevant preliminaries. In Section 4, the definitions of the system model and security model are given; then, we give the details of the proposed blockchain-based efficient highway toll paradigm (BEHT) in Section 5. The analysis of BEHT in terms of security and performance can be found in Section 6. Finally, the conclusion is given in Section 7.

2. Related Works

The opportunistic platoon belongs to dynamic platoon management category. There are two different kinds of dynamic platoon. One is the *real-time platoon*, and the other one is the *opportunistic platoon* [3]. The opportunistic platoon refers to the vehicles that are close in a certain distance and have the similar interest or features with each other that form a temporary platoon without prior planning. The opportunistic platoon formation strategy is complicated. It requires not only the cooperation of vehicles in terms of maneuver but also the robust formation protocols. Besselink et al. give an overall review on cyber-physical control of road freight transportation [4]. They also discuss the possibility and precondition of forming an opportunistic platoon. Sokolov et al. considered the platoon formation maximization by coordinating the centralized routing and departure time. They also present a concrete simulation result as well as an

optimization model [5]. Since the deployment of platooning technology is not widespread, the potential benefits or opportunistic platoon has not been discovered. Therefore, some previous planning is required. Zeng et al. proposed a joint communication and control mechanism for wireless autonomous vehicular platoon systems; both the communication delay and the stability of control system are considered [6]. Through utilizing the Markovian jumping system theory, Wen and Guo proposed a sampled-data control system for connected vehicles subject to switching topologies, communication delays, and external disturbances [7]. Alam et al. discussed the significance of heavy-duty vehicle platooning that would help enhance safety and efficiency in global trade. They also evaluated the fuel saving, controller performance, and affectivity of changeable weather conditions. Moreover, the future of freight transportation system is given, in which the author believes that cooperation and platooning play an important role [8]. Boysen et al. investigated the platooning of trucks along an identical path since they found that the efficiency of platooning cannot only be guaranteed by the platooning technique but also be impacted by the platoon formation process [9]. Gong et al. developed a novel car-following control scheme for a platoon of connected and autonomous vehicles on a straight highway; they also constructed dual-based distributed algorithms to compute optimal solutions with proven convergence [10]. After that, Gong et al. proposed a series of research on how to optimize the AVP with human-driven vehicles in real world [11–13].

In the dynamic platoon management scenario, a newcomer may not willingly follow the protocols. For example, it may take advantage of the platoon to decrease the fuel consumption, but refuse to pay the platooning service charge, or it may propagate some phishing information to harm the platoon security. To deal with these security problems, some blockchain-based platoon management schemes have also been put forward. Wagner and McMillin proposed a physical action verification scheme with blockchain [14]. They mainly focus on integrity verification when the roadside unit is absent. When malicious vehicles try to join or leave the platoon, the protocol proceeds only when the vehicles can be sensed in a certain range. Ledbetter et al. proposed a practical protocol for leadership incentives for a heterogeneous and dynamic platoon [15]. Through incentive mechanism, vehicles are encouraged to participate in the platoon leader election. Calvo and Mathar proposed a blockchain-based secure communication scheme for connected vehicles; they utilize the ring signature to verify the identity of the vehicles that joined, and then, the information can be shared among authenticated vehicles through a multiparty smart contract. But they only provided the theoretical analysis but failed to give the experiment evaluation. Moreover, they introduced the microtransaction concept to deal with the low efficiency of consensus in the bitcoin network [16]. Zhang et al. presented an onionchain-based VANET framework to integrate the traceability of intermediate variables generated during the transactions [17]. For the purpose of encouraging vehicles to participate in the building of an effective vehicular announcement network, Li et al.

proposed a privacy-preserving blockchain-based incentive announcement network for communication of smart vehicles. They designed the consensus phases based on the Byzantine fault tolerance algorithm to meet the needs of reaching an agreement in a short period of time [18]. Kang et al. proposed an optimized consensus management using reputation and contract theory to tackle the challenge of voting collusion. They used delegated proof of stake to realize consensus [19]. Cheng et al. integrated attribute-based encryption with blockchain to balance the tradeoff between the availability and the privacy preservation on the Internet of Vehicles (IoVs) [20].

3. Preliminaries

3.1. Ethereum. Ethereum is an open-source, distributed computing platform based on a public blockchain with smart contracts' scripting capabilities [21]. With the use of smart contracts, Ethereum extends the range of application, making blockchains from purely distributed repositories to open, compilable blockchain development projects. Ethereum owns a powerful Turing complete development language, and it supports a modified version of the Nakamoto consensus through transaction-based state transitions. Miners use a consistent algorithm to mine and verify transactions for generating a new block. The Ethereum protocol moves far beyond currency. The currency named ether provides a liquidity layer to allow for efficient exchange of digital assets between public accounts and a mechanism for paying transaction fees.

3.2. Merkle Tree. A Merkle tree is a binary tree, in which every leaf node is labeled with the hash (e.g., SHA-256) of a data block, and every nonleaf node is labeled with the cryptographic hash by concatenating its child nodes as shown in Figure 1. The layer-by-layer operations from bottom to top, in turn, generate a unique node Merkle root. It is used to describe the integrity of all data block information stored.

Once the leaf node is modified, it will cause the change of the hash value on its parent node, which in turn affects the change of the Merkle root. According to its characteristic, in the blockchain, multiple transactions are used as data blocks of leaf nodes to build a Merkle tree. Any change in the transaction will cause a change in the Merkle root, and the integrity of all transactions can be verified by the Merkle root [22].

3.3. Elliptic Curve Digital Signature Algorithm. The Elliptic Curve Digital Signature Algorithm (ECDSA) offers a variant of the Digital Signature Algorithm (DSA) which uses elliptic curve cryptography. ECDSA has two processes for digital signature and signature verification [23]. The elliptic curve parameter is $T = (p, a, b, G, n)$, and the elliptic curve is defined as $y^2 = (x^3 + ax + b) \bmod p$, where p is a large prime number, F_p is a finite field, a and b are integers, G is the base point on $E(F_p)$, n is a prime number that is the order of the base point G , the private key of the PL is d , the public key $Q = G^d$, k is the chosen random integer, e is the value of the

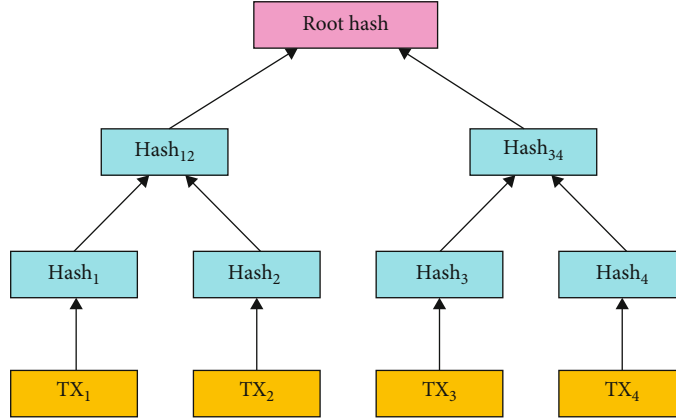


FIGURE 1: Merkle tree.

hash operation of the message m , and r are the remainders of x to n in the point (x, y) on the elliptic curve.

- (1) **ECDSA signature generation.** A signs the message m . The steps are as follows:

$$\begin{aligned}
 k &= \text{RandomInteger}[1, n - 1] \\
 G^k &= (x, y) \\
 r &= x \bmod n \\
 e &= \text{Hash}(m) \\
 s &= k^{-1}(e + dr) \bmod n \\
 \text{signature} &= (r, s)
 \end{aligned} \tag{1}$$

- (2) **ECDSA signature verification.** After B receives the signature data (r, s) of A, to verify the signature of A on message m , the following steps are required:

Verify that r and s are integers in the interval $[1, n - 1]$

$$\begin{aligned}
 e &= \text{Hash}(m) \\
 w &= s^{-1} \bmod n \\
 u_1 &= ew \bmod n \text{ and } u_2 = rw \bmod n \\
 X &= G^{u_1} Q^{u_2}
 \end{aligned}$$

If X is the point at infinity, then reject the signature.

Otherwise, convert the x coordinate of X to an integer \bar{x} .

$$v = \bar{x} \bmod n$$

If $v = r$, accept the signature, otherwise abort.

(2)

3.4. Aggregate Signature. The aggregate signature can create a signature on arbitrary distinct messages $M_i \in \{0, 1\}^*$ [24]. In this scheme, G_1 and G_2 are two multiplicative cyclic groups of prime order p . G_1 and G_2 , their respective generators g_1 and g_2 , the computable isomorphism Ψ from G_2 to G_1 , and the bilinear map $e : G_1 \times G_2 \rightarrow G_T$, with target group G_T , are system parameters. The scheme includes five algorithms: *KeyGen*, *Sign*, *Verify*, *Aggregate*, and *Aggregate Verify*.

- (1) *Key Generation.* For a particular user, pick random $x \leftarrow^R Z_p$, and compute $v \leftarrow g_2^x$. The user's public key is $v \in G_2$. The user's secret key is $x \in Z_p$.
- (2) *Signing.* For a particular user, given the secret key x and a message $M \in \{0, 1\}^*$, compute $h \leftarrow H(M)$, where $h \in G_1$ and $\sigma \leftarrow h^x$. The signature is $\sigma \in G_1$.
- (3) *Verification.* Given the user's public key v , a message M , and a signature σ , compute $h \leftarrow H(M)$; accept if $e(\sigma, g_2) = e(h, v)$ holds.
- (4) *Aggregation.* For the aggregating subset of users $U \in \text{Users}$, assign to each user an index i , ranging from 1 to $k = |U|$. Each user $u_i \in U$ provides a signature $\sigma_i \in G_1$ on a message $M_i \in \{0, 1\}^*$ of his choice. The messages M_i must all be distinct. Compute $\sigma \leftarrow \prod_{i=1}^k \sigma_i$. The aggregate signature is $\sigma \in G_1$.
- (5) *Aggregate Verification.* We are given an aggregate signature $\sigma \in G_1$ for an aggregating subset of users U , indexed as before, and are given the original messages $M \in \{0, 1\}^*$ and public keys $v_i \in G_2$ for all users $u_i \in U$. To verify the aggregate signature σ

- (i) ensure that the messages M_i are all distinct and reject otherwise
- (ii) compute $h_i \leftarrow H(M_i)$ for $1 \leq i \leq k = |U|$ and accept if $e(\sigma, g_2) = \prod_{i=1}^k e(h_i, v_i)$ holds

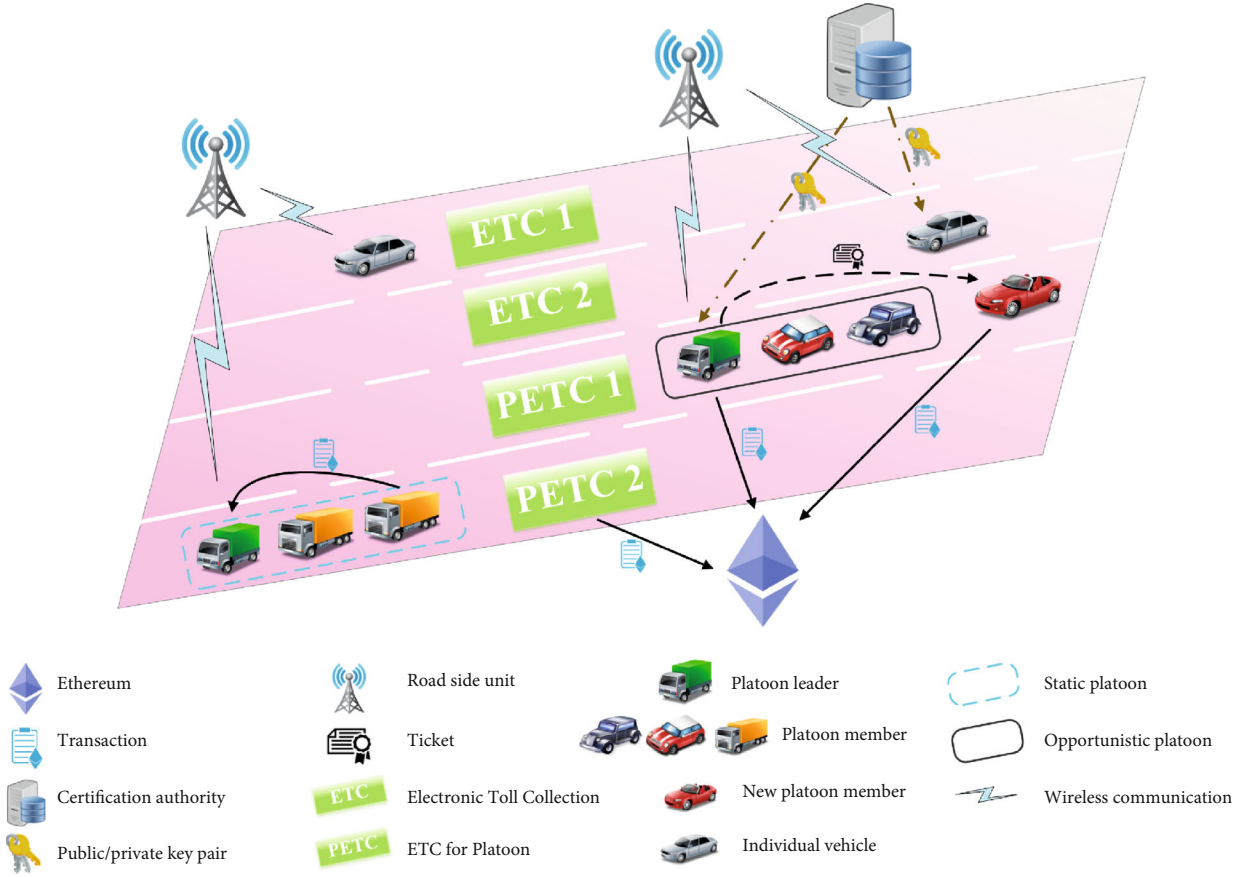


FIGURE 2: System model.

Using the properties of the bilinear map, the left-hand side of the verification equation expands:

$$\begin{aligned}
 e(\sigma, g_2) &= e\left(\prod_{i=1}^k h_i^{x_i}, g_2\right) = \prod_{i=1}^k e(h_i, g_2)^{x_i} \\
 &= \prod_{i=1}^k e(h_i, g_2^{x_i}) = \prod_{i=1}^k e(h_i, v_i),
 \end{aligned}
 \tag{3}$$

which is the right-hand side, as required.

4. System Model

Our system consists of six entities: platoon leader (PL), platoon member (PM), roadside unit (RSU), certificate authority (CA), Ethereum, and Electronic Toll Collection (ETC) System, as shown in Figure 2.

(1) *Platoon Leader*. The platoon leader is the core of a platoon. It can create a new platoon, release commands in the platoon, communicate with RSU, and distribute tickets for the vehicles that want to join the platoon. However, it may release wrong commands on purpose.

(2) *Platoon Member*. New platoon member (NPM) can join into a platoon as a PM after authentication. The PM receives commands from the PL. The PM passes through the ETC behind the PL. The PM is assumed to be dishonest. It tries to escape from the payment in the platoon. In a static platoon, members are fixed. The opportunistic platoon allows any vehicle to join.

(3) *Roadside Unit*. The roadside unit releases traffic information and verifies the deduction transaction status. It can be recognized just as a trusted facility.

(4) *Certificate Authority*. Certificate authority is responsible for releasing public/private key pairs for each vehicle. All the vehicles should register at the CA before entering the blockchain. The CA does not have to be online during the entire platoon journey. The authentication work is entrusted to the Ethereum. The CA is assumed to be fully trusted.

(5) *Ethereum*. Ethereum is responsible for the vehicle authentication and ETC payment through the use of the smart contract. Its internal data structure Merkle tree can effectively verify transactions. The Ethereum is assumed to be fully trusted.

- (6) *Electronic Toll Collection*. Electronic Toll Collection assigns the optimal ETC channel to platoons and deducts the registered owner's account without requiring them to park. It is a credible public facility like RSU. In particular, ETC for Platoon (PETC) is a channel specifically open for platoons.

5. Proposed BEHT Scheme

5.1. Design Goal. The goal of our proposal is to speed up the platoon's payment at the Electronic Toll Collection. In the real environment, the process of deducting the vehicle through the ETC, in turn, consumes a significant amount of time on the road. Our proposal puts several vehicles into a platoon with mutual distrust, passing through the ETC at just one time. In this way, the working time of the ETC is reduced exponentially, which dramatically improves its working efficiency.

To form a platoon not based on trust, we take advantage of the fact that the blockchain does not require trust, and the smart contract of the blockchain handles the process of building a platoon. Before the platoon travels to the ETC, the RSU can interact with the PL under our design protocol. The PL sends the result of the interaction to the PMs as instructions, which requires the function of intraplatoon communication in the platoon virtual environment. To prevent false messages issued by the PL, all communication contents will be permanently recorded on the blockchain. Then, the PL leads the PMs through the ETC in accordance under the requirements of RSU. The ETC charges the PL, and the required cost includes the sum of all vehicles in the entire platoon, which means that the PL pays for all the PMs at this time. At the end of the trip, the PM pays the service fee and ETC fee according to the transaction record to the PL through the blockchain.

In order to prevent PMs from refusing to admit that the PL paid for them, the signatures of all PMs will be aggregated to determine which vehicles have passed through the ETC. In addition, these payment transactions will also be recorded on the blockchain for later use as evidence.

5.2. Details of Scheme. In order to implement the scheme, we designed the platoon as a virtual area where private communication is possible. Moreover, we also proposed a protocol between RSU, platoon, and ETC to complete the payment through multiple interactions. Our scheme contains the following five modules.

- (1) *PL Registration*. The PL applies for registering a platoon.
- (2) *Ticket Generation*. The PL generates a dedicated ticket for NPM. The NPM uses the ticket to register into the platoon.
- (3) *Communication*. The platoon members can communicate with each other while the PL can communicate with nearby RSUs through DSRC. All communication data will be backed up on the blockchain for later verification check.

- (4) *Payment Interaction Protocol*. When the platoon is approaching the ETC, the RSU and the platoon perform multiple interaction confirmations according to the protocol we designed, and finally, the ETC completes the payment operation.

- (5) *Credibility Mechanism*. This reputation mechanism is directly related to the amount of punishment in order to effectively reduce the probability of the platoon violating the protocol.

In the following paragraphs, we elaborate on these five modules in detail.

5.2.1. PL Registration. In our blockchain-based scheme, each vehicle needs to register in the blockchain to obtain account address and the public key and private keys from CA in advance. After successfully registering the blockchain account, the vehicle can initiate transactions and invoke smart contracts in the blockchain network.

Each device on the vehicle for ETC payment owns a unique identification code, which allows the vehicle to verify identity. The PL chooses a *platoonID* and provides its own ETC identification code *IC* to register in the smart contract that we deployed on the blockchain. If the smart contract detects that the *platoonID* and *IC* have never been used, the registration will succeed. The PL had applied for a virtual platoon on the smart contract.

5.2.2. Ticket Generation. The PL interacts with the vehicle who wants to join the platoon as a PM through the Dedicated Short-Range Communications (DSRC). After confirmation, it decides whether to allow the vehicle to join the platoon and generate a ticket to it. Vehicle should register in the blockchain network to obtain a blockchain account and public/private keys. The PM provides its own ETC identification code *IC*, *platoonID* of target platoon, and the blockchain account address *addr*. The corresponding PL of the platoon integrates the data and then digitally signs it by ECDSA digital signature algorithm *E* using the private key *pk* and returns the generated signature $sign = E_{pk}(IC||platoonID||addr)$ to the PM as a ticket. With the ticket, the PM is eligible to join this platoon.

The PM sends the ETC identification code *IC*, *platoonID*, blockchain account address *addr*, and ticket to the smart contract on the blockchain to register into the platoon. If the smart contract detects that the *IC* has not been registered, the *platoonID* exists, and the ticket has never been used, smart contract uses the *platoonID* to query the public key *PK* of the PL to verify the digital signature of the ticket by ECDSA digital signature verification algorithm $V_{PK}(IC||platoonID||addr, ticket)$. If the verification succeeds, the PM is permitted to enter the platoon. All registration rules are shown in Algorithm 1.

After the PMs join into the platoon, the platoon is formed. In this scheme, the platoon includes static platoon and opportunistic platoon. All vehicles in the static platoon are unchanged and do not receive other vehicles to apply for entering the platoon. At this time, the PL does not distribute the exclusive ticket to the vehicle. The opportunistic

```

if IC.ExistInSmartContract() $\wedge$ (8)
  addr.ExistInSmartContract() then
    return Error(9)
  end.
if Vehicle.type = PL then
  if platoonID.ExistInSmartContract() then
    return Error(10)
  end
end
else
  if Vehilce.type = PM then
    if !platoonID.ExistInSmartContract() $\wedge$ 
      VerifyTicket(ticket) = failed $\wedge$ ticket.Used() then
      return Error
    end
  end
end
  RegisteIntoContract()
  
```

ALGORITHM 1: The Smart Contract Registration Rules

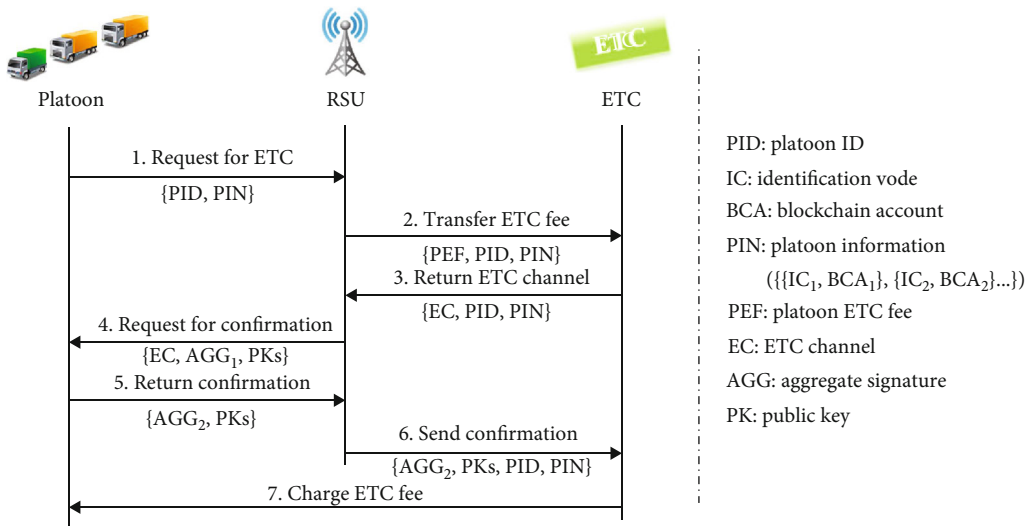


FIGURE 3: Interaction payment protocol.

platoon is a temporary set of multiple vehicles on the road, with substantial variability. Vehicles who want to join the platoon apply for the PL, and the PL returns ticket to each vehicle to join the platoon.

5.2.3. Communication. In our previous research results [25], the efficiency of uploading data to the blockchain is too slow, so we take a new approach that platoon members use DSRC to communicate with low latency directly and slowly back up the communication data on the blockchain permanently for later verification as a deposit. To achieve resource sharing, different platoons can also send and receive data through PLs using DSRC. Moreover, the PL can also communicate with nearby RSUs which send messages back, which enables multiparty data forwarding sharing.

5.2.4. Interaction Payment Protocol. In our scheme, when the platoon travels to ETC, in order to reduce the vehicle decel-

eration time significantly, the PL leads all PMs to pass ETC quickly and complete the deduction process at the same time. In such a short period, ETC cannot perform autosensing recognition for each vehicle. So the solution we adopted is that ETC only needed to identify the PL and deduct the costs of all platoon members from the PL. That is, the PL pays the ETC fee for PMs. The RSU, ETC, and platoon interact to complete the deduction protocol. Figure 3 demonstrates these interaction payment protocols. The specific processes are as follows:

- (1) When each vehicle gets onto the highway and passes through the ETC, ETC identifies the vehicle’s information and records the original station on the blockchain. On the highway, multiple vehicles build their platoons. When the platoon reaches about 1 kilometer before the ETC, the PL of the platoon sends a

request to nearby RSUs, including detailed vehicle information of the platoon

- (2) Having obtained the platoon's information, the RSU queries the blockchain for the original station of each vehicle and calculates the cost that each vehicle should pay to the ETC. Combine the results into a set of data for each vehicle, such as

{plate number:A.0001, IC: 001122, original station: New York station, terminal station: Washington station, ETC fee: 10 dollars}

The RSU sends the total fee and information of all platoon members to the ETC system

- (3) The ETC system would store the fee amount and return the assigned optimal ETC channel for the platoon to RSU
- (4) The RSU adds the ETC channel into the data set and digitally signs each data with its private key for aggregating all the digital signatures together to generate an aggregate signature. The RSU sends the aggregate signature, the original data set for verification, and its public key to the PL as requesting for confirmation, which in turn forwards it to all PMs
- (5) The PM verifies the aggregate signature and confirms the amount of the deduction. If the PM confirms that the amount of its deduction is correct, it will digitally sign the set of data with its private key and then send it to the PL. After the PL collects the digital signatures of all the PMs, it aggregates them to generate an aggregate signature. Finally, the PL sends the aggregate signature to the RSU as confirmation. The RSU sets a waiting time. If no confirmation response is received after the timeout, the RSU will retransmit, which is called the timeout retransmission mechanism
- (6) After receiving the aggregate signature, the RSU queries the public key of each vehicle on the blockchain and combines the original data to verify the validity of the aggregate signature. If the verification passes, the confirmation step for the payment is completed. All confirmation records would be sent to ETC
- (7) ETC would store all records in the blockchain and monitor the platoon for charging

The PL must lead the platoon to the designated ETC channel according to the instructions from RSU. The PL is driving in front of the platoon. The ETC automatically recognizes the ETC device of the PL then charges the PL according to the total cost fee sent from RSU before and upload the transaction record to the blockchain. In a short time, the platoon passed this ETC.

Based on the speed of the platoon, RSU estimates that the platoon has passed ETC to complete the deduction in a few minutes. At this time, the RSU queries the blockchain for a

payment record and verifies the correctness of it. If the valid deduction information did not exist, the RSU communicates with the PL again to query the current platoon status and recomplete the payment protocol. At the end of the trip, the PM pays the service fee and ETC fee according to the transaction record to the PL through the blockchain. This incentive mechanism will promote more vehicles to undertake the tasks of the PL.

5.2.5. Credibility Mechanism. To maintain the orderly operations of the scheme, we will introduce the concept of vehicle credibility value (CV) and impose late penalties on vehicles that violate the protocol. The credibility value directly affects the penalty amount.

If the platoon does not interact with the RSU according to the protocol or fails to pass the designated channel based on the RSU's instructions, this will reduce the credibility value of each platoon member and impose a penalty charge according to the penalty standard. If the platoon does not provide an aggregate signature to ETC or RSU, the ETC cannot confirm each member of the platoon. That is to say, ETC only deducts the required fee of the PL, and PMs evade the deduction operation. According to the current highway penalty mechanism, the PM for this situation is charged on the farthest distance of the current highway, and the amount to be paid is recorded for later punishment. In our proposal, the credibility value of each platoon member will be reduced, and the penalty amount for breaching the protocol will be deducted. Simultaneously, if the platoon completes the protocol, the credibility value of the platoon members will also increase. The relationship between the credibility value and the penalty amount is shown in Figure 4. We set two thresholds for the penalty amount and the rapid penalty amount. The maximum credibility value is 100, and the minimum is 0. The specific segmentation function is as follows:

- (1) $CV > 90$. The vehicle follows the protocol well and does not require additional deductions. It is the last piece with the green line in Figure 4.
- (2) $50 \leq CV \leq 90$. The possibility of occasional mistakes in the platoon leads to a violation of the protocol. The amount of the deduction is linear, and its slope is small. It is the middle piece with the blue line in Figure 4.
- (3) $CV < 50$. We can conclude that this vehicle often makes mistakes, even deliberately violates the protocol. We will make severe punishment for this kind of vehicle, and the penalty amount will increase exponentially with the decrease of CV. It is the first piece with the yellow line in Figure 4.

6. Simulations and Performance Analysis

6.1. Security Analysis. The scheme proposed in this paper enables the opportunistic autonomous vehicle to form a platoon and pass the ETC together. In this environment of mutual distrust, we focus on the attacks of platoon by bogus information injection and repudiation.

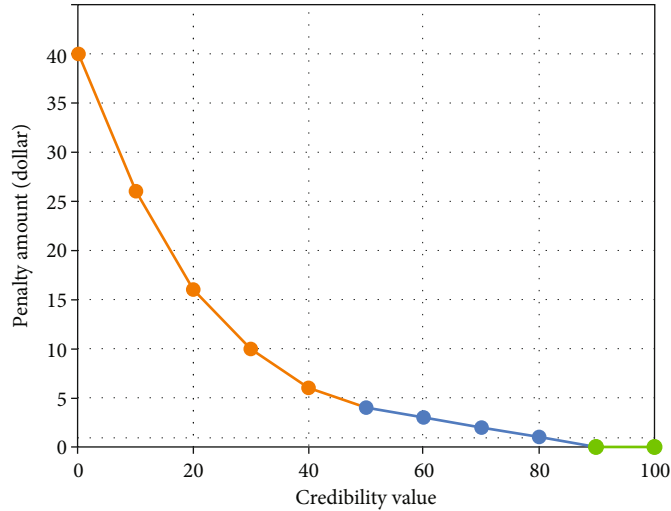


FIGURE 4: Penalty rules. Low credibility value leads to increased penalties.

6.1.1. Bogus Information Injection. Bogus information injections occur when the vehicle or RSU sends the wrong message. When a vehicle wants to join a platoon, it is possible that it just sent a false message to the PL remotely and is not nearby. In the ticket generation phase, the PL must interact with the vehicle to allow it to join the platoon. Therefore, the request of the attacking vehicle to join the platoon will be rejected if it is not in the sensing range.

In the communication of the scheme, bogus information injection will happen with many chances. The PL may send a fake message or instruction inside the platoon. After receiving the commands from the PL, the PM will use its sensing equipment to verify the feasibility of the commands and then decide whether to perform operations. At the same time, all messages inside the platoon will be recorded in the blockchain to prevent members from posting malicious messages.

In the interaction payment protocol, the data to be confirmed would be sent and received between the RSU and the PM, but it needs to be sent to the PL first and then forwarded to the PM by the PL. In our protocol, RSU digitally signs the data with its private key and generates an aggregated signature, ensuring that the PL cannot modify or falsify the data during this period. If the PL intentionally loses the digital signature, this will be detected by the RSU, and then, its credibility value will be lowered, and a penalty will be imposed on it. The RSU calculates the ETC fee for the vehicle. The vehicle's original station and terminal station on the highway are recorded on the blockchain. RSU deliberately adjusts the ETC fee privately and will be detected through traceability of the blockchain.

6.1.2. Repudiation. In the opportunistic autonomous vehicle scenario, repudiation occurs between mutually distrusted vehicles. They refused to admit having done something. Our scheme is based on blockchain, and repudiation is no longer a problem. Because the blockchain is immutable, any operation on the blockchain will be recorded for verification.

After the platoon passed the ETC, the PL temporarily paid the ETC fees for all PMs, and any PM may refuse to

admit that it owed the PL some expenses. The traceability of the blockchain keeps the transaction of the PL payment permanently on the blockchain. The PL can download the transaction for use as a credential.

In the interaction payment protocol, when the platoon is passing ETC, ETC only needs to detect the PL and ignore the PMs to complete the payment transaction. However, at this time, there will be a vehicle outside the platoon closely following the platoon through the ETC to evade payment. ETC did not detect the vehicle, and the transaction for the vehicle deduction did not occur. After a few minutes, the RSU could not detect the transaction of the vehicle deduction on the blockchain. When the RSU communicates with the vehicle and requests to recomplete the protocol, the vehicle has passed the ETC, and the protocol cannot be completed. In this case, the vehicle violates the normal execution of the protocol. We will reduce its credibility value and make a fine. Moreover, for the penalty for the current highway penalty mechanism, the PM for this situation is charged on the farthest distance of the current highway for later punishment.

6.2. Experimental Evaluation. We design a detailed experimental evaluation of each functional module. All of the experiments are the result of averaging 100 trails. The experiment environment consists of an Ubuntu 18.04 laptop equipped with an Intel Core i5-4590 CPU @ 3.30 GHz (4 virtual cores), 4 GB RAM, and an Ubuntu 18.04 workstation equipped with an Intel Core i5-7200U CPU @ 2.50 GHz (4 virtual cores), 8 GB RAM. The workstation is used to simulate the Ethereum environment and run the smart contract. The laptop performs as the Ethereum node client.

We use Ganache CLI to simulate the Ethereum environment, which applies `ethereumjs` to fulfill all Ethereum client behaviors. It does not require computational effort to mine the blocks, which makes it easier to run smart contracts written in the Solidity language and node clients using C++. The interaction between blockchain and node is realized by `QJsonRpc`, which is a Qt implementation of the JSON-RPC

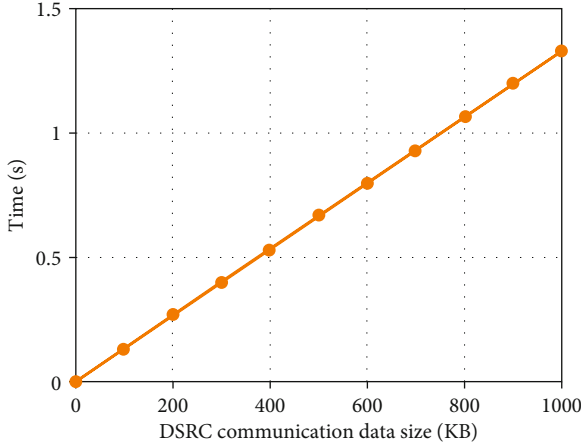


FIGURE 5: Time consumption of DSRC.

protocol (remote procedure call protocol). We can complete all the Ethereum operations using these tools.

6.2.1. Time Consumption for DSRC. In our scheme, DSRC technology is used in many scenarios, including data communication between the platoon members and RSUs. To reach the efficiency requirements of autonomous vehicles, we test the time consumption of DSRC wireless communication technology. According to the 802.11p standard, we set the parameters as $\text{ChannelDataRate} = 6 \text{ Mbps}$ and $\text{PropagationDelay} = 2 \mu\text{s}$ in transmission range about 500 m. The theoretical calculation results are recorded in Figure 5. As the data size increases, the DSRC communication delay has a linear extension.

In the interaction payment protocol, the platoon initiates a request for ETC payment to RSU that would confirm the payment between them through DSRC data transferring. We calculate the data time consumption of each function, as demonstrated in Table 2.

6.2.2. Time Consumption for Blockchain. Figure 6 gives the time cost of platoon member registration. The abscissa is the number of members, including one leader. The ordinate is the time of registrations for this platoon. The time consumption of the PL and PM is almost constant. Therefore, the platoon members' registration time increases linearly with the number of platoon members. The histogram of the number of the member(s) = 2 demonstrates that registering a PM requires more time than a PL. That is because the PM should upload the ticket mentioned above to the smart contract when registering. The smart contract requires additional time to verify and store the ticket backup to prevent duplicate registrations.

All communication data in a platoon even through DSRC would be uploaded to the virtual platoon zone in the smart contract for later verification as a deposit. We designed an experiment where the independent variable is the length of data generated randomly. Its unit is KB. The dependent variable is the time required to upload and download data from blockchain's smart contract. After 100 repeated experiments, we obtained the experi-

TABLE 2: Time consumption for each function in interaction payment protocol.

Feature	Request for ETC payment	Request for confirmation	Return confirmation
Data size (KB)	297	110	333
Time (ms)	0.396	0.147	0.444

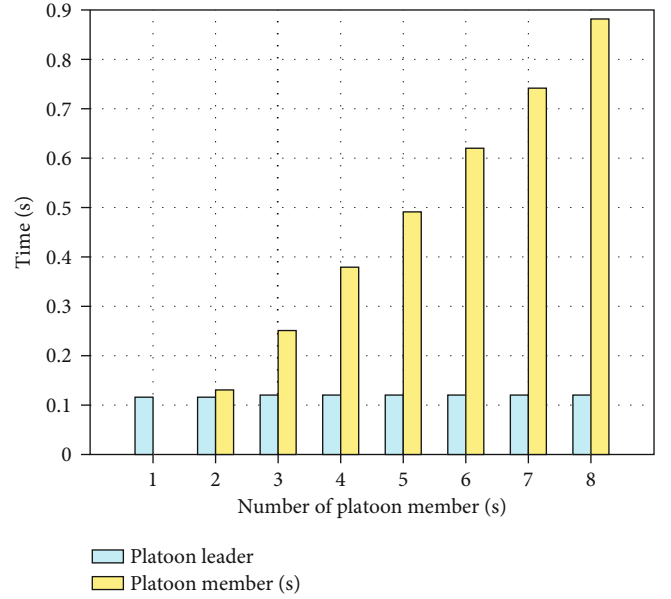


FIGURE 6: Registration time.

mental results described in Figure 7. As the data length increases, the time required increases almost linearly. The client should encode the data before calling function to the smart contract, which requires computational power. The smart contract decodes the data after receiving, and stores it on the smart contract. However, the overhead of the smart contract is considerable. It takes resources to store data on smart contracts, and the overhead increases with the amount of data. Therefore, the time consumption will increase with the data size. In the meantime, the time for downloading data is almost unchanged. Clients just initiate *eth_call* requests to the blockchain that would send corresponding data back, which consumes some query and network time. With the data size increasing, the time for downloading would increase little.

During the interaction payment protocol phase, the PL on behalf of the platoon agreed to deduct fees from RSU through ETC, and then, the platoon could pass ETC. After ETC deducts the platoon, it will upload transaction records to the blockchain for deposition inquiry as required. Ethereum's public chain has low efficiency and instability when linking new blocks. In this case, the platoon can query transaction records after a while. Figure 8 records the time for linking ten consecutive randomly selected blocks. The abscissa is each block index, and the ordinate is the consensus time consumed when the block was generated.

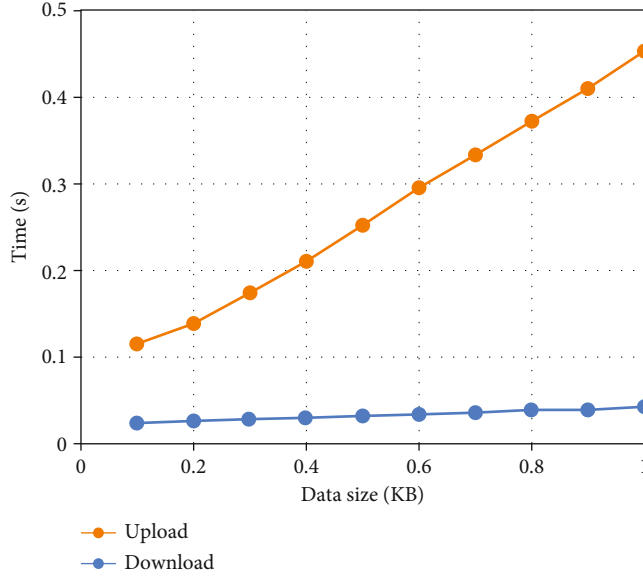


FIGURE 7: Data transmission time in blockchain.

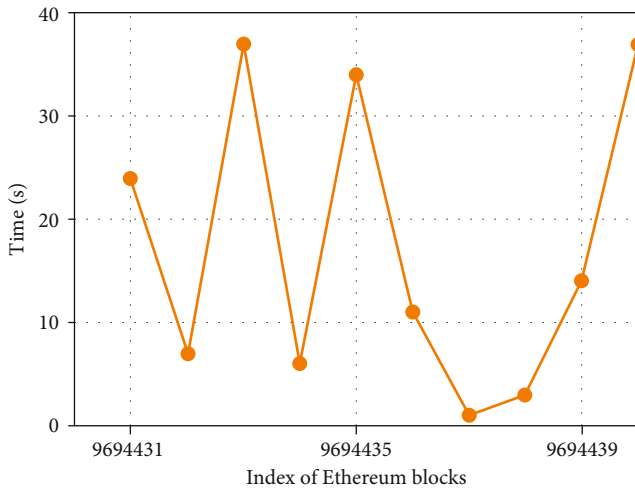


FIGURE 8: Ten consecutive Ethereum blocks' consensus time.

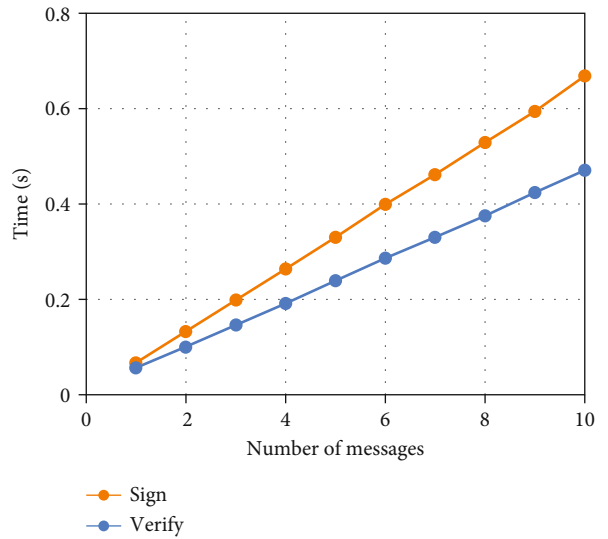


FIGURE 9: Aggregate signature time consumption.

6.2.3. *Time Consumption for Aggregate Signature.* The aggregate signature aggregates the digital signatures obtained by signing each user's message. The experimental horizontal coordinate we designed is the number of randomly generated messages, which can also be understood as the number of users. In our proposed method, it is the number of all platoon members. Each member use its secret key x and a message $M_i \in \{0, 1\}^*$ to compute $h_i \leftarrow H(M_i)$, $\sigma_i \leftarrow h^x$. The signature is σ_i . The randomly generated message is digitally signed, and we can obtain the average time required for a single signature from the repeated experiments. We record the average signing time as T_{sign} . The process of aggregating all the signatures σ_i only takes time T_{aggr} to compute $\sigma \leftarrow \prod_{i=1}^k \sigma_i$. This computation does not require too much computing power compared with the power consumed by digital signatures.

Let n be the number of messages, and then, the summary of the consumed time is

$$T_{\text{sum}} = n \times T_{\text{sign}} + (n - 1) \times T_{\text{aggr}} = (T_{\text{sign}} + T_{\text{aggr}}) \times n - T_{\text{aggr}}. \quad (4)$$

As shown in Figure 9, the summary time required to digitally sign messages and aggregate the digital signatures increases linearly with the number of messages. The reason for this increase is that it consumes much computational power when digitally signing and aggregating each message.

With the number of original messages generating aggregate signature increasing, the process of verifying the correctness of the aggregate signature also consumes

more time. The number and time have a linear relationship. In the process of aggregate signature verification, we use an aggregate signature σ indexed as before for members, the original messages $M_i \in \{0, 1\}^*$, and public keys v_i for members to compute all $h_i \leftarrow H(M_i)$ and then judge if $e(\sigma, g_2) = \prod_{i=1}^k e(h_i, v_i)$. Here, we need to hash each message. Because the message is randomly generated, multiple experiments can be performed to get the average time T_{hash} required for each hash algorithm. In the $\prod_{i=1}^k e(h_i, v_i)$ operation before the correctness, the multiplication requires T_{multi} , and the total time required to verify n messages is

$$T_{\text{sum}} = n \times T_{\text{hash}} + (n - 1) \times T_{\text{multi}} = (T_{\text{hash}} + T_{\text{multi}}) \times n - T_{\text{multi}}. \quad (5)$$

The performance in the graph is the linear relationship of growth.

6.2.4. Performance of Scheme. Our proposed scheme enables the platoon to pass ETC at one time, and we compare this scheme with the ETC for single vehicle in the actual scenario.

According to the highway ETC standard, the vehicle should comply with the rules of vehicle speed v and distance D_v between two vehicles when passing the ETC. Assume that m cars build a platoon, and the distance between the vehicles inside the platoon is D_p . For n cars with an average length of l , we compare the time T_{platoon} required to pass the ETC after they build a platoon with the time T_{vehicle} required for a single vehicle to pass ETC:

$$\begin{aligned} T_{\text{vehicle}} &= \frac{n \times l + (n - 1) \times D_v}{v} = \frac{n \times (l + D_v) - D_v}{v}, \\ T_{\text{platoon}} &= \frac{(n/m) \times [m \times l + (m - 1) \times D_p] + ((n/m) - 1) \times D_v}{v} \\ &= \frac{n \times (l + D_p) + (n/m) \times (D_v - D_p) - D_v}{v}. \end{aligned} \quad (6)$$

According to the standard parameters of ETC, let $v = 20$ km/h, $D_v = 35$ m, $D_p = 15$ m, and $l = 5$ m, and assume $m = 8$ cars in a platoon. For different n , the experimental results are shown in Figure 10. The time spent by single vehicles passing through ETC and platoon through ETC increases linearly with the number of vehicles, which is consistent with the calculation result of the formula. When the number of vehicles is the same, it is evident that the platoon consumes less time. As the number of vehicles increases, the time gap between the two lines becomes larger and larger. From this, we can conclude that our scheme saves the time overhead of ETC payments on the highway.

7. Conclusions

In this paper, we propose an efficient highway toll paradigm based on blockchain for opportunistic autonomous vehicle platoon. Vehicles can autonomously build a platform to form a virtual secure communication area relying on blockchain

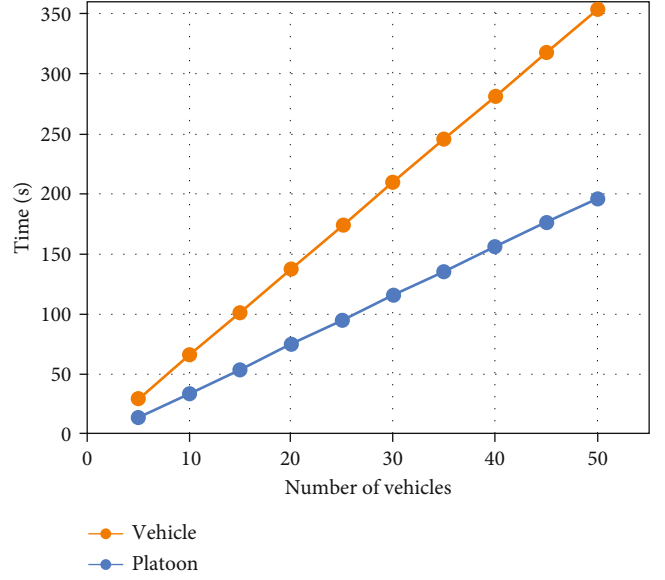


FIGURE 10: ETC passing time consumption.

technology. The platoon interacts with RSU to complete the ETC payment preparation phase, and then, the platoon leader leads all platoon members to pass the ETC for finishing the payment quickly. For the bogus information injection and repudiation attacks that may occur between mutually untrusted vehicles, we conducted a detailed security analysis to conclude that our designed protocol can defend against these attacks. Moreover, the experimental results show that the scheme is highly efficient for autonomous vehicles and dramatically reduces the time for ETC deductions.

Data Availability

The simulation data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the IAF-PP WP5: Design of Security Functionality for V2X Networks Grant for the project of A19D6a0053 by the Agency for Science, Technology and Research, Singapore.

References

- [1] A. Driving, *Levels of Driving Automation are Defined in New SAE International Standard J3016: 2014[J]*, SAE International, Warrendale, PA, USA, 2014.
- [2] B. Li, "Stochastic modeling for vehicle platoons (I): Dynamic grouping behavior and online platoon recognition[J]," *Transportation Research Part B: Methodological*, vol. 95, pp. 364–377, 2017.
- [3] A. K. Bhoopalam, N. Agatz, and R. Zuidwijk, "Planning of truck platoons: a literature review and directions for future

- research,” *Transportation Research Part B: Methodological*, vol. 107, pp. 212–228, 2018.
- [4] B. Besselink, V. Turri, S. H. van de Hoef et al., “Cyber–physical control of road freight transport,” *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1128–1141, 2016.
- [5] V. Sokolov, J. Larson, T. Munson, J. Auld, and D. Karbowski, “Platoon formation maximization through centralized routing and departure time coordination,” 2017, <http://arxiv.org/abs/1701.01391>.
- [6] T. Zeng, O. Semiari, W. Saad, and M. Bennis, “Joint communication and control for wireless autonomous vehicular platoon systems,” *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7907–7922, 2019.
- [7] S. Wen and G. Guo, “Sampled-data control for connected vehicles with Markovian switching topologies and communication delay,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 2930–2942, 2020.
- [8] A. Alam, B. Besselink, V. Turri, J. Mårtensson, and K. H. Johansson, “Heavy-duty vehicle platooning for sustainable freight transportation: a cooperative method to enhance safety and efficiency,” *IEEE Control Systems Magazine*, vol. 35, no. 6, pp. 34–56, 2015.
- [9] N. Boysen, D. Briskorn, and S. Schwerdfeger, “The identical-path truck platooning problem,” *Transportation Research Part B: Methodological*, vol. 109, pp. 26–39, 2018.
- [10] S. Gong, J. Shen, and L. Du, “Constrained optimization and distributed computation based car following control of a connected and autonomous vehicle platoon,” *Transportation Research Part B: Methodological*, vol. 94, pp. 314–334, 2016.
- [11] S. Gong and L. Du, “Cooperative platoon control for a mixed traffic flow including human drive vehicles and connected and autonomous vehicles,” *Transportation Research Part B: Methodological*, vol. 116, pp. 25–61, 2018.
- [12] C. Wang, S. Gong, A. Zhou, T. Li, and S. Peeta, “Cooperative adaptive cruise control for connected autonomous vehicles by factoring communication-related constraints,” *Transportation Research Part C: Emerging Technologies*, vol. 38, pp. 242–262, 2019.
- [13] S. Gong, A. Zhou, and S. Peeta, “Cooperative adaptive cruise control for a platoon of connected and autonomous vehicles considering dynamic information flow topology,” *Transportation Research Record*, vol. 2673, no. 10, pp. 185–198, 2019.
- [14] M. Wagner and B. McMillin, “Cyber-physical transactions: a method for securing VANETs with blockchains,” in *2018 IEEE 23rd Pacific Rim International Symposium on Dependable Computing (PRDC)*, pp. 64–73, Taipei, Taiwan, December 2018.
- [15] B. Ledbetter, S. Wehunt, M. A. Rahman, and M. H. Manshaei, “LIPs: a protocol for leadership incentives for heterogeneous and dynamic platoons,” in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, vol. 1, pp. 535–544, Milwaukee, WI, USA, July 2019.
- [16] J. A. L. Calvo and R. Mathar, “Secure blockchain-based communication scheme for connected vehicles,” in *2018 European Conference on Networks and Communications (EuCNC)*, pp. 347–351, Ljubljana, Slovenia, June 2018.
- [17] Y. Zhang, J. Weng, J. Weng, M. Li, and W. Luo, “Onionchain: towards balancing privacy and traceability of blockchain-based applications,” 2019, <http://arxiv.org/abs/1909.03367>.
- [18] L. Li, J. Liu, L. Cheng et al., “Creditcoin: a privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2204–2220, 2018.
- [19] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, and J. Zhao, “Toward secure blockchain-enabled Internet of vehicles: optimizing consensus management using reputation and contract theory,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2906–2920, 2019.
- [20] L. Cheng, J. Liu, G. Xu et al., “SCTSC: a semicentralized traffic signal control mode with attribute-based blockchain in IoVs,” *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1373–1385, 2019.
- [21] G. Wood, “Ethereum: a secure decentralised generalised transaction ledger,” *Ethereum Project Yellow Paper*, vol. 151, no. 2014, pp. 1–32, 2014.
- [22] J. Xu, L. Wei, Y. Zhang, A. Wang, F. Zhou, and C. Z. Gao, “Dynamic fully homomorphic encryption-based Merkle tree for lightweight streaming authenticated data structures,” *Journal of Network and Computer Applications*, vol. 107, pp. 113–124, 2018.
- [23] D. Johnson, A. Menezes, and S. Vanstone, “The elliptic curve digital signature algorithm (ECDSA),” *International Journal of Information Security*, vol. 1, no. 1, pp. 36–63, 2001.
- [24] L. Cheng, Q. Wen, Z. Jin, H. Zhang, and L. Zhou, “Cryptanalysis and improvement of a certificateless aggregate signature scheme,” *Information Sciences*, vol. 295, pp. 337–346, 2015.
- [25] Z. Ying, M. Ma, and L. Yi, “BAVPM: practical autonomous vehicle platoon management supported by blockchain technique,” in *2019 4th International Conference on Intelligent Transportation Engineering (ICITE)*, Singapore, Singapore, September 2019.

Research Article

Achieving Privacy-Preserving Group Recommendation with Local Differential Privacy and Random Transmission

Hanyi Wang ^{1,2}, Kun He,^{1,2} Ben Niu ¹, Lihua Yin,³ and Fenghua Li^{1,2}

¹*Institute of Information Engineering, Chinese Academy of Sciences, China*

²*School of Cyber Security, University of Chinese Academy of Sciences, China*

³*Cyberspace Institute of Advanced Technology (CIAT), Guangzhou University, China*

Correspondence should be addressed to Ben Niu; niuben@iie.ac.cn

Received 16 March 2020; Revised 5 June 2020; Accepted 31 July 2020; Published 5 September 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Hanyi 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.

Group activities on social networks are increasing rapidly with the development of mobile devices and IoT terminals, creating a huge demand for group recommendation. However, group recommender systems are facing an important problem of privacy leakage on user's historical data and preference. Existing solutions always pay attention to protect the historical data but ignore the privacy of preference. In this paper, we design a privacy-preserving group recommendation scheme, consisting of a personalized recommendation algorithm and a preference aggregation algorithm. With the carefully introduced local differential privacy (LDP), our personalized recommendation algorithm can protect user's historical data in each specific group. We also propose an Intra-group transfer Privacy-preserving Preference Aggregation algorithm (IntPPA). IntPPA protects each group member's personal preference against either the untrusted servers or other users. It could also defend long-term observation attack. We also conduct several experiments to measure the privacy-preserving effect and usability of our scheme with some closely related schemes. Experimental results on two datasets show the utility and privacy of our scheme and further illustrate its advantages.

1. Introduction

With the development of the social networks and the mobile devices like smartphone or IoT terminals [1, 2], recommender systems, which are used to recommend for each individual, play an important role in our daily life. Nowadays, more and more people tend to gather together to enjoy social activities, such as watching movie or hiking together, having dinner with friends, or using apps such as Meetup or Douban to find out group activities with common-interest-strangers. With the latest development and popularity of smart devices and social networking services, it is convenient for people to form a group. Such trends bring new challenges to the existing recommender system: which items or events (e.g., movie/attraction/restaurant) should be recommended, in order to satisfy all/most of the group members? Under this circumstance, group recommendation has gradually become one of the hotspots in the field of recommender systems [3–8].

However, the privacy issue is obstructing the development of recommender system. In 2018, about 87 million Facebook users' information was leaked to Cambridge Analytica company. By utilizing such information, the company built user model and obtain users' personal preferences. Based on these preferences, the company recommended targeted promotion content about the US election to these users. It affects users' votes to some extent, causing serious violations of human rights. Therefore, we should pay much attention to protect users' privacy on recommender systems, including both of their historical data and personal preference.

Many privacy-preserving approaches have been proposed to protect users' historical data during recommendation over recent years. Yang et al. [9] proposed a framework called PrivRank. In PrivRank, both historical data and activity data of users are protected by obfuscation. In the mean time, these obfuscated data could provide high-quality personalized ranking-based recommendation services. In the

eHealthcare system, Xu et al. [10] proposed a privacy-preserving online medical service recommendation scheme. This scheme uses Paillier encryption algorithm to match each patient's needs with the information of doctors. It recommends appropriate doctors to patients, without knowing patients' exact demands.

Although these work protect users' historical data during recommendation, they ignore that results of personalized recommendation also leak user's privacy. For example, if an untrusted server learns a specific user's preference on restaurants, without knowing the exact restaurants that the user went to, the server could easily figure out user's food taste, places that he or she usually went, and how much he or she may cost on food. And the server probably infers user's salary or even home address according to these preferences.

For this concern, we take the privacy of personal preference into consideration. In that case, there are three challenges we need to solve. First, we need to hide both historical data and preferences of users from honest but curious parties during the recommendation. We also need to guarantee the accuracy of recommendation results at the same time. Second, since a group is dynamic, we should focus on the long-term observation attacks launched by the malicious users in the group recommendation. Third, how to evaluate our recommendation scheme is also a tricky issue.

The main contributions of our paper are summarized as follows:

- (i) We propose a privacy-aware group recommendation scheme that protects each user's historical data and personal preferences at the same time. The scheme consists of two algorithms. The former focuses on the personalized recommendation problem, which guarantees ϵ -LDP for each user. The latter is a privacy-preserving preference aggregation algorithm called IntPPA, designed to solve the group recommendation problem. By adding noise on each member's preference profile, and transferring them within the group, the user's preferences can be protected, and the accuracy of the group recommendation is also guaranteed.
- (ii) In IntPPA algorithm, no one could easily infer the members' preferences through long-term observation attacks. We also adopt a median aggregation strategy to prevent malicious group members from tampering data.
- (iii) We conduct several experiments on two real-world datasets. We use RMSE and F-score to measure the utility, and use RMSE and "matched pairs" to measure the privacy-preserving effect under long-term observation attacks. The results show our scheme has good utility and privacy-preserving effect.

The rest of this paper is organized as follows. In Section 2, we review some recent work. Section 3 introduces the preliminaries. In Section 4, we elaborate the details of our proposed methods. We also demonstrate our evaluation results in Section 5. Finally, we conclude our paper in Section 6.

2. Related Work

2.1. Privacy-Preserving Personalized Recommendation. Differential privacy (DP) has been widely utilized in the personalized recommender systems to protect users' privacy. Under DP, the recommender systems could not infer a user's profile from the changes of recommendation results. McSherry and Mironov [11] proposed a differential privacy method based on collaborative filtering algorithm. They added noise not only on the sum, count, or average but also on the covariance matrix and then sent it to the recommender system. Even though the method could protect users' privacy, the performance is of low utility since too much noise has been added. Then, Liu et al. [12] and Balu and Furon [13] also proposed recommendation algorithms based on DP and improved the performance. However, most of them assumed that the recommender system is a trusted third party.

Hua et al. [14] achieved differentially private matrix factorization based on a gradient descent algorithm, which prevents users' private ratings from the untrusted recommender. Shen et al. [15] proposed a differentially private framework without any trusted third parties. By solving optimization problems, the method could perturb users' historical records, and guarantee users' category preferences under ϵ -DP.

Some solutions utilized local differential privacy to construct privacy-preserving recommender systems. Shin et al. [16] proposed a personalized recommendation algorithm based on matrix factorization and used a LDP mechanism on the gradient of users' profiles. The algorithm not only protected users' ratings, but also the items that were rated by the users. Shin et al. [16] also enhanced their algorithm via dimension reduction in order to improve the accuracy and decrease the computing costs. Since this algorithm obtains recommendation results on the user side, only users can learn the recommended result.

2.2. Group Recommendation. Group recommendations try to satisfy a group of users' preferences. It could be applied in various areas such as video services [17], shopping [18], traveling [19], having dinner [20], and even in IoT scenario [21]. How to extract the common preferences of group members, reduce the preference conflict among group members, and make the recommendation results satisfy the needs of all group members as much as possible, are the key problems in group recommendation system.

Solutions to group recommendation are usually divided into two categories. One category aggregates the profiles of the group members into one profile, then regards it as a user profile, and makes personalized recommendations based on the aggregated profile [22, 23]. On the contrary, the other methods first makes personalized recommendations for each group member, respectively, and then aggregates their recommendation results as the group recommendation result by aggregation strategies [24, 25]. Comparing with the former methods, the latter ones have better flexibility and also obtain more attentions.

A few work solve the privacy problems in group recommendation. Luo and Chen [26] proposed a group-based privacy-preserving method. The algorithm perturbed the

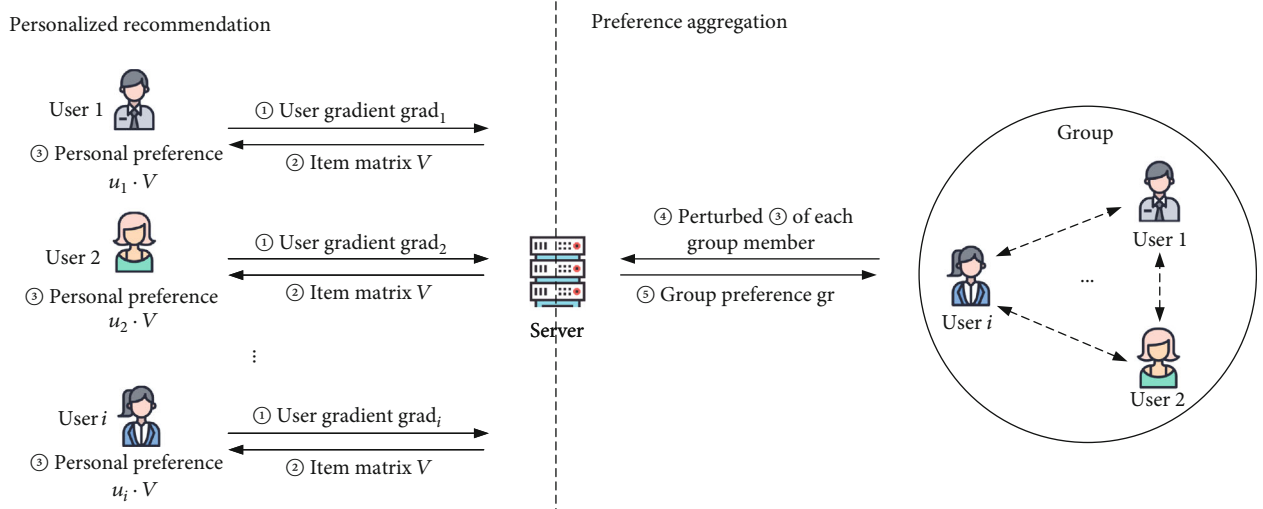


FIGURE 1: Framework of our proposed scheme.

preference data of each group member and then utilized CF-based group recommendation to aggregate their data. Shang et al. [27] proposed a ranking-based privacy-preserving group recommendation method. Group members exchanged their obfuscated profiles to protect sensitive information. And they constructed a hybrid collaborative filtering model based on Markov random walks to provide recommendations and predictions to group members.

3. Preliminaries

3.1. Motivation and Basic Idea. There is a huge privacy risk in personalized recommendation. Once the attacker collects the user's historical data in the recommendation, the user's sensitive information could be inferred. For example, we often rate the restaurants we have visited in the Dianping or Yelp. According to these data, the attacker can easily infer the user's historical geographical location and even further speculate the user's financial conditions and home address.

On the other hand, with the development of social network, the interaction between users increases. Recommender systems are no longer limited to recommending items to one user, but gradually extended to group-oriented services.

Most group recommendation algorithms need to integrate the personalized recommendation results of each group members, so the risks in personalized recommendation also hold true for group recommendation. However, in addition to historical data, users' personalized recommendation results are also very sensitive. Personalized recommendation results reflect users' preferences and behavior habits, involving food preferences, entertainment interests, social network relationships, political tendencies, and other aspects, thus forming a "user portrait" of people.

Existing privacy-preserving algorithms of recommendation system only protect user's historical data but ignore the disclosure risk of personalized recommendation results. Moreover, a group tends to change and update, so the attacker is likely to analyze more user information through the long-term observation. How to further protect the per-

sonalized recommendation results and resist the long-term observation attack in group recommendation is the main problem to be solved in this paper.

We design a privacy-aware group recommendation scheme. Figure 1 further illustrates our basic idea, which could be divided into two parts. Steps 1, 2, and 3 show the personalized recommendation part. Users interact with the recommendation server under local differential privacy, in order to train the personalized recommendation model via matrix factorization mechanism without privacy leakage, and then the personalized recommendation results will be generated on the client side. After the personalized recommendation part, each user obtains his or her own preferences. Steps 4 and 5 represent the preference aggregation part. When a group needs a recommendation, the group will execute the IntPPA algorithm. Each group member's personal preference is transferred inside the group with perturbation before sending to the server. In this case, the real preferences of group members are hidden during the random transmission and will not be exposed to others. After receiving all members' perturbed preferences, the server adopts the median aggregation strategy to fuse them and obtains the final group preference.

3.2. Local Differential Privacy. Local differential privacy is a state-of-the-art notion that guarantees users to share private data safely.

Definition 1. A privacy mechanism \mathcal{M} satisfies ϵ -local differential privacy ($\epsilon \geq 0$) if for any two different records $t, t' \in \text{Domain}(\mathcal{M})$, and for any output $t^* \in \text{Range}(\mathcal{M})$,

$$\Pr(\mathcal{M}(t) = t^*) \leq e^\epsilon \Pr(\mathcal{M}(t') = t^*). \quad (1)$$

Local differential privacy has a sequential composability property same as differential privacy. Sequential composability means the privacy budget ϵ could be assigned in different steps of an algorithm.

Property 2. For n privacy-preserving algorithms \mathcal{M}_i , $1 \leq i \leq n$. If algorithm \mathcal{M}_i satisfies ε_i -local differential privacy, algorithm $M = \{\mathcal{M}_1, \dots, \mathcal{M}_n\}$ satisfies ε -LDP, which $\varepsilon = \sum_{i=1}^n \varepsilon_i$.

3.3. Problem Statement. We take the movie recommender system as an example to introduce our scheme in this paper. We assume that there is a movie recommender system consisting of m movies and n users. We denote the rating generated by user i for movie j as $b_{i,j}$. If user i has not rated movie j , $b_{i,j} = \text{null}$. The movie recommender system has many groups that are formed by users. We denote a typical group with g users as G . The notations are listed in Table 1.

In a group recommender system, it first collects users' historical data b_i to predict preference rating b_i^* for movies that are not rated by users and then aggregates $b_{i,j}^*$ (for all $i \in G$) for each movie j to obtain group preference rating gr_j for movie j . Usually, the sensitive historical ratings and preference ratings will be transmitted between users and server without protection, which increases the risk of users' privacy leakage. In our work, we assume that the server is untrusted and each user only trusts themselves. And the adversaries are assumed to have access to the output data of all users and know the privacy-preserving scheme adopted by the users. Since groups tend to change and update, attackers also try to find out the connection between users' identities and their sensitive data by carrying out the long-term observation attacks. In these cases, how to build a group recommender system without privacy leakage is our main problem.

We are facing several challenges to solve the problem. First, how to protect users' personal preference ratings $b_{i,j}$ from leakage to others except the rating owner and guarantee the accuracy of group recommendation results at the meantime? Second, how to defend against long-term observation attacks? Third, how to measure the privacy-preserving effect against long-term observation attack?

4. Our Proposed Scheme

4.1. Personalized Recommendation under Local Differential Privacy Algorithm. To fully protect users' personal profiles, the privacy-preserving algorithm should not only prevent users' historical data from leakage but also need to pay attention to their personalized recommendation results, during the personalized recommendation step. So based on the algorithm proposed by Shin et al. [16], we refine the matrix factorization method to fit in with our group recommendation scenario.

To solve the matrix factorization problem without overfitting, we minimize the following formula:

$$\sum_{i,j} \frac{y_{i,j} (b_{i,j} - u_i v_j)^2}{N} + \lambda \left(\sum_{i=1}^n \|u_i\|^2 + \sum_{j=1}^m \|v_j\|^2 \right). \quad (2)$$

In formula (2), user vector $u_i = (u_{i,1}, u_{i,2}, \dots, u_{i,d})$ represents user i 's relations with d latent factors separately. These factors also have inter-dependencies with movies, which are denoted as item vectors $v_j = (v_{1,j}, v_{2,j}, \dots, v_{d,j})^T$ for movie j .

TABLE 1: Notations.

Notation	Meaning
n	Number of users
m	Number of movies
d	Number of latent factors
$b_{i,j}$	Rating generated by user i for movie j , which could be <i>null</i>
b_i^*	User i 's preference ratings predicted by personalized recommendation under LDP, $b_i^* = (b_{i,1}^*, \dots, b_{i,j}^*, \dots, b_{i,m}^*)$
G	A group
g	Number of group members
gr_j	Group preferences of movie j
k	Number of iterations
N	Total number of ratings
u_i	User i 's user vector
v_j	Movie j 's item vector
λ	Regularization coefficient
γ	Learning rate
ε	Privacy budget
∇_v	Gradient matrix for V at each iteration
Δf	Change value during the transmission in IntPPA
P_0	The probability that user's profile is sent to group members
$t_{\text{start}} (t_{\text{end}})$	Group members need to execute IntPPA algorithm after (before) this time

Matrix $Y = \{y_{i,j}\}_{n \times m}$ represents whether a user has given a rate to a movie. If $b_{i,j} = \text{null}$, $y_{i,j} = 0$, else $y_{i,j} = 1$. $N = \sum_{i,j} y_{i,j}$, which is the total number of the ratings in the system.

We utilize a privacy-preserving gradient descent algorithm to approach the minimum of formula (2).

Users separately interact with the server for k times to achieve the gradient descent algorithm. According to Property 2, if our algorithm needs to realize ε -LDP, then each iteration needs to realize ε/k -LDP. Before each iteration, users and the server need to initialize u_i and V on their own sides. During each iteration, V is sent to each user first, then user i computes a $d \times m$ matrix X_i , where each element $(x_i)_{l,j} = -2 y_{i,j} u_{i,l} (b_{i,j} - u_i v_j)$. It is relative to the gradient of formula (2) on user vectors. According to Shin's work, each $(x_i)_{l,j}$ needs to be normalized into range $[0, 1]$. The min-max normalization is usually used to normalize, but it will leak users' gradient range. So, we propose a normalization method without too much information leakage. Each user i computes $\text{Max}_i = \max(|(x_i)_{0,0}|, \dots, |(x_i)_{d,m}|)$, then $(x_i)_{l,j} = (x_i)_{l,j} / \text{Log}_i$ is normalized in range $[0, 1]$, where $\text{Log}_i = 10^{\lceil \log_{10}(\text{Max}_i) \rceil}$. Then, each user uses LDP algorithm to perturb $(x_i)_{l,j}$. User i randomly chooses two elements $l_i \in [1, d]$, $j_i \in [1, m]$, and computes $c_i = (x_i)_{l_i, j_i}$.

Next, user i perturbs c_i in the following distribution and obtains c'_i . We denote $t(\varepsilon) = md\text{Log}_i(e^{\varepsilon/k} + 1)/(e^{\varepsilon/k} - 1)$:

$$\Pr(c'_i = c_0) = \begin{cases} \frac{c_i(e^{\varepsilon/k} - 1)}{2(e^{\varepsilon/k} + 1)} + \frac{1}{2} & \text{when } c_0 = t(\varepsilon), \\ \frac{-c_i(e^{\varepsilon/k} - 1)}{2(e^{\varepsilon/k} + 1)} + \frac{1}{2} & \text{when } c_0 = -t(\varepsilon). \end{cases} \quad (3)$$

After that, each user i sends c'_i , l_i , j_i to the server, then updates user vector $u_i = u_i - \gamma[-2/N \sum_j y_{i,j} v_j^T (b_{i,j} - u_i v_j) + 2\lambda u_i]$, where γ is the learning rate of the gradient descent algorithm. After receiving information from all users, the server adds each c'_i/N to the row l_i and column j_i of matrix $\nabla_v = \{0\}_{d \times m}$, in order to estimate the gradients of V . Then, server updates $V = V - \gamma(\nabla_v + 2\lambda V)$ and sends it back to all users.

Users and the server interact k times as above, and finally each user i computes $b_{i,j}^* = u_i V$ on their own sides and obtains each items' predicted ratings.

The following theorems illustrate the privacy and usability of this algorithm, which have been proved similarly in [16], so we are not going to prove them here.

Theorem 3. *For each user, the algorithm in Section 4.1 satisfies ε -local differential privacy.*

Theorem 4. *There is $\mathbb{E}(v_{i,j}) = v'_{i,j}$, $v'_{i,j}$ represents the true value of the item matrix under no privacy-preserving algorithm. $v_{i,j}$ represents the perturbed value of item matrix elements under this algorithm.*

Theorem 3 indicates the functionality on the privacy protection of our algorithm, while Theorem 4 shows its utility, since $v_{i,j}$ is the unbiased estimation of $v'_{i,j}$.

4.2. IntPPA Algorithm. This algorithm corresponds to steps 4 and 5 in Figure 1. It consists of an inter-group random transmission and an aggregation strategy. We first introduce why we choose median strategy to be our aggregation strategy.

There are lots of aggregation strategies in group recommendation such as average strategy. Under this strategy, the average of group members' ratings represents the group preference. Also, there is a typical strategy called least misery. Under this strategy, the minimum value of group members' ratings represents the group preference.

Suppose a group consists of five members: Alice, Bob, Charles, David, and Eric. Each member has his or her own preferences on five movies (predicted via personalized recommendation), which is shown in Table 2. The server supposes to select two movies to satisfy most of the members. If the average strategy is chosen, the group preferences for these 5 movies will be 3.2, 3.4, 3.8, 2.8, and 2.8, separately. And Movies 2 and 3 will be recommended to the group. If the least misery strategy is chosen, the group preferences for these 5 movies will be 1, 2, 2, 1, and 1, separately, and

TABLE 2: A rating matrix.

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5
Alice	5	4	5	3	5
Bob	3	4	3	2	3
Charles	5	4	2	1	3
David	2	3	5	3	1
Eric	1	2	4	5	2
Gloria	1	1	1	5	1

the server will also recommend Movies 2 and 3. In this paper, we choose median strategy to measure the group preference. Under this strategy, the median value of members' ratings represents the group preference. So, the group preference will be 3, 4, 4, 3, and 3, separately, and Movies 2 and 3 will be recommended. According to the definition, median strategy is less affected by extreme ratings. We assume there is a dishonest user Gloria, she is a new member of this group, and she wants the server to choose Movie 4, so she changed her predicted ratings into 1, 1, 1, 5, 1. In this case, the group will recommend Movies 3 and 4 under average aggregation strategy. While under least misery strategy, all movies has the same preference, which is indistinguishable for the server, so this strategy is unavailable under this situation. For our median aggregation strategy, Gloria's plan does not work out. Movies 2 and 3 still have the highest ratings, which means Movie 4 will not be recommended. Based on these analysis, it is hard for the malicious users in the group to affect the group preference by changing their predicted scores under median aggregation strategy.

Next, we are going to introduce the random transmission. Suppose there is a group G , who has g group members. Group members aim to send profile b_i^* to the server without privacy leakage, and at the meantime, the server could obtain the group preference of each item. In order to prevent the server from knowing users' personal preferences, we design a random transmission mechanism inside the group, called IntPPA.

Server sets two time parameters: t_{start} and t_{end} . All the members have to start this IntPPA algorithm after time t_{start} and finish before time t_{end} . We propose a perturbed algorithm $pf(b_i^*) = (b_{i,1}^* + f_1, b_{i,2}^* + f_2, \dots, b_{i,m}^* + f_m)$, and for each f_j , $\Pr(f_j = \Delta f) = 0.5$, $\Pr(f_j = -\Delta f) = 0.5$. Δf is a fixed value, and we call it as "change value."

Group members first perturb their profile b_i^* into $b'_i = p f(b_i^*)$. Then, they send perturbed profiles to the server with the probability of $1 - p_0$ and send to each group member with the probability of p_0/g . When a member receives others' profiles, he/she also needs to use the perturbed algorithm to update $b'_i = pf(b'_i)$ first and then sends b'_i to group members or server according to the probability distribution described above. A member's profile will not stop to be transmitted until the profile is sent to the server or the time exceeds t_{end} .

The server finally receives all b'_i and computes group recommendation results $gr = (gr_1, gr_2, \dots, gr_m)$, in which $gr_j = \text{median}(b'_i, b'_{2,j}, \dots, b'_{i,j}, \dots, b'_{g,j})$. gr_j represents the group

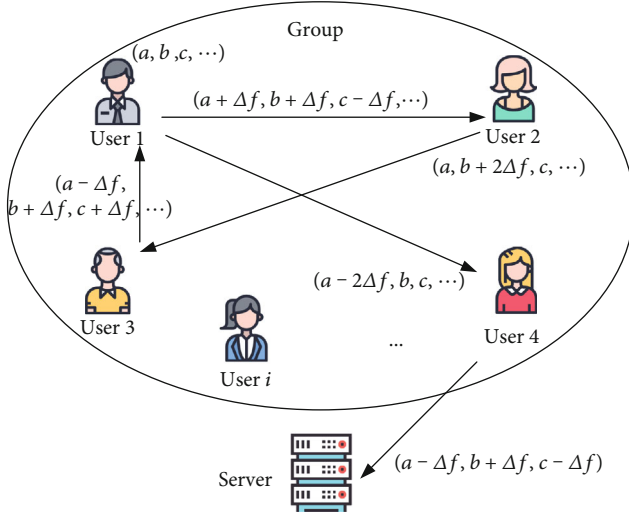


FIGURE 2: IntPPA algorithm process. User 1 has profile (a, b, c, \dots) . During the algorithm, user 1 perturbs the profile into $(a + \Delta f, b + \Delta f, c - \Delta f, \dots)$ and transmits to user 2. Then, user 2 perturbs it into $(a, b + 2\Delta f, c, \dots)$ and transmits to user 3. After that, user 3 transmits $(a - \Delta f, b + \Delta f, c + \Delta f, \dots)$ to user 1, and then user 1 transmits $(a - 2\Delta f, b, c, \dots)$ to user 4. Finally, user 4 submits $(a - \Delta f, b + \Delta f, c - \Delta f, \dots)$ to the server.

preferences of movie j . Figure 2 shows an example of the process of IntPPA. User 1 holds a profile (i.e., personalized recommendation results) (a, b, c, \dots) . He/she first executes the perturbed algorithm and obtains $(a + \Delta f, b + \Delta f, c - \Delta f, \dots)$ and transmits the perturbed profile to user 2. Then, the profile is sent to user 2 \rightarrow user 3 \rightarrow user 1 \rightarrow user 4 \rightarrow server. Finally the server receives $(a - \Delta f, b + \Delta f, c - \Delta f, \dots)$.

Under IntPPA algorithm, neither the server nor users could recognize whom the profiles are exactly from. When a new group member joins in this group, the server also cannot find out the fresh member’s profile, according to the difference between results computed before and after the member’s attendance.

5. Performance Evaluation

5.1. Experimental Settings

5.1.1. Datasets. We evaluate our experiments on two datasets. The first dataset is MovieLens-100k [28], which contains one hundred thousand movie ratings from approximately 1000 users on over 1500 movies. The second dataset is FilmTrust [29]. FilmTrust is a dataset crawled from the entire FilmTrust website. Table 3 shows the details of the two datasets.

5.1.2. Comparison Methods. In Section 2.2, we mentioned several privacy-preserving group recommendation work. However, we did not compare our scheme with theirs. Although their works protect users’ ratings and even could protect users’ preferences to some extent, they did not consider the long-term observation attack during the group recommendation. While we assume a stronger attack model, it is not appropriate to compare with them. In this case, we compare our scheme (priv-MF-IntPPA)

TABLE 3: Dataset statistics.

Dataset	MovieLens-100k	FilmTrust
Total users	943	1508
Total items	1682	2071
Total ratings	100,000	35,497
Range	[3, 25]	[0.5, 4]
Density	6.3%	1.1%

with some baselines in our experiments: priv-MF-med, priv-MF-avg, and priv-MF-lm.

The difference between ours and the other three methods is the aggregation strategy. Priv-MF-med, priv-MF-avg, and priv-MF-lm choose median, average, and least misery aggregation strategies, respectively. They have the same privacy-preserving personalized recommendation methods as ours, while they do not apply privacy protection to the preference aggregation part.

5.1.3. Utility Metrics. We evaluate the data utility of personalized recommendation algorithm by computing the RMSE (root mean squared error) between the item scores predicted by the training set and the actual score of the test set. We evaluate the group recommendation accuracy performance using precision ($\text{Pre}@G_0$) and recall ($\text{Rec}@G_0$). Here, G_0 is the number of the recommended items. We evaluate the experiment with $G_0 = \{5, 10, 15\}$. $\text{Pre}@G_0$ is the fraction of top G_0 recommendations selected by the group, and $\text{Rec}@G_0$ is the fraction of true items retrieved in the top G_0 recommendations.

5.1.4. Privacy Metrics. We use two metrics to evaluate the privacy-preserving effect of IntPPA algorithm. We compute the RMSE between each user’s profile before and after the IntPPA algorithm. We also propose a “matched pair” metric to describe the privacy-preserving effect against long-term observation attack.

We execute the IntPPA algorithm on the same group for multiple times and compute the differences between the perturbed profiles. For example, if user i ’s profile in execution one is closest to user i' ’s in execution two, we will call them as “a pair.” If these two members have the same identity, which means user i is actually user i' , we will call that pair “matched.” We denote the number of matched pairs in a group as mp and normalize it to interval $[0, 1]$. Obviously, the smaller mp is, the better the privacy protection performs against the long-term observation attack.

5.2. Evaluation Results. We follow the evaluation method of privacy protection proposed in Li et al. [30] to measure our scheme.

We first did some pilot experiments to screen appropriate parameters of the method. According to the pilot experiments, we set number of latent factors $d = 10$ and number of iterations $k = 10$. We also set $\gamma = 6 \times 10^{-6}$, $\lambda = 2 \times 10^{-3}$ in MovieLens’ experiments and $\gamma = 8 \times 10^{-6}$, $\lambda = 2 \times 10^{-3}$ in FilmTrust’s.

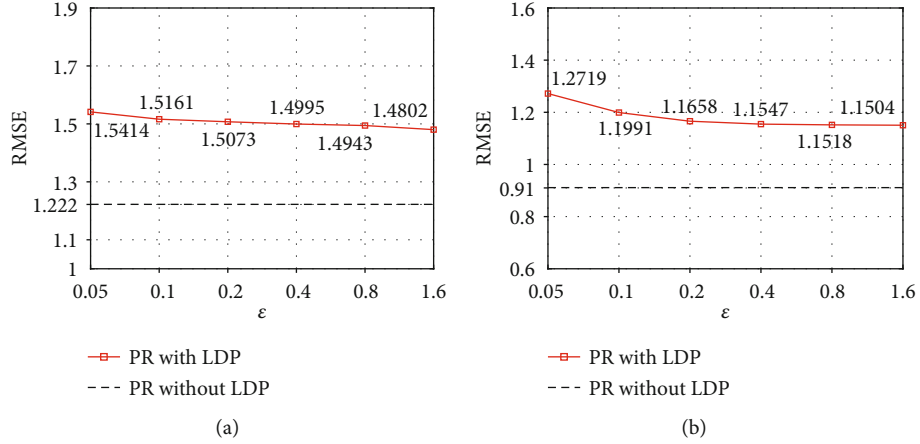


FIGURE 3: Utility of personalized recommendation vs. ϵ . (a) MovieLens. (b) FilmTrust.

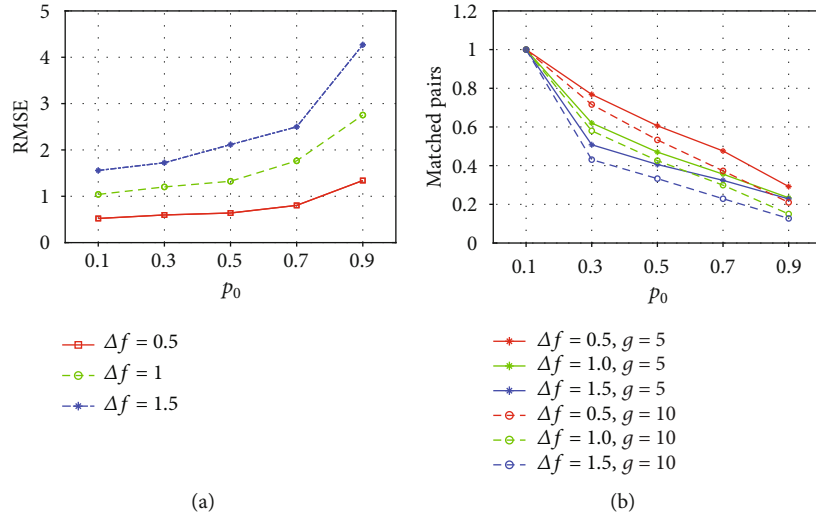


FIGURE 4: Privacy-preserving effect of IntPPA algorithm. (a) RMSE. (b) Matched pairs.

5.2.1. *Personalized Recommendation Utility Analysis.* We divide the dataset into five folds, using four folds as training set and the other as testing set. We use RMSE to measure the utility of personalized recommendation algorithm. In Figures 3(a) and 3(b), the dotted line displays the utility of non-privacy matrix factorization algorithm, while the red line displays the utility of ours. In both datasets, the red line approaches to the dotted line as the privacy budget ϵ increases, which means ϵ has a negative correlation with the utility of our personalized recommendation algorithm. As we know, ϵ has a positive correlation with the privacy-preserving effect. So, in order to balance the privacy and utility, we set $\epsilon=0.4$ for both datasets in the following experiments.

5.2.2. *IntPPA Algorithm Privacy-Preserving Effect Analysis.* We utilize two methods to measure the privacy-preserving effect of IntPPA. Since these measurements are only relevant with parameter Δf , p_0 , and g , we test it on MovieLens dataset.

Figure 4(a) shows that larger Δf or larger probability p_0 achieves better performance on privacy protection. In terms of parameter selection, when the data in the dataset has wider range, we need to increase Δf or p_0 to guarantee a suitable perturbation.

Second, we utilize “matched pairs” mp to measure the privacy. Figure 4(b) shows that when transfer probability p_0 increases, “matched pairs” will decrease, which means a better performance on privacy protection. Change value Δf and group size g are also negatively correlated with mp . However, once p_0 is smaller than 0.1, the IntPPA algorithm could not resist long-term observation attack.

According to the above experiments, we choose parameter $\Delta f=0.5, p_0=0.7$ for both MovieLens and FilmTrust datasets.

5.2.3. *Group Recommendation Accuracy Analysis.* Since both MovieLens and FilmTrust do not contain group information, we extract groups that are randomly chosen from the users to build groups.

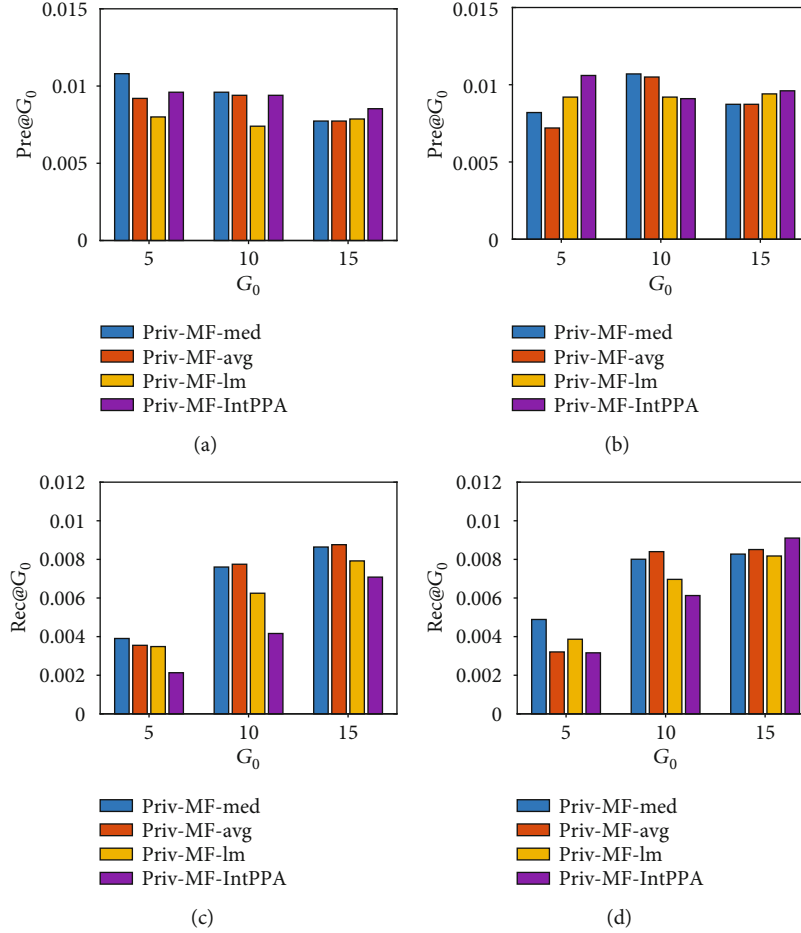


FIGURE 5: Performance of group recommendation methods for the MovieLens dataset. (a) Group size = 5. (b) Group size = 10. (c) Group size = 5. (d) Group size = 10.

We use $\text{Pre}@G_0$ and $\text{Rec}@G_0$ to evaluate our scheme's recommendation accuracy. In MovieLens (FilmTrust), if a group member gives score 4 (3) or above to a movie, we assume that the movie is adopted by the member.

Figures 5 and 6 report the $\text{Pre}@G_0$ and $\text{Rec}@G_0$ values for the two datasets with $G_0 = \{5, 10, 15\}$ and group size $g = \{5, 10\}$. We compare our scheme (priv-MF-IntPPA) with the other group recommendation algorithms (priv-MF-med, priv-MF-avg, and priv-MF-lm) described in Section 5.1. We observe from the two figures that

- (i) For priv-MF-IntPPA, when G_0 becomes larger, $\text{Pre}@G_0$ has a trend to become smaller, but $\text{Rec}@G_0$ becomes larger.
- (ii) Group size g has little effect on the recall and precision value of our group recommendation.
- (iii) In the MovieLens dataset, the precision of our scheme is close to others', and sometimes even be better than them due to the randomness. However, the recall of our scheme is a little weak. In Figure 5(c), when $G_0 = 5$, the recall values of priv-MF-IntPPA and priv-MF-med are 0.0021 and

0.0039. Priv-MF-med improves 85.7% over our scheme. But when $G_0 = 15$, the recall values are 0.0070 and 0.0086; priv-MF-med only improves 22.9% over our scheme. So, as group size becomes larger, the differences will be acceptable.

- (iv) In the FilmTrust dataset, the precision and recall values of each method are much lower than the values in the MovieLens dataset. It is probably because FilmTrust has lower rating density. Figure 6(c) shows that our scheme performs much better than other methods, which means that our algorithm is more suitable for low sparsity data than other algorithms.

From the above, our scheme can provide accurate group recommendation under the premise of ensuring privacy.

5.3. Communication and Time Cost. As for personalized recommendation, we only analyze the communication cost in one iteration. For each user, no matter how many items the user rates, only three elements are transmitted to the server, which is about 24 B for both MovieLens and FilmTrust datasets. During each iteration, server transmits a matrix V of dm

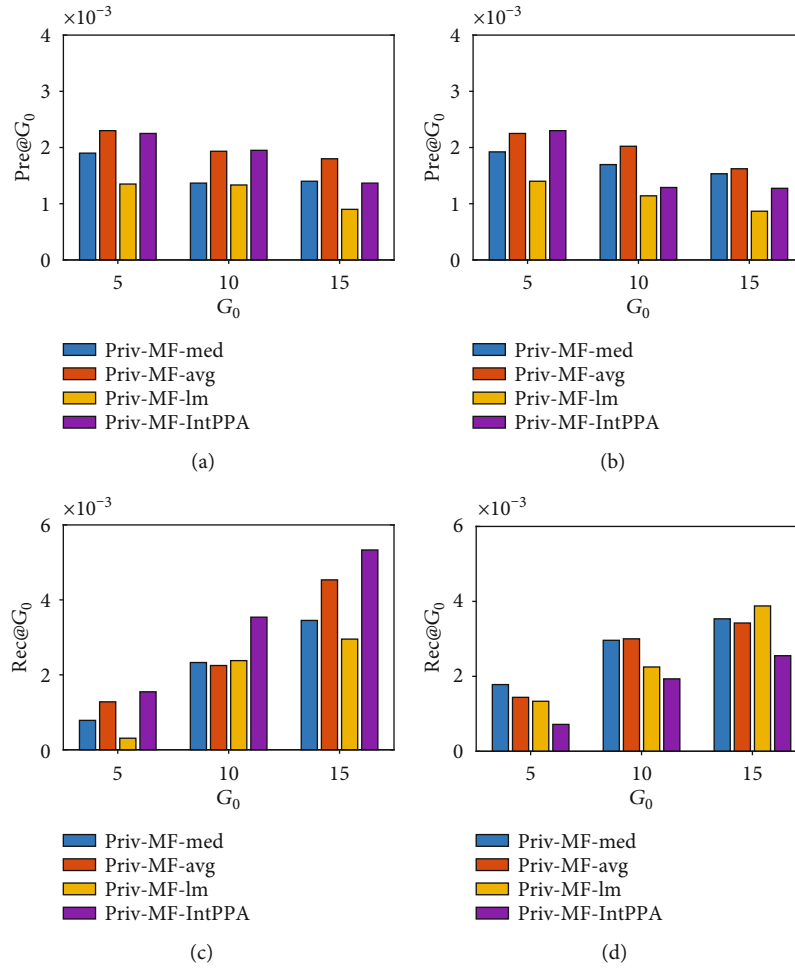


FIGURE 6: Performance of group recommendation methods for the FilmTrust dataset. (a) Group size = 5. (b) Group size = 10. (c) Group size = 5. (d) Group size = 10.

elements to each user, which are about 0.15 MB for MovieLens and 0.47 MB for FilmTrust.

In IntPPA algorithm, server has no communication cost. Users transmit elements to the server and group members, which are less than 1 MB during our algorithm in both datasets.

According to our test, under the MovieLens (FilmTrust) dataset, the time of each iteration in personalized recommendation is no more than 90 (180) seconds. The time consumed by other parts can be ignored.

6. Conclusions

In this paper, we proposed a privacy-aware group recommendation scheme, consisting of a personalized recommendation algorithm and a preference aggregation algorithm. For the personalized recommendation, we employed local differential privacy to protect user's historical data and prevent predicted preferences from leakage. We also designed an intra-group transfer privacy-preserving preference aggregation algorithm called IntPPA. IntPPA could not only protect users' privacy but also defend against long-term observation attacks. Moreover, we presented several experi-

ments to measure the privacy effect and usability of our proposed scheme. The effect and efficiency of our proposed scheme on the MovieLens and FilmTrust datasets show our advantages.

Data Availability

Our datasets are from MovieLens-1M dataset. The MovieLens-1M data supporting our recommendation system are from previously reported studies and datasets, which have been cited [28].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (61932015, 61872441, 61872100), the National Key R&D Program of China (2017YFB0802203), and the Youth Innovation Promotion Association, Chinese Academy of Sciences (2018196).

References

- [1] X. Chen, A. Li, X. Zeng, W. Guo, and G. Huang, "Runtime model based approach to iot application development," *Frontiers of Computer Science*, vol. 9, no. 4, pp. 540–553, 2015.
- [2] Y. Yang, X. Zheng, and C. Tang, "Lightweight distributed secure data management system for health internet of things," *Journal of Network and Computer Applications*, vol. 89, pp. 26–37, 2017.
- [3] S. A. Yahia, S. B. Roy, A. Chawlat, G. Das, and C. Yu, "Group recommendation: Semantics and efficiency," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 754–765, 2009.
- [4] L. A. M. C. Carvalho and H. T. Macedo, "Users' satisfaction in recommendation systems for groups: an approach based on noncooperative games," in *Proceedings of the 22nd ACM International Conference on World Wide Web*, pp. 951–958, 2013.
- [5] V. Rakesh, W. C. Lee, and C. K. Reddy, "Probabilistic group recommendation model for crowdfunding domains," in *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*, pp. 257–266, 2016.
- [6] T. D. Q. Vinh, T. A. N. Pham, G. Cong, and X. L. Li, *Attention-based group recommendation*, CoRR, 2018.
- [7] X. Wang, Y. Liu, J. Lu, F. Xiong, and G. Zhang, "TruGRC: trust-aware group recommendation with virtual coordinators," *Future Generation Computer Systems*, vol. 94, pp. 224–236, 2019.
- [8] Q. Yuan, G. Cong, and C. Y. Lin, "COM: A generative model for group recommendation," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 163–172, 2014.
- [9] D. Yang, B. Qu, and P. Cudré-Mauroux, "Privacy-preserving social media data publishing for personalized ranking-based recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 3, pp. 507–520, 2019.
- [10] C. Xu, J. Wang, L. Zhu, C. Zhang, and K. Sharif, "Ppmr: a privacy-preserving online medical service recommendation scheme in ehealthcare system," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5665–5673, 2019.
- [11] F. McSherry and I. Mironov, "Differentially private recommender systems: building privacy into the netflix prize contenders," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 627–636, 2009.
- [12] Z. Liu, Y. Wang, and A. Smola, "Fast differentially private matrix factorization," in *Proceedings of the 9th ACM Conference on Recommender Systems*, pp. 171–178, 2015.
- [13] R. Balu and T. Furon, "Differentially private matrix factorization using sketching techniques," in *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*, pp. 57–62, 2016.
- [14] J. Hua, C. Xia, and S. Zhong, "Differentially private matrix factorization," in *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 1763–1770, 2015.
- [15] Y. Shen and H. Jin, "EpicRec: towards practical differentially private framework for personalized recommendation," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 180–191, 2016.
- [16] H. Shin, S. Kim, J. Shin, and X. Xiao, "Privacy enhanced matrix factorization for recommendation with local differential privacy," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 9, pp. 1770–1782, 2018.
- [17] X. Zhou, L. Chen, Y. Zhang et al., "Enhancing online video recommendation using social user interactions," *The VLDB Journal*, vol. 26, no. 5, pp. 637–656, 2017.
- [18] J. Xu, Z. Yan, G. Cao, and J. Zhao, "Family shopping recommendation system using behavior sequence data and user profile," in *Proceedings of the 10th International Conference on Internet Multimedia Computing and Service*, pp. 1–7, 2018.
- [19] T. D. Pessemier, J. Dhondt, and L. Martens, "Hybrid group recommendations for a travel service," *Multimedia Tools and Applications*, vol. 76, no. 2, pp. 2787–2811, 2017.
- [20] C. Zhang, H. Zhang, and J. Wang, "Personalized restaurant recommendation method combining group correlations and customer preferences," *Information Sciences*, vol. 454–455, pp. 128–143, 2018.
- [21] J. S. Lee and I. Y. Ko, "Service recommendation for user groups in internet of things environments using member organization-based group similarity measures," in *Proceedings of the IEEE International Conference on Web Services*, pp. 276–283, 2016.
- [22] F. Ortega, A. Hernando, J. Bobadilla, and J.-H. Kang, "Recommending items to group of users using matrix factorization based collaborative filtering," *Information Sciences*, vol. 345, pp. 313–324, 2016.
- [23] H. Zhao, Q. Liu, Y. Ge, R. Kong, and E. Chen, "Group preference aggregation: a nash equilibrium approach," in *Proceedings of the IEEE 16th International Conference on Data Mining*, pp. 679–688, 2016.
- [24] L. Boratto, S. Carta, and G. Fenu, "Discovery and representation of the preferences of automatically detected groups: exploiting the link between group modeling and clustering," *Future Generation Computer Systems*, vol. 64, pp. 165–174, 2016.
- [25] J. Castro, J. Lu, G. Zhang, Y. Dong, and L. Martinez-López, "Opinion dynamics-based group recommender systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 12, pp. 2394–2406, 2018.
- [26] Z. Luo and Z. Chen, "A privacy preserving group recommender based on cooperative perturbation," in *Proceedings of the IEEE International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, pp. 106–111, 2014.
- [27] S. Shang, Y. Hui, P. Hui, P. Cuff, and S. Kulkarni, "Beyond personalization and anonymity: towards a group-based recommender system," in *Proceedings of the 29th Annual ACM Symposium on Applied Computing*, pp. 266–273, 2014.
- [28] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *ACM Transactions on Interactive Intelligent Systems*, vol. 5, no. 4, pp. 1–19, 2016.
- [29] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel bayesian similarity measure for recommender systems," in *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pp. 2619–2625, 2013.
- [30] F. Li, H. Li, B. Niu, and J. Chen, "Privacy computing: concept, computing framework, and future development trends," *Engineering*, vol. 5, no. 6, pp. 1179–1192, 2019.

Research Article

Secure and Intelligent Energy Data Management Scheme for Smart IoT Devices

Tianqi Zhou,^{1,2} Jian Shen ,^{2,3} Sai Ji,⁴ Yongjun Ren,¹ and Leiming Yan¹

¹The School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China

²Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University, Guiyang 550025, China

³Cyberspace Security Research Center, Peng Cheng Laboratory, Shenzhen 518000, China

⁴The Network Information Center, Nanjing University of Information Science and Technology, Nanjing 210044, China

Correspondence should be addressed to Jian Shen; s_shenjian@126.com

Received 24 June 2020; Revised 17 July 2020; Accepted 28 July 2020; Published 1 September 2020

Academic Editor: Ximeng Liu

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

The renewable energy plays an increasingly important role in many fields such as lighting, automobile, and electric power. In order to make full use of the renewable energy, various smart Internet of Thing (IoT) devices are deployed. However, in the field of energy management, the two-way mismatch between the demand and the supply of the renewable energy will greatly affect the efficiency of the renewable energy. In addition, the security threat of the energy data and the privacy leakage of the user may hinder the further development of smart IoT devices. Therefore, how to achieve consistency and balance between the demand and the renewable energy supply and how to guarantee the security and privacy of smart IoT devices become the key problems of the energy-efficient smart environment. In this paper, a secure and intelligent energy data management scheme for smart IoT devices is proposed. It is worth noting that, with the help of artificial intelligence (AI) technologies and secure cryptography primitives, the proposed scheme realizes high-efficient and secure energy utilization in a smart environment. Specifically, the proposed scheme aims at improving the efficiency of the energy utilization in the multidimensions of a smart environment. In order to realize the fine-grain energy management of smart IoT devices, strategies of three different dimensions are considered and realized in the proposed scheme. Moreover, technologies in AI are applied and integrated into the energy management scheme. The analysis shows that the proposed scheme can make full use of the renewable energy in smart IoT devices.

1. Introduction

With the development of human industrialization and the demand of all kinds of electronic intelligent products in our life, electric energy has become one of the essential resources for human beings. As a new field in power system, the smart grid provides more possibilities for power system monitoring, operation, and optimization [1].

The demand for electric energy in industry and life determines the need for a large amount of energy supply. However, today's world is experiencing an acute energy shortage. In order to make full use of the renewable energy, various smart Internet of Things (IoT) devices are deployed in a smart envi-

ronment such as smart cities, smart healthcare, and smart grid. Therefore, how to deal with the problem of energy supply will be the key problem that affects the development of the smart grid. In this paper, we first point that renewable energy [2] can be used as a source of energy supply for smart IoT devices. Then, technologies of artificial intelligence (AI) [3] are introduced into our scheme to design an AI-based energy management scheme for smart IoT devices. Finally, the effectiveness and practicability of the proposed scheme are proved by simulation.

Employing the technology of AI, the proposed scheme realizes energy management at three dimensions, which is presented as follows.

(i) Interlayer control

Energy collection base stations are classified according to its real-time output and its influencing factors (e.g., temperature, humidity, pressure, time, and season). In particular, it can be divided into three categories: energy sufficiency, energy shortage, and no energy supply. After that, the traffic load of the base station and the number of users to serve are determined by the classification.

(ii) Intralayer control

Prediction of the energy arrival rate through the analysis of the historical data. After that, smart IoT devices can adjust in advance according to the forecast results, which can effectively save energy. Moreover, the congestion can be avoided and quality of service (QoS) [4] of smart IoT devices can be improved.

(iii) Caching and pushing

Implementation of caching and push strategies can be achieved by analyzing user usage habits and history data of smart IoT devices. The recommendation algorithm realizes more accurate and effective caching and pushing. Then, their commendation data can be transmitted in advance when the energy is sufficient. It reduces the probability of congestion while saving energy and improves the user experience.

The objective of this paper is to match the demand of smart IoT devices and the energy supply of renewable energy. To this end, an AI-based energy management scheme for multidimension smart IoT devices is proposed, which leverages the technology in AI to realize a more intelligent energy-efficient smart environment.

1.1. Our Contributions. Employing the technology of AI, the proposed scheme realizes energy management at three dimensions in a smart environment. The main contributions of this paper are listed as follows.

- (i) The multidimension energy management is proposed. The energy management strategies of smart IoT devices are categorized into three dimensions. Each dimension has different classifications and control principles, thereby achieving efficient and fine-grain energy management for smart environment
- (ii) AI technologies are introduced in the multidimension energy management. AI technologies are employed in each dimension of smart IoT devices to realize energy-saving and intelligent energy management. In particular, the three dimensions of smart IoT devices are interlayer control with logistic regression and clustering analysis, intralayer control with regression algorithm and caching, and pushing with recommendation algorithm

The reminder of the article is organized as follows. An overview of energy-saving technologies in smart IoT devices is first introduced. Then, the status of renewable energy and

technologies in AI are given. Next, the proposed scheme is presented in detail and followed by the discussions of future works. Finally, the paper is summarized in the conclusion.

2. Deployment and Energy-Saving Technologies of Smart IoT Devices

The smart IoT devices can provide more convenience and service for our life. Whereas, it also enlarges the network scale and introduces highly dynamic topology, which requires more efficient and intelligent management of smart IoT devices. In this section, the deployment of smart IoT devices in a renewable energy-based smart environment is presented, which includes existing and promising future applications of smart IoT devices. In addition, existing network management technologies from all aspects are also discussed.

2.1. Deployment of Smart IoT Devices. Smart IoT devices serve various applications in every aspect of our lives and industrial manufacturing. Generally speaking, the deployment of smart IoT devices in renewable energy-based smart environment is illustrated in Figure 1. The smart environment is composed of many parts, which can be divided into smart substation, smart distribution network, smart meter, smart terminal, and new energy storage system. The intelligent substation uses advanced intelligent equipment to automatically complete the basic functions of information collection, measurement, control, protection, and monitoring.

The function of intelligent electric energy meter is the two-way metering function under the two-way interactive power supply mode. The smart terminal is the key equipment of the smart grid, which monitors and manages the electric equipment, guides the users to use the electricity reasonably, adjusts the peak and valley load of the power grid, and realizes the intelligent interaction between the power grid and the users. In addition, the following services can be supported by a smart environment.

(i) Smart home

A smart home is a kind of system, which is designed to control equipment in home. For example, audio, refrigerator, electric cooker, air conditioner, sweeping robot, and other household equipment can be remotely controlled through a network connection. The smart grid provides energy supply for devices in the smart home.

(ii) Smart cities

The concept of a smart environment was proposed in 1999 [5]. However, it has not been rapidly developed until recent years. Nowadays, the concept of smart cities [6] has been put forward. In the deployment of smart cities, technologies, and energy from all aspects, especially the energy from smart grid should be integrated in order to support smart cities.

(iii) Vehicle-to-grid (V2G) network

V2G is the relationship between electric vehicles and power grid [7]. That is, the on-board battery supplies power

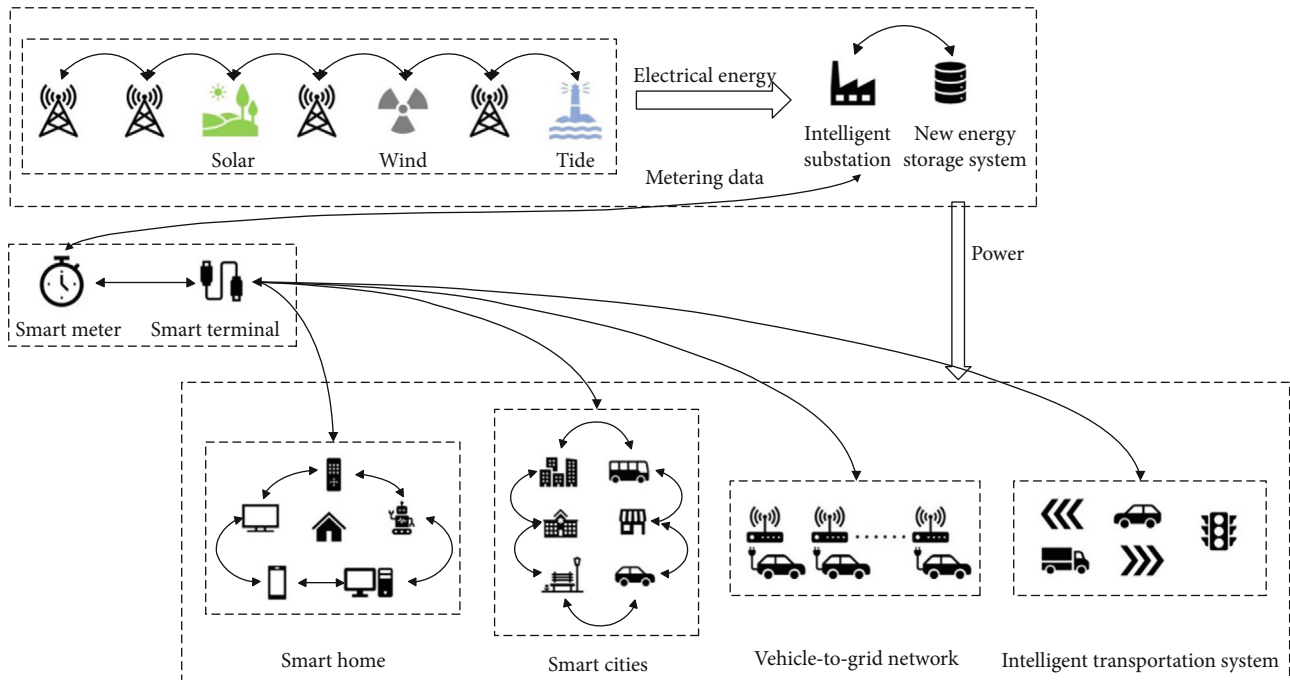


FIGURE 1: The deployment of smart IoT devices.

to the grid system when the vehicle is free. In contrast, the current flows from the grid to the vehicle when the on-board battery needs to be charged. It is obvious that the energy management of the smart grid forms the important guarantee of the V2G network.

(iv) Intelligent transportation system

The intelligent transportation system is a comprehensive transportation management system. It works an omnidirectional role across the region with real-time, accuracy, and efficiency. Unsurprisingly, the advancement of the smart grid, especially the energy management strategies of smart grid can further promote the development of an intelligent transportation system.

2.2. Related Works. Energy and cost efficiencies are ultimately important for a smart environment, which needs precise control and provides services for devices in various scenarios. However, the smart environment that integrates various renewable sources and technologies in information science will definitely become more complex [8]. Moreover, the high penetration from the renewable energy may bring adverse effects on the smart environment. Therefore, many methods and solutions have been put forward to reasonably use renewable energy in a smart environment [9–13]. The wide-area protection and control (WAPAC) [14–16] and the wide-area monitoring system (WAMS) [17–19] are two advanced concepts in the smart grid. The phasor measurement unit (PMU) [20–22] is an important component for WAPAC, which can be deployed in the smart grid to monitor various parameters of the smart grid. In order to obtain the accurate and real-time measurements, WAMS not only employs sensors nodes in the smart but also uses a global

positioning system (GPS) [23, 24]. In addition to the above structural and physical technologies, the information and communication technology (ICT) [25–28] also plays an essential role in the smart grid. The data from both suppliers and consumers are collected and analyzed by the ICT. In addition, according to the analyzed results, intelligent and efficient solutions will be proposed. Moreover, AI as an emerging technique plays an increasingly important role in various research fields. The smart grid can also apply techniques in AI to optimize the energy efficiency [29–32].

Generally, a smart environment brings all possibilities for our future life. However, various obstacles (e.g., energy consumption, security issues [28, 33], high-speed data, and high-mobility) hinge the development of it. Many technologies have been proposed to solve these difficulties, especially energy-saving technologies have been widely used in smart IoT devices. In this paper, we achieve an energy management scheme for smart IoT devices. It is worth noting that by employing the advanced technologies in AI, the proposed scheme is a novelty and performs well than existing energy saving schemes.

3. Key Issues for AI-Based Energy-Efficient Networks

The increasing scale and highly dynamic topology of smart IoT devices require energy-efficient and intelligent management. In the proposed scheme, we focus on energy-saving smart IoT devices supported by renewable energy. In addition, AI technologies are employed to achieve efficient and intelligent energy management. Therefore, in this section, the characteristics of renewable energy and technologies in AI are summarized.

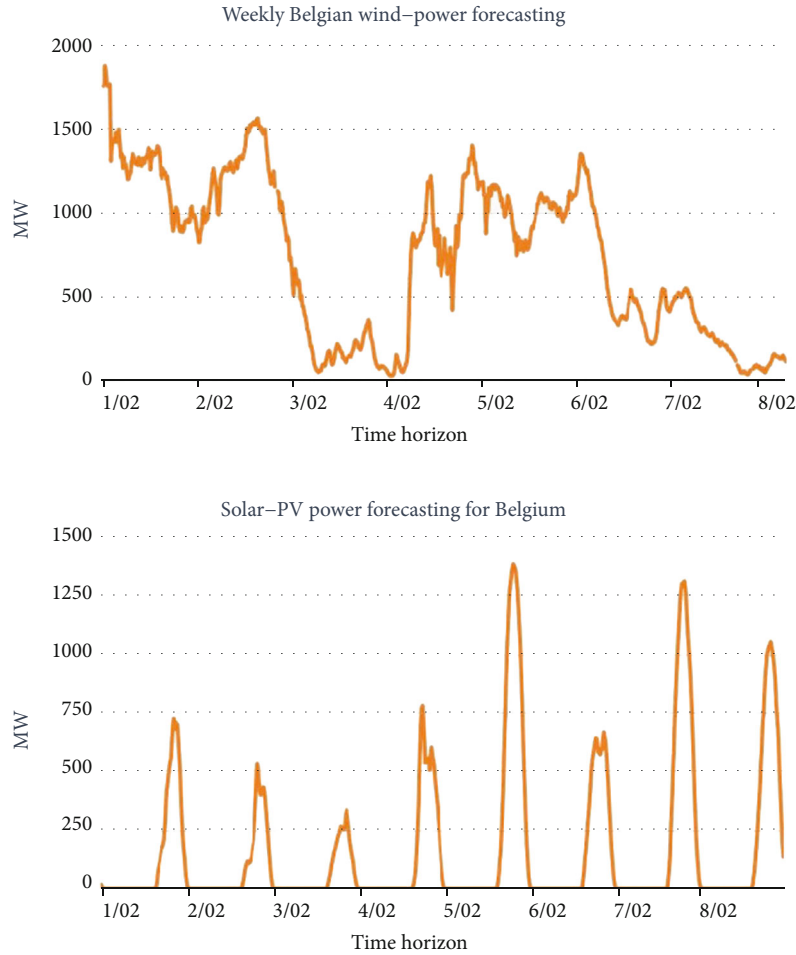


FIGURE 2: Wind-power and solar-PV power.

3.1. Renewable Energy. Renewable energy, including sunlight, wind, rain, tides, waves, and geothermal heat, is theoretically inexhaustible and will become progressively more important as time goes on. Today, renewable energy has become an energy source in the fields of electricity, air and water heating/cooling, transportation, and rural (off-grid) energy services [34]. In 2015, more than US\$286 billion was invested in renewable technologies about wind, hydro, solar, and bio-fuels. In the same year, according to statistics, renewable energy contributed 23.7% to humans' generation of electricity. By 2015, more than half of the world's new power capacity installation is renewable.

Elia [35], a Belgium's transmission system operator, can generate power either in "traditional" ways (e.g., in nuclear power stations, combined-cycle gas turbine facilities, or combined heat and power plants) or in units using renewable energy sources (like wind or solar farms and thermal or hydroelectric power stations). Then, the created power is injected into the (high-voltage) transmission system. Finally, it is delivered to the end user.

Specifically, according to the measured data of Elia, the energy generated by wind and solar in the first week of February 2018 is illustrated in Figure 2. In Figure 2, megawatts, usually abbreviated as MW, is a unit of power. It refers to

the electricity generated by generators in the unit time under rated conditions.

It can be observed from Figure 2 that the wind power reaches a maximum of about 1800 MW on the early morning of February 1 and then drops down. After a period of fluctuation, it hits the bottom on the 4th of February. In the following days, the wind power goes through volatility increase and decrease, eventually, reaches trough again at dusk on the 7th of February. By contrast, solar power is much more regular, which can be seen in Figure 2. With sunrise and sunset, the output of solar energy changes regularly. In particular, during the day, the solar power begins to climb up from the morning and reach the highest point around mid-day. Then, it begins to fall and finally shuts down at evening. Although the wind energy supply is not very regular and there is no supply of solar energy at night, they can also contribute a considerable energy supply to the power grid. Therefore, the development and utilization of renewable energy will be a promising application field, especially for the smart grid, which requires more energy consumption. In summary, it is indicated in Figure 2 that on the one hand, renewable energy provides considerable power for the smart grid. On the other hand, it implicates the mismatch between the grid load demand and renewable energy supply.

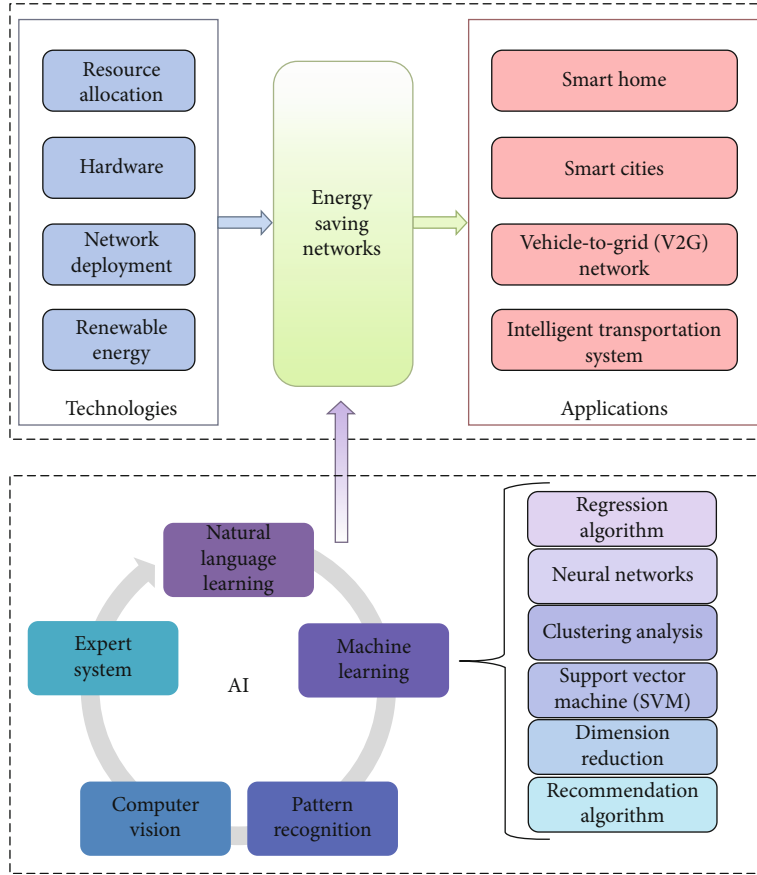


FIGURE 3: The architecture of the future networks and artificial intelligence.

3.2. Artificial Intelligence Technology. Due to the mismatch between the grid load and the supply of the renewable energy, many researchers have been devoted to the improvement of the renewable energy utilization. In the following, technologies in AI are introduced. These technologies can be employed in smart IoT devices to improve the utilization of the renewable energy. Artificial intelligence, a new science simulates cognitive functions of humans, was founded as an academic discipline in 1956. The intelligent automation of electrical distribution networks has driven the future power system development, which will bring change in both network design and network operation. In the future, AI will bring change in both network design and network operation. As shown in Figure 3, artificial intelligence mainly includes the branches of machine learning, computer vision, pattern recognition, expert system, and natural language learning.

In this paper, we mainly employ technologies of machine learning in the proposed AI-based energy management scheme, which are displayed in Figure 3. In the following, technologies of machine learning that will be used in this paper are introduced.

3.2.1. Regression Algorithm. In the energy management of smart IoT devices, the regression algorithm can be employed to derive the mathematical model of the energy. Thus, achieving more efficient energy management in smart environment. The regression algorithms include linear regression

[36–38] and logistic regression [39]. Linear regression models the relationship between a dependent variable and one or more independent variables. Linear regression can be fitted using the least-squares approach which has many practical uses like prediction. By contrast, logistic regression is a regression model where the dependent variable is categorical, which has been widely used on the field like machine learning, medical, and social sciences. In general, the basic formulas of the regression algorithm are

$$\theta_j = \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x_i) - y_i) x_i^j, \quad (1)$$

and the sigmoid function

$$\sigma(z) = \frac{1}{1 + e^{-z}}. \quad (2)$$

Even though the regression algorithm can be understood and implemented easily, neural network can deal with more complicated and nonlinear cases than the regression algorithm. In some complicated cases of smart IoT devices, neural network can perform well. In the following, the basis of the neural network is presented.

3.2.2. Neural Networks. Artificial neural network (ANN) [40] is a nonlinear and adaptive information processing system

consisting of a large number of interconnected processing units. Without any prior knowledge, ANN is able to evolve a set of relevant knowledge system from the learning material that it processed. ANN has been widely applied to various fields such as machine translation, computer vision, speech recognition, and medical diagnosis. In general, the basic formulas of the neural networks are

$$\tanh(x) = \frac{e^{(x)} - e^{(-x)}}{e^{(x)} + e^{(-x)}}, \quad (3)$$

$$\tanh(x)' = 1 - (\tanh(x))^2. \quad (4)$$

The scores of each layer in the neural network are calculated as

$$\text{score} = \sum_{i=0}^{d^{l-1}} w_{ij}^l x_i^{l-1}. \quad (5)$$

In particular, the formula of the gradient descent method in a neural network is as follows.

$$\begin{aligned} e_n &= (y_n - \text{NNet}(\mathbf{x}_n))^2 = (y_n - s_1^{(L)})^2 \\ &= \left(y_n - \sum_{i=0}^{d^{(L-1)}} w_{i1}^{(L)} x_i^{(L-1)} \right)^2. \end{aligned} \quad (6)$$

Then, the partial derivative result of output layer is calculated as

$$\frac{\partial e_n}{\partial w_{i1}^{(L)}} = \frac{\partial e_n}{\partial s_1^{(L)}} \cdot \frac{\partial s_1^{(L)}}{\partial w_{i1}^{(L)}} = -2(y_n - s_1^{(L)}) \cdot (x_i^{(L-1)}). \quad (7)$$

Finally, the optimization can be achieved based on the following formula

$$\begin{aligned} E_{\text{in}}(w) &= \frac{1}{N} \sum_{n=1}^N \text{err} \left(\left(\dots \tanh \left(\sum_i w_{ik}^{(2)} \right. \right. \right. \\ &\quad \left. \left. \left. \cdot \tanh \left(\sum_i w_{ij}^{(1)} x_{n,i} \right) \right) \right), y_n \right). \end{aligned} \quad (8)$$

For some data, a linear decision boundary cannot be found. In this case, SVM can be used to achieve classification. In particular, in the concept of SVM, the third dimension is introduced into the two-dimensional plane to achieve the classification of nonlinear cases.

3.2.3. Support Vector Machine (SVM). In machine learning, support vector machines (SVM) [41] are supervised learning models with associated learning algorithms which are mainly used to solve the problem of data classification in pattern recognition. Given a set of training examples, an SVM training algorithm builds a model that can classify new examples. Note that SVM can not only perform linear classification, it can also efficiently perform a nonlinear classification by

employing a kernel trick, which maps the inputs into high-dimensional feature spaces to achieve a nonlinear classification. Various real-world problems including text and hyper-text categorization, classification of images, and biological can resort to SVM. In general, the basic formulas of the SVM are

$$\begin{cases} \omega^T x_i + \gamma \geq 1 & \forall y_i = 1, \\ \omega^T x_i + \gamma \leq -1 & \forall y_i = -1. \end{cases} \quad (9)$$

The above formula is equivalent to

$$y_i(\omega^T x_i + \gamma) \geq 1 \quad \forall x_i. \quad (10)$$

3.2.4. Clustering Analysis. Clustering is a kind of unsupervised learning [42], and the purpose is to classify a set of data points. Clustering analysis is distinguished from the supervised classification analysis, where no training data are available. Specifically, the clustering algorithm is a method of automatically dividing a pile of unlabeled data into several classes, which ensures similar features of the same class of data, in which the data of the same class have similar features. Clustering analysis plays an important role in pattern recognition, data compression, computer graphics, etc. Generally speaking, clustering learning algorithms include K -means, Agglomerative, and DBSCAN. The key part of the clustering algorithm is the calculation of the distance. For example, in K -means algorithm, the distance is usually calculated as

$$\text{dist}(a, b) = \sum_i^N (|a_i - b_i|^p)^{1/p}. \quad (11)$$

3.2.5. Recommendation Algorithm. The recommendation algorithm [43] is an information filtering algorithm that seeks to predict the preference of users. It mainly includes demographic-based recommendation, content-based recommendation, and collaborative filtering. In general, the basic formulas of the recommendation algorithm are

$$\begin{aligned} \rho_{x,x} &= \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y} = \frac{E((X - \mu_x)(Y - \mu_y))}{\sigma_x \sigma_y} \\ &= \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}}. \end{aligned} \quad (12)$$

The above formula is the Pearson correlation coefficient, which is equivalent to

$$\rho_{x,y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}. \quad (13)$$

Also, two commonly used distance formulas are Euclidean distance and cosine distance, which are shown as follows.

$$\text{similarity}(X, Y) = 1 / \left(1 + \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \right), \quad (14)$$

$$\text{sim}(X, Y) = \cos \theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \cdot \|\vec{y}\|}. \quad (15)$$

4. AI-Based Energy Management Scheme for the Multidimension Smart IoT Devices

The proposed AI-based energy management scheme can be categorized into three dimensions. In this section, each dimension of the scheme combined with AI technologies is presented in detail.

4.1. Interlayer with Logistic Regression and Clustering Analysis. The main goal of interlayer control is to classify the energy collection base stations so that they can match the demand of smart IoT devices. Specifically, the classification can be achieved by employing logistic regression and clustering analysis.

4.1.1. Interlayer with Logistic Regression. Using logistic regression analysis in the interlayer, the energy collection base station can be divided into available and unavailable. In fact, logistic regression is constructed by embedding the sigmoid function [44] into linear regression, which converts the numerical results to a probability between 0 and 1. After that, predictions and classifications can be derived based on this probability and the threshold set in advance.

Figures 4 and 5 demonstrate the classification in the interlayer, which is realized with the help of the logistic regression. Figure 4 shows the relation between the influence factors of renewable energy and energy output, which are the original training data. Based on the original training data, a logistic regression model can be trained.

4.1.2. Inter layer with Clustering Analysis. In addition to logistic regression, clustering analysis can also be applied in interlayer energy management for smart IoT devices. Compared with logistic regression, clustering analysis is unsupervised learning, which is able to classify unlabeled data automatically. In the proposed scheme, the energy collection base stations are classified according to the time and meteorological data.

Figures 6 and 7 depict the clustering analysis of the energy collection base stations classification, which can be used for guiding the smart IoT device control. It is worth noting that no training data are available in clustering analysis. Figure 6 shows the original data based on time and meteorological data. Then, the original data can be automatically divided into four classes by the clustering algorithm, which is shown in Figure 7. Specifically, the purple, red, green, and blue points in Figure 7 represent four kinds of the energy collection base stations, which are sufficient, barely enough, low, and no energy supply, respectively. After that, the load of smart IoT devices can be assigned according to the classification.

From the above discussion, it is not difficult to observe that the interlayer control of the proposed scheme can be achieved with the help of the regression and the clustering algorithms.

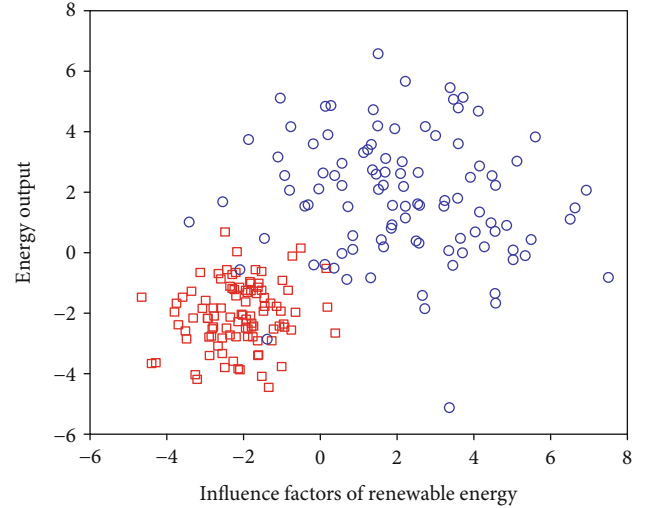


FIGURE 4: Original training data.

4.2. Intralayer Control with Regression Algorithm. The main purpose of intralevel is to determine the distribution of demands and the number of users served by predicting the supply of energy. Therefore, this requires a more accurate prediction of the supply of energy. Generally speaking, the linear function fitting has already been realized by the least square method. However, from the analysis of Section 3, we can know that the supply of renewable energy is nonlinear. Fortunately, neural networks and SVM can fit nonlinear data. In general, a cost function needs to be constructed first. Then, gradient descent algorithm is applied to approach the lowest point (i.e., the minimum point of the cost function). Finally, the optimal solution of the weight is obtained, which is also the optimal solution of fitting.

In particular, a visualized example of the cost function is illustrated in Figure 8, which shows the cost function with two parameters. In Figure 8, the X and Y-axis represent parameters θ_1 and θ_2 , respectively. Namely, the factors that determine the output of renewable energy. Accordingly, the Z-axis is the cost $J(\theta_1, \theta_2)$, namely, the deviation between the fitting results and the practical data. Therefore, the minimum point of the cost function is the smallest error and is also the optimal solution of the fitting. Moreover, finding the minimum point of the cost function can be achieved by the widely-used gradient descent algorithm.

It can be seen that the output of renewable energy can be predicted by regression analysis. It is helpful for the energy-efficient smart IoT devices.

4.3. Caching and Pushing with Recommendation Algorithm. Caching and pushing in the energy-efficient smart IoT devices aims at taking advantage of the superfluous resource. In the energy-efficient smart IoT devices with renewable energy, the mismatch between the demand and energy supply is an obstacle to making full use of renewable resources. For example, solar energy reaches the maximum at noon and is unable to provide energy at night. However, the smart IoT devices generally have a large demand for power in the evening. Thus, the supply of renewable energy in a direct

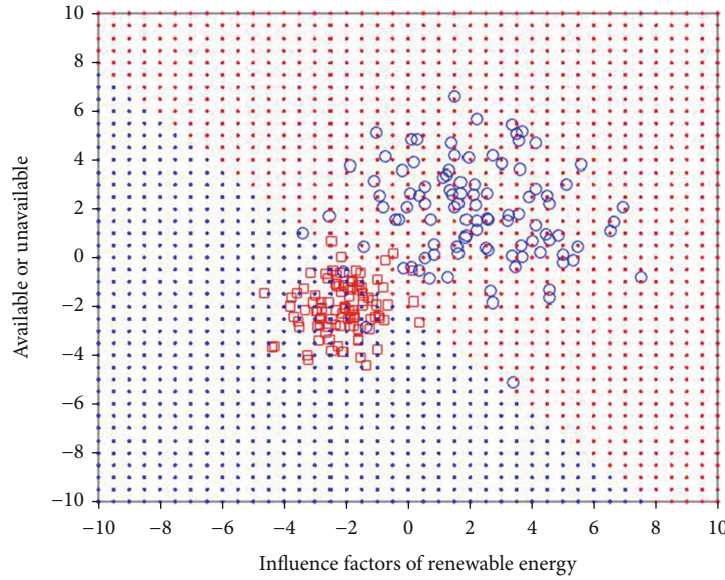


FIGURE 5: Classification results on training data.

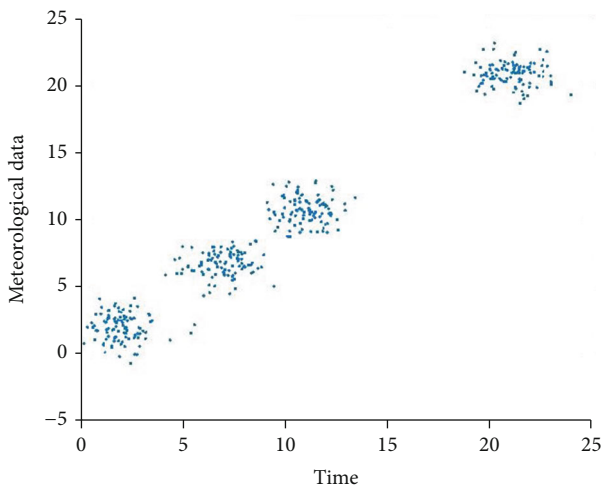


FIGURE 6: Original data.

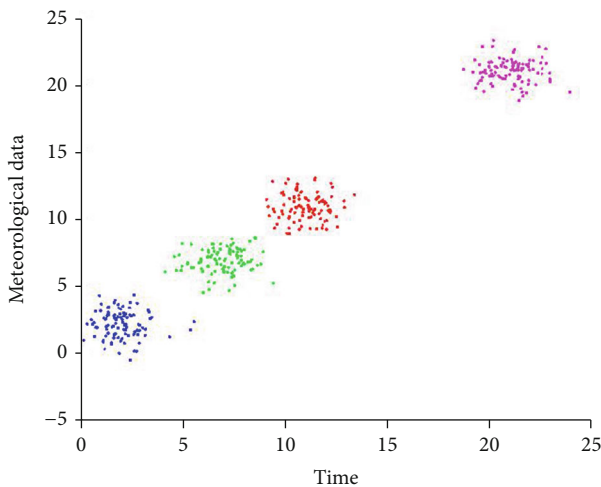


FIGURE 7: Clustering results.

way will cause a waste of energy. The problem of mismatch between energy supply and demand is alleviated on time scale by caching technology. Essentially, the demand that is needed in the evening can be cached in advance when the resource is sufficient. Moreover, in order to accurate and effective caching, a recommendation algorithm could be added. By analyzing users' information, using habits, and historical data, the recommended algorithms can implement more efficient caching and pushing for smart IoT devices.

In the proposed scheme, two kinds of collaborative filtering algorithms: user-based collaborative filtering and item-based collaborative filtering are employed. Note that the base idea of user-based collaborative filtering and item-based collaborative filtering has been changed in accordance with the context.

5. Future Research Directions

There are many challenges to be solved in the energy-efficient smart IoT devices, which are briefly discussed in the following.

(i) Computation overhead and storage overhead

For one thing, the introduction of AI technology will lead to greater computing overhead. For another, the caching strategy will also provide a higher requirement for storage resources. Therefore, in order to diminish the computing overhead and storage overhead, new technologies need to be considered. For example, cloud computing and cloud storage [45].

(ii) Reliability

In the network, it is required to provide services at any time. However, the supply of renewable energy is discontinuous and fluctuant. The smooth transition from energy

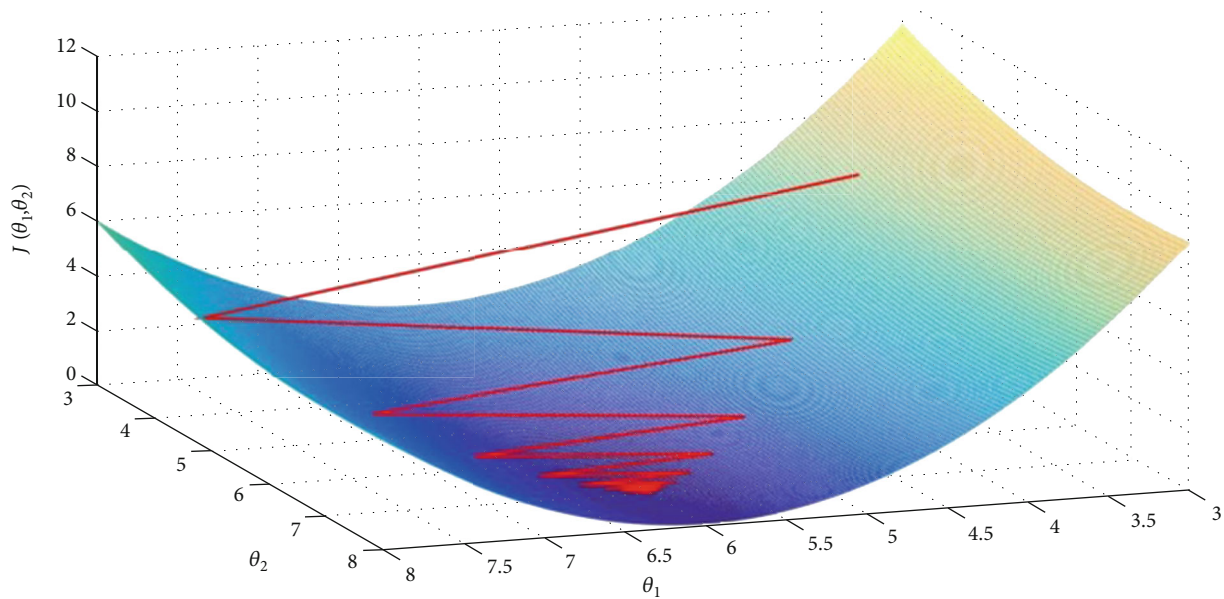


FIGURE 8: Figure of the cost function.

supply to smart IoT devices still requires more researches and works to be devoted.

(iii) Quality of Service (QoS)

User-friendly and environmentally-friendly guarantee the sustainable development of the energy-efficient smart IoT devices. Therefore, QoS should be considered in the design of an AI-based energy management scheme.

(iv) Security

Nowadays, with the development of information technology, all kinds of security threats are also flooded with various fields touched by the network. The smart IoT devices are no exception. For example, the user of smart IoT devices may be performed by an attacker. In order to ensure the normal running of smart IoT devices, the security mechanism needs to be added to the energy management of smart IoT devices.

6. Conclusion

In this paper, we concentrate on improving the energy efficiency of smart IoT devices. The existing schemes that improve the efficiency of energy use in smart IoT devices are analyzed. Moreover, the characteristics and current status of renewable energy and technologies in AI were studied and introduced. After that, the energy management scheme of three dimensions was proposed. Remarkably, in order to obtain a satisfactory result, technologies in AI were embedded in the three dimensions. Namely, interlayer control with logistic regression and clustering analysis, intralayer control with regression algorithm, and caching and pushing with recommendation algorithm. The analysis shows that the energy management scheme combined with AI technologies can greatly improve the utilization of renewable energy in a

fine-grain scale. Finally, some open issues are discussed in the future works.

Data Availability

The related data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grants No. U1836115, No. 61672295, No. 61922045, No. 61877034, and No. 61962009, the Natural Science Foundation of Jiangsu Province under Grant No. BK20181408, the Major Scientific and Technological Special Project of Guizhou Province under Grant No. 20183001, and the Foundation of Guizhou Provincial Key Laboratory of Public Big Data under Grant No. 2018BDFJ003.

References

- [1] T. Dragičević, P. Siano, and S. R. Prabaharan, "Future generation 5g wireless networks for smart grid: a comprehensive review," *Energies*, vol. 12, no. 11, p. 2140, 2019.
- [2] A. Q. Huang, "Power semiconductor devices for smart grid and renewable energy systems," in *Power Electronics in Renewable Energy Systems and Smart Grid: Technology and Applications*, pp. 85–152, 2019.
- [3] F. Jiang, Y. Jiang, H. Zhi et al., "Artificial intelligence in healthcare: past, present and future," *Stroke and Vascular Neurology*, vol. 2, no. 4, pp. 230–243, 2017.

- [4] M. W. Khan and M. Zeeshan, "QoS-based dynamic channel selection algorithm for cognitive radio based smart grid communication network," *Ad Hoc Networks*, vol. 87, pp. 61–75, 2019.
- [5] M. Weiser, R. Gold, and J. S. Brown, "The origins of ubiquitous computing research at parc in the late 1980s," *IBM system-journal*, vol. 38, no. 4, pp. 693–696, 1999.
- [6] T. Zhou, J. Shen, D. He, P. Vijayakumar, and N. Kumar, "Human-in-the-loop-aided privacy-preserving scheme for smart healthcare," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–10, 2020.
- [7] J. Shen, T. Zhou, F. Wei, X. Sun, and Y. Xiang, "Privacy-preserving and lightweight key agreement protocol for v2g in the Social Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2526–2536, 2018.
- [8] M. M. Eissa, "Challenges and novel solution for wide-area protection due to renewable sources integration into smart grid: an extensive review," *IET Renewable Power Generation*, vol. 12, no. 16, pp. 1843–1853, 2018.
- [9] M. H. Rehmani, M. Reisslein, A. Rachedi, M. Erol-Kantarci, and M. Radenkovic, "Integrating renewable energy resources into the smart grid: recent developments in information and communication technologies," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 2814–2825, 2018.
- [10] M. Hossain, N. Madlool, N. Rahim, J. Selvaraj, A. Pandey, and A. F. Khan, "Role of smart grid in renewable energy: An overview," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1168–1184, 2016.
- [11] H. Lund, "Renewable heating strategies and their consequences for storage and grid infrastructures comparing a smart grid to a smart energy systems approach," *Energy*, vol. 151, pp. 94–102, 2018.
- [12] A. M. Eltamaly, M. A. Mohamed, and A. I. Alolah, "A novel smart grid theory for optimal sizing of hybrid renewable energy systems," *Solar Energy*, vol. 124, pp. 26–38, 2016.
- [13] J. Shen, C. Wang, T. Li, X. Chen, X. Huang, and Z.-H. Zhan, "Secure data uploading scheme for a smart home system," *Information Sciences*, vol. 453, pp. 186–197, 2018.
- [14] S. Blair, G. Burt, N. Gordon, and P. Orr, "Wide area protection and fault location: review and evaluation of pmu-based methods," *The 14th International Conference on Developments in Power System Protection (DPSP)*, 2018.
- [15] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1389–1407, 2017.
- [16] V. K. Singh, A. Ozen, and M. Govindarasu, "Stealthy cyber attacks and impact analysis on wide-area protection of smart-grid," *2016 North American Power Symposium (NAPS)*, pp. 1–6, 2016.
- [17] Y. Liu, S. You, W. Yao et al., "A distribution level wide area monitoring system for the electric power grid-FNET/GridEye," *IEEE Access*, vol. 5, pp. 2329–2338, 2017.
- [18] W. Yao, D. Zhou, L. Zhan et al., "GPS signal loss in the wide area monitoring system: Prevalence, impact, and solution," *Electric Power Systems Research*, vol. 147, pp. 254–262, 2017.
- [19] D. Bhor, K. Angappan, and K. M. Sivalingam, "Network and power-grid co-simulation framework for smart grid wide-area monitoring networks," *Journal of Network and Computer Applications*, vol. 59, pp. 274–284, 2016.
- [20] S. R. Samantaray, I. Kamwa, and G. Joos, "Phasor measurement unit based wide-area monitoring and information sharing between micro-grids," *IET Generation, Transmission & Distribution*, vol. 11, no. 5, pp. 1293–1302, 2017.
- [21] W. Meng, X. Wang, Z. Wang, and I. Kamwa, "Impact of causality on performance of phasor measurement unit algorithms," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1555–1565, 2017.
- [22] H. H. Müller and C. A. Castro, "Genetic algorithm-based phasor measurement unit placement method considering observability and security criteria," *IET Generation, Transmission & Distribution*, vol. 10, no. 1, pp. 270–280, 2016.
- [23] R. N. Gore and S. P. Valsan, "Wireless communication technologies for smart grid (wams) deployment," *2018 IEEE International Conference on Industrial Technology (ICIT)*, pp. 1326–1331, 2018.
- [24] T. J. Mary, K. Ragavendran, S. R. Raja, and Z. M. Saqib, "Power system advancement with utility customer interface system," *2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)*, pp. 591–594, 2017.
- [25] A. Kovendan and D. Sridharan, "Development of smart grid-system in India: a survey," *Proceedings of the International Conference on Nano-electronics, Circuits & Communication Systems*, pp. 275–285, 2017.
- [26] J. Shen, T. Zhou, D. He, Y. Zhang, X. Sun, and Y. Xiang, "Block design-based key agreement for group data sharing in Cloud computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 6, pp. 996–1010, 2019.
- [27] J. Shen, T. Zhou, F. Wei, X. Sun, and Y. Xiang, "Privacy-preserving and lightweight key agreement protocol for v2g in the social internet of things," *IEEE Internet of things Journal*, vol. 5, no. 4, pp. 2526–2536, 2017.
- [28] T. Zhou, J. Shen, X. Li, C. Wang, and H. Tan, "Logarithmic encryption scheme for cyber-physical systems employing Fibonacci Q-matrix," *Future Generation Computer Systems*, vol. 108, pp. 1307–1313, 2020.
- [29] S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings, "Putting the 'smarts' into the smart grid," *Communications of the ACM*, vol. 55, no. 4, pp. 86–97, 2012.
- [30] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, 2015.
- [31] E. S. Rigas, S. D. Ramchurn, and N. Bassiliades, "Managing electric vehicles in the smart grid using artificial intelligence: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1619–1635, 2014.
- [32] J. Shen, T. Zhou, K. Wang, X. Peng, and L. Pan, "Artificial intelligence inspired multi-dimensional traffic control for heterogeneous networks," *IEEE Network*, vol. 32, no. 6, pp. 84–91, 2018.
- [33] L. Liu, O. DeVel, Q.-L. Han, J. Zhang, and Y. Xiang, "Detecting and preventing cyber insider threats: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1397–1417, 2018.
- [34] F. C. Kilic, "Recent renewable energy developments, studies, incentives in Turkey," *Energy Education Science and Technology Part A: Energy Science and Research*, vol. 28, no. 1, pp. 37–54, 2011.
- [35] Elia, Belgium's electricity transmission system operator, <http://www.elia.be/en/about-elia>.

- [36] S. A. Shah, U. Nadeem, M. Bennamoun, F. Sohel, and R. Togneri, "Efficient image set classification using linear regression based image reconstruction," *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 99–108, 2017.
- [37] M. L. King, "Testing for autocorrelation in linear regression-models: a survey," *Specification analysis in the linear model*, pp. 19–73, 2018.
- [38] M. Schleich, D. Olteanu, and R. Ciucanu, "Learning linear regression models over factorized joins," *Proceedings of the 2016 International Conference on Management of Data*, pp. 3–18, 2016.
- [39] W. Jiang, J. Josse, M. Lavielle, and T. Group, "Logistic regression with missing covariates parameter estimation, model selection and prediction within a joint-modeling framework," *Computational Statistics & Data Analysis*, vol. 145, article 106907, 2020.
- [40] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science*, vol. 355, no. 6325, pp. 602–606, 2017.
- [41] T. Han, D. Jiang, Q. Zhao, L. Wang, and K. Yin, "Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 8, pp. 2681–2693, 2017.
- [42] J. Blackhurst, M. J. Rungtusanatham, K. Scheibe, and S. Ambulkar, "Supply chain vulnerability assessment: A network based visualization and clustering analysis approach," *Journal of Purchasing and Supply Management*, vol. 24, no. 1, pp. 21–30, 2018.
- [43] L. Xiaojun, "An improved clustering-based collaborative filtering recommendation algorithm," *Cluster Computing*, vol. 20, no. 2, pp. 1281–1288, 2017.
- [44] A. Kanakia, J. Klingner, and N. Correll, "A response threshold sigmoid function model for swarm robot collaboration," *Distributed Autonomous Robotic Systems*, pp. 193–206, 2016.
- [45] X. Chen, X. Huang, J. Li, J. Ma, W. Lou, and D. S. Wong, "New algorithms for secure outsourcing of large-scale systems of linear equations," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 1, pp. 69–78, 2014.

Research Article

A Data Encryption and Fast Transmission Algorithm Based on Surveillance Video

Shi Qiu ¹, Ying Cui ², and XianJia Meng ³

¹Key Laboratory of Spectral Imaging Technology CAS, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, 710119, China

²College of Equipment Management and Support, Engineering University of PAP, Xi'an, 710086, China

³School of Information Science and Technology, Northwest University, Xi'an, 710027, China

Correspondence should be addressed to Ying Cui; cuiying_1224@163.com and XianJia Meng; xianjiam@nwu.edu.cn

Received 3 June 2020; Revised 24 June 2020; Accepted 21 July 2020; Published 5 August 2020

Academic Editor: Ximeng Liu

Copyright © 2020 Shi Qiu 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.

Video surveillance is an effective way to record current events. In view of the difficulty of efficient transmission of massive surveillance video and the risk of leakage in the transmission process, a new data encryption and fast transmission algorithm is proposed in this paper. From the perspective of events, the constraints of time and space dimension is broken. First, a background and moving object extraction model is built based on video composition. Then, a strong correlation data encryption and fast transmission model is constructed to achieve efficient data compression. Finally, a data mapping mechanism is established to realize the decoding of surveillance video. Our experimental results show that the compression ratio of the proposed algorithm is more than 60% under the premise of image confidentiality.

1. Introduction

Video surveillance system has a wide range of application value in many fields such as security defense, traffic management, and environmental detection. The massive surveillance video data encryption and fast transmission is the current problem to be solved [1, 2]. It is reported that the video surveillance data is limited by the size of storage space, and only the video data within a certain time range is saved (generally, the video data stored in public places for 1 to 2 months, such as shopping malls or corridors, and the video data stored in special places for 3 to 6 months, such as gas stations and banks).

From the perspective of security, the surveillance information needs to be retained for forensics and security screening as long as possible. In terms of video data encryption, Xiao et al. [3] design an encryption algorithm from the perspective of hardware. Li et al. [4] design an optional encryption protection mode. Aljawarneh and Yassein [5] use a threshold to encrypt video based on the big video data. Xu [6] considers the data confidentiality and compression

together. Khelif et al. [7] evaluate the video encryption effect. In order to realize the efficient storage of surveillance video, scholars put forward the theory of compressed sensing. The basic idea is to reduce the dimension of video data by sampling the signal at the rate of under Nyquist and recover the signal by using the prior knowledge of the signal. Main algorithms include Chen et al. [8] propose a distributed compression sensing algorithm to balance the weight of decoding and encoding. Canh and Jeon [9] propose the Kronecker model to alleviate the complexity of high dimension measurement. Adler et al. [10] block the image and compresses the image in different regions. Xu and Ren [11] construct a multiscale compression framework to achieve dynamic compression. Huang et al. [12] use sliding windows to find similar areas for compression. Zhong et al. [13] use the deep learning to select image data blocks for compression. Biswas et al. [14] propose a SIFT model to describe the change of temporal correlation of video sequence to achieve compression. Zheng et al. [15] use the sparse coding to compress data. Fei et al. [16] compress the multiview image fusion. Rahaman and Paul [17] use different coding methods to compress data

according to its importance. Chaudhari and Dhok [18] transform video into frequency domain analysis and coding. Abbas et al. [19] use optical code division multiple-access networks. Liu et al. [20] propose a Cloud computing data security algorithm. Yu et al. [21] propose a novel three-layer QR code based on the secret sharing scheme and liner code.

All of the above algorithms are based on the image frame compression algorithm, and the compression efficiency is limited under the premise of ensuring the video effect.

The surveillance video is a kind of video which is formed by shooting fixed scenes with fixed cameras. The data has a strong correlation in time and space dimensions. The content of surveillance video can be considered as a dynamic superposition of moving object and static background. If the moving object and static background can be saved, it has a strong guiding role in compression. At present, the main algorithms for background establishment and moving object extraction are as follows: Li et al. [22] use the difference between the foreground frame and the background frame to extract the moving object. He et al. [23] build an optical flow model based on the motion information. Sengar and Mukhopadhyay [24] introduce the target boundary extraction mechanism based on the optical flow model to extract the moving object more completely. Ou et al. [25] build a GMM model and introduces learning factors to dynamically update the foreground and background. Chavan and Gengaje [26] combine a GMM model with an optical flow method to extract the target hierarchically. Yeh and Lin [27] establish a three-layer discrimination mechanism to locate the moving target area in real-time. Shijila et al. [28] regard video as a low-rank matrix, and then extracts moving objects and background regions. However, the above algorithm only considers the characteristics of the image itself, so it is difficult to build a pure background. When the moving object is extracted, the situation of moving tailing and submerged in the background will appear.

Therefore, we carry out in-depth research on surveillance video data, analyze it from a new perspective of the minimum unit of object, establish a time and space constraint model to compress image data substantially, and establish the encryption mapping relationship between the compression and the original video to realize the safe and fast transmission of surveillance data.

2. Data Encryption and Fast Transmission

Video data is composed of limited frame image data F and attribute W :

$$V = F + W. \quad (1)$$

Surveillance video data is collected by a fixed camera, and the visual surveillance data is composed of the pure background image B and the moving object D .

$$F = B + D. \quad (2)$$

According to the strong correlation between image frames, the moving object D has

$$\begin{cases} D_n = D_c + D_n^u, \\ D_{n+1} = D_c + D_{n+1}^u, \end{cases} \quad (3)$$

where D_n and D_{n+1} are two adjacent images, D_c is the same part, D_n^u and D_{n+1}^u are the specific part of D_n and D_{n+1} , and D_c is saved, D_n^u and D_{n+1}^u are encoded to achieve the compression.

The proposed algorithm uses moving targets in multi-frames and the background in the single frame. The great change and the background light intensity mutation of surveillance video do not occur in a short time, so the gradual change is ignored in this paper. Finally, the background with gradual changes is not the main information.

Based on the above analysis and derivation, the key steps of surveillance video information compression are to establish a pure background image B , accurately calculate the moving object D , and quickly encode and decode the specific part. The flow chart designed is shown in Figure 1: (1) The frame difference model of visual perception is constructed to realize the establishment of pure background and the fast extraction of the motion region. (2) The compression mechanism of intraframe and interframe is established to break the constraints of the global time axis and spatial axis and realize the high compression of data in the space-time dimension. (3) Establish the corresponding relationship between the compressed video and the original video and quickly reconstruct the original video.

2.1. Video Information Extraction. According to the cognitive principle, the moving object and background reconstruction are extracted in the moving area after the visual perception, and the background is solidified. Regardless of the subsequent changes, the moving object will not be regarded as the background due to long-term static. Through this principle, the video sequence can be regarded as a completely static background (pure background) and a moving target.

The main idea of frame difference algorithm is to select two images to calculate the absolute value of pixel value difference point by point, measure the difference with a specific threshold T , and get the difference image $d_n(x, y)$ to determine the difference area.

$$d_n(x, y) = \begin{cases} 1 & f_n(x, y) - b_n(x, y) \geq T, \\ 0 & f_n(x, y) - b_n(x, y) < T. \end{cases} \quad (4)$$

In recent years, scholars have made a series of improvements on the frame difference method. Shang et al. [29] propose a three-frame difference for background detection. Zaharin et al. [30] established a background subtraction and frame difference model for pedestrian detection. Guo et al. [31] dynamically update the background frame and extract the moving object in real-time. The research of the above algorithm mainly focuses on the long-time motion of the moving object, without considering the long-time static situation of the object, resulting in the object will be submerged in the background frame.

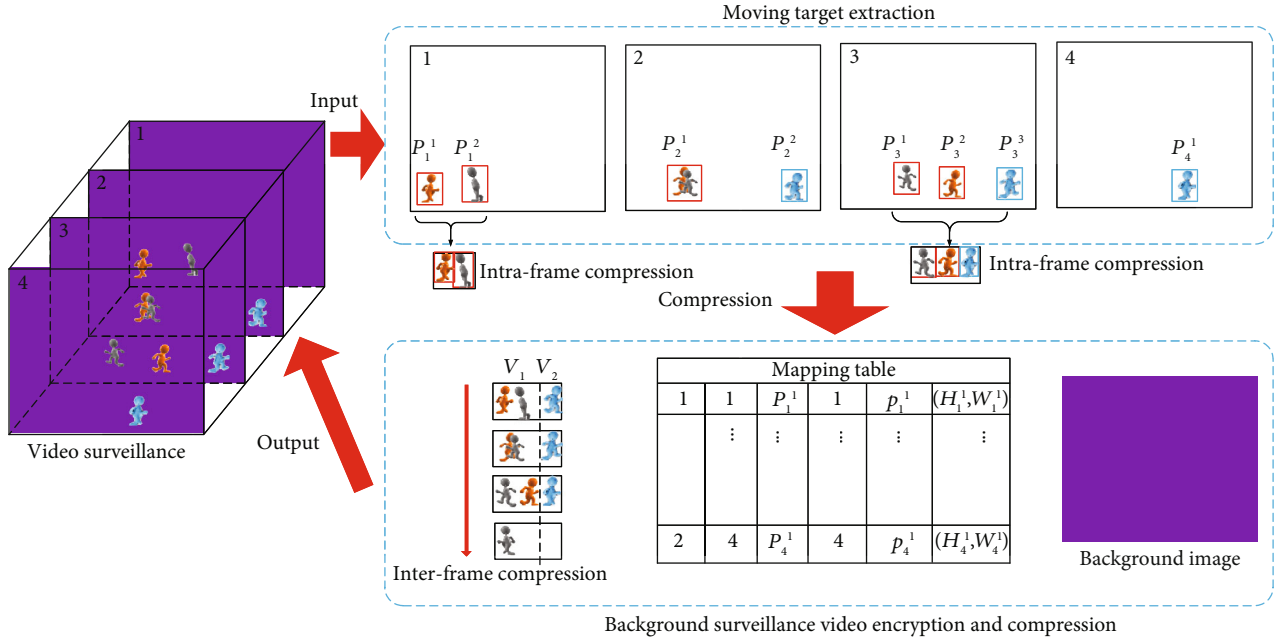


FIGURE 1: The proposed algorithm flow chart.

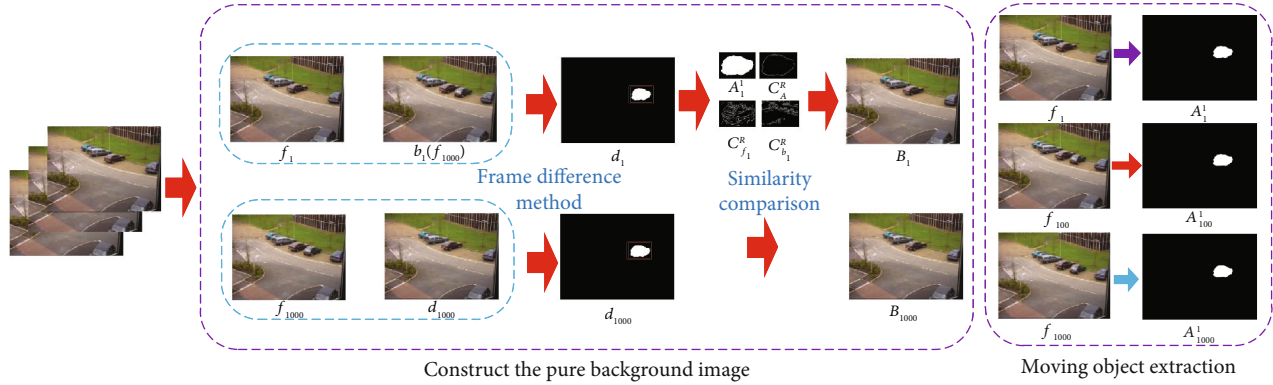


FIGURE 2: Video information extraction flow chart.

B refers to a completely static image in the video, which can only be constructed after observing the whole sequence. Therefore, the frame difference method is used to detect the difference area and determine the pure background from the side. The flow chart is shown in Figure 2.

Step 1. according to the inverse ratio of the correlation between video frames and time interval, the background is initialized with $b_1(x, y) = f_N(x, y)$.

Step 2. select the image frame F_n in sequence according to the video sequence and calculate the difference area according to Eq.(4). The area is opened by mathematical morphology and the set of areas is recorded as $\{A_n\}_m$, which means that m moving areas are detected in the n -th frame, and the i -th moving area is recorded as A_n^i .

Step 3. visual perception is mainly through color and texture features. A_n is detected as a moving area in Step2 through

color features, but it is uncertain whether A_n^i is in F_n or G_n . Take the smallest rectangular area of A_n^i as R , and the Canny operator is used to extract FN, GN. A_n^i is recorded as CR_f , CR_g , and CR_A at the boundary of R . The common pixel points of CR_f , CR_g , and CR_A are counted, respectively, to measure the similarity of CR_f , CR_g , and CR_A . If the similarity between CR_g and CR_A is high, it means that A_n^i is in G_n , which needs to be updated.

Step 4. since there is a certain light intensity difference between F_n and G_n , the corresponding area of A_n^i will be updated directly, and there will be abrupt phenomenon. According to the characteristics of uniform distribution of light, according to

$$E = \frac{\sum_{x,y} (f_n(x, y) - b_n(x, y))}{\text{Nom}}; (x, y) \in \{R - A_n^i\}. \quad (5)$$

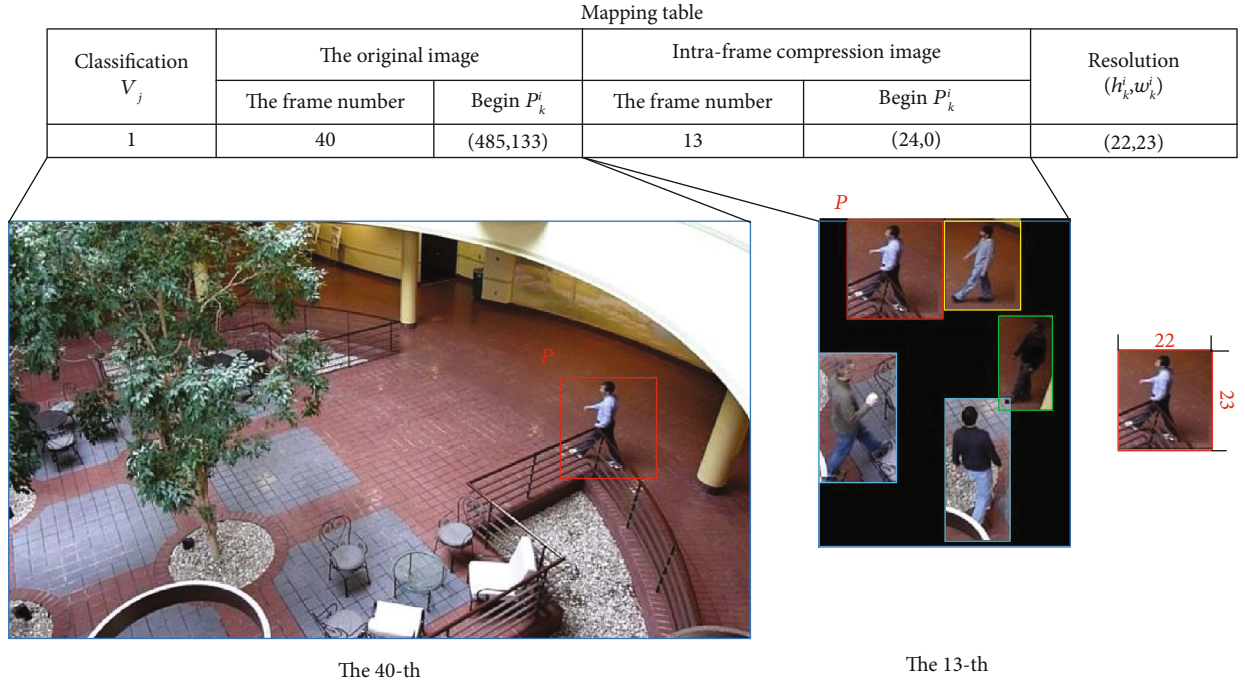


FIGURE 3: The correspondent relationship map.

Calculate the mean value E to simulate the light distribution, where Nom is the number of pixel points that meet the conditions, then the background is updated to

$$b_n(x, y) = f_n(x, y) - E; \quad (x, y) \in A_n^i, \quad (6)$$

where b_n is the pure background B .

Step 5. the frame difference method is used to calculate the difference between $f_n(x, y)$ and B . The difference is regarded as a moving area and the set is recorded as $\{D_n\}_m$.

Based on the above, most of the information contained in the surveillance video is in B and D , and the data can be compressed and decompressed on the basis of B and D , which can greatly reduce the amount of data storage.

2.2. Data Encryption and Transmission. After the processing of the proposed algorithm in the last section, all useful information in the image has been extracted from the video surveillance data, and the amount of image data has been greatly reduced. However, there are still a lot of redundancies in time and space. For this reason, the proposed algorithm is different from the traditional one, which breaks the limitation that the image frame is the minimum compression unit, instead uses the moving object as the minimum unit to compress and store.

At present, the existing video compression and encryption methods are all based on videos, that is, getting videos equals to getting the whole content. The proposed algorithm regards one video as image information and text information. The video information cannot be effectively restored

TABLE 1: Experimental data.

Image revolution	Image frame rate	Scene
640 × 480	15/25/30	Type 1, 2, 3
1280 × 720	15/25/30	Type 1, 2, 3

by acquiring image or text information alone, so the data encryption can be realized. Additionally, the region of interest is only the moving target and pure background in the surveillance video. The pure background does not change for a period of time, instead only the moving target changes. Thus, we only extract the moving target and save them to achieve the compression, and the location information is saved in the form of text.

Since the object motion is shown as continuity and uncertainty of object motion in the image, the 3D connecting area of the image area is marked as V_i , which is used to distinguish the independent path based on the time axis.

Interframe compression aims to reduce the number of stored frames, break the global time axis, and compress in the internal time sequence of V_j to obtain the compressed frame number L .

The goal of intraframe compression is to reduce the size of storage space. Because of the continuity of moving objects, there is a strong correlation between frames, which shows a strong image similarity in the image. In order to break the global space axis and compress the relative position of V_j , the compressed image sequence is obtained. The storage space is H^m in height and W^m in width. The calculation formula is as follows:







Frame	The 1-st frame	The 50-th frame	The 88-th frame	Background reconstruction
Input image sequence				
Traditional frame difference method				
GMM				
Ours				

FIGURE 4: Type 1 results images.

$$\begin{aligned}
 H^m &= \max \left\{ \sum_{i=1}^{L(V_m)} h_k^i \right\}, \\
 W^m &= \max \left\{ \sum_{i=1}^{L(V_m)} w_k^i \right\},
 \end{aligned} \tag{7}$$

where h_k^i and w_k^i represent the length and width of the i -th moving object of the k -th frame image, respectively.

The proposed algorithm compresses the video from two aspects: interframe and intraframe. The amount of data is greatly reduced, and the original video sequence becomes M small sequences and pure background frame B . In order to fully consider the nontamperability of video, M small sequences are spliced into compressed sequence Q , the resolution of the image is $H \times W$, and the minimum resolution of a single-frame image including M small videos is satisfied.

$$\{H, W\} = \min \text{Area}(\{H^1, W^1\}, \{H^2, W^2\}, \dots, \{H^m, W^m\}, \dots, \{H^M, W^M\}). \tag{8}$$

In order to restore the video, we need to match the storage information with the time and space information of the original video. Therefore, we build a mapping list to keep the original video secret and transmit it quickly, as shown in Figure 3.

Since Q only contains moving objects, and the proposed compression algorithm saves each individual moving sequence in a specific area, it can fully guarantee the strong similarity between image frames for further compression of subsequent coding. The traditional residual video compression perception (RVCs) makes every N frames into one

group, and the first frame of each group is the key frame, which encodes the residual part of each frame. RVCs are compressed by the frame, and the selection of n and d_c will directly affect the compression efficiency.

On the basis of extracting image sequences of interest, the RVC algorithm is used to compress image blocks. The proposed algorithm makes full use of interframe and intraframe information, breaks the global time and space correspondence, and only needs to save compressed video sequence and pure background frame image. The corresponding relationship between the storage and the original video is established, and the original video information is retained on the basis of greatly reducing the storage space.

3. Experiment and Result Analysis

The surveillance video as shown in Table 1. They can be regarded as a mixture of the following three situations:

Type 1. : the first image is a pure background, and then the object moves all the time.

Type 2. : the first image contains a moving object, and then the object moves all the time.

Type 3. the first image is a pure background, and then the moving object is static for a long time.

The experiment uses the following database, under the platform of Windows 7 and VS 2015 compiler.

3.1. Comparison of Video Information Extraction Algorithms. To verify the performance of the proposed algorithm in the interest area, we compare the proposed algorithm with the traditional frame difference algorithm and GMM algorithm

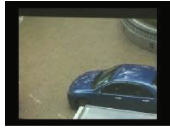

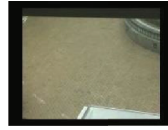












Frame	The 1-st frame	The 411-th frame	The 800-th frame	Background reconstruction
Input image sequence				
Traditional frame difference method				
GMM				
Ours				

FIGURE 5: Type 2 results images.
















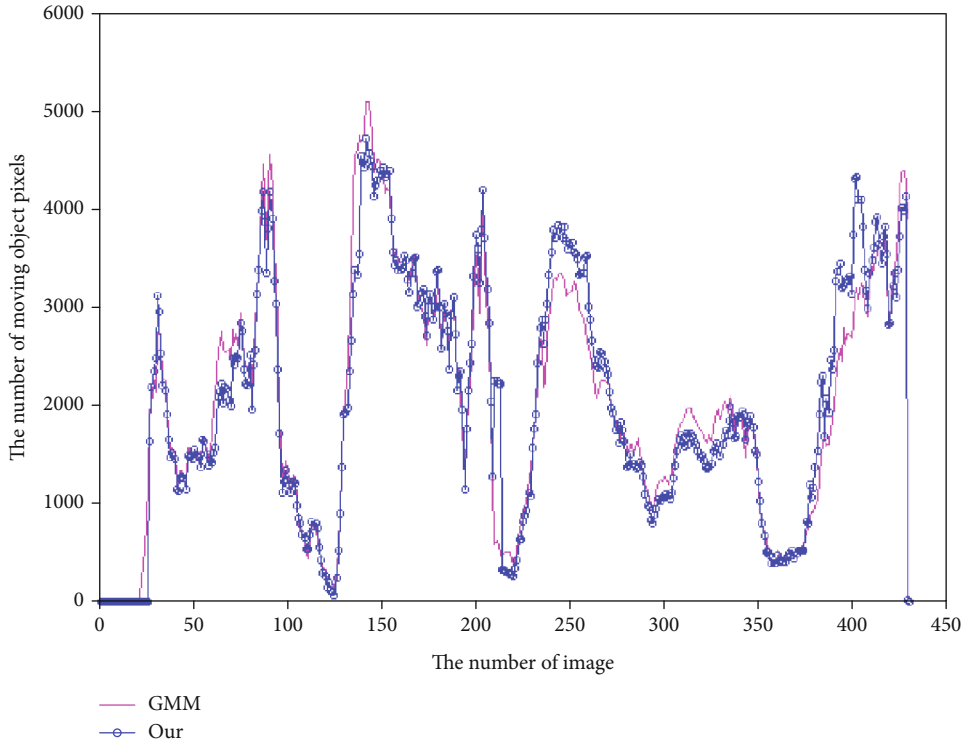
Frame	The 1-st frame	The 670-th frame	The 800-th frame	Background reconstruction
Input image sequence				
Traditional frame difference method				
GMM				
Ours				

FIGURE 6: Type 3 results images.

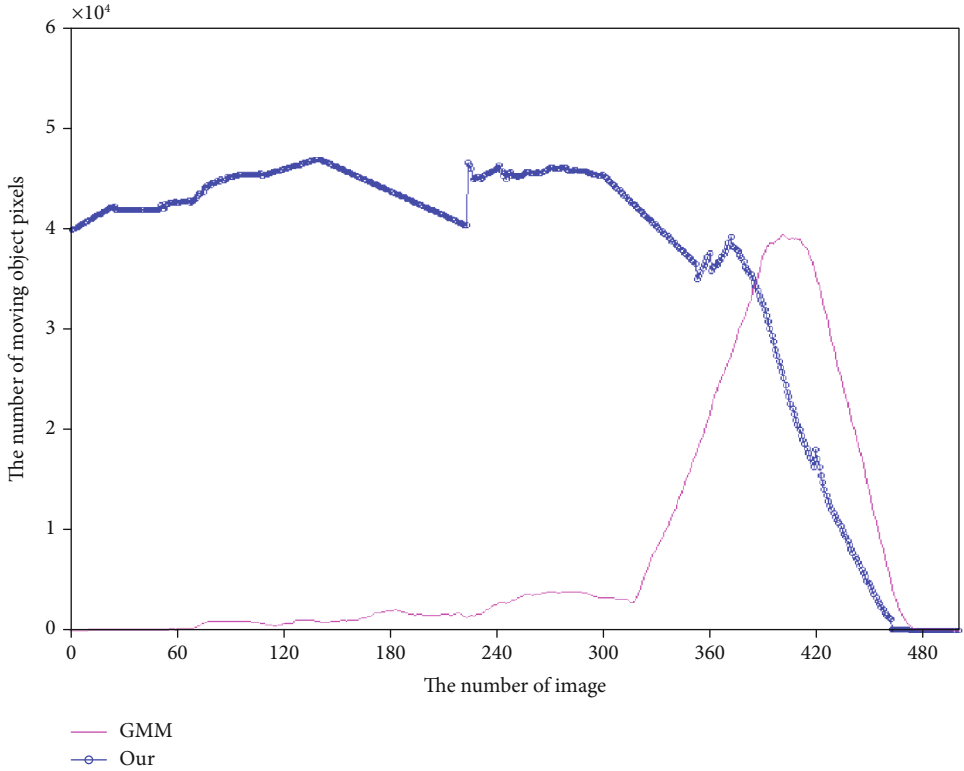
based on the database. To ensure the optimization of the proposed algorithm, the optimal parameters mentioned in the references are applied. The traditional frame difference method uses the first frame as the background frame, the

mixed dimension of the GMM algorithm is $k = 5$, this paper is $t = 30$.

The detection effect of type 1 is good, as shown in Figure 4. Since the first frame is a pure background, the

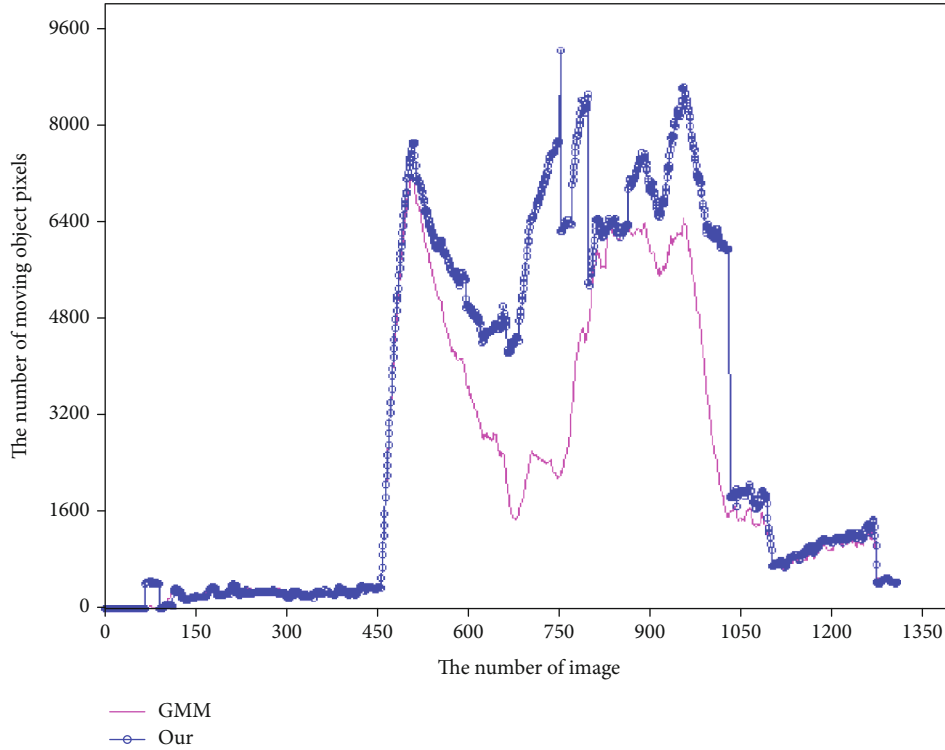


(a) Type 1



(b) Type 2

FIGURE 7: Continued.



(c) Type 3

FIGURE 7: Statistical image of moving object pixels.

traditional frame difference method can extract the moving target better. Because the object is always moving, the Gaussian model established by the GMM algorithm can effectively distinguish the moving area and the background area. The detection effect of type 2 is shown in Figure 5. Because the first frame contains moving objects, the traditional frame difference method does not consider background update, resulting in error detection. The GMM algorithm introduces learning factors to dynamically update the background and moving area, but the learning time takes too long, which will lead to the situation that the tail and some moving objects are still in the background. The detection effect of type 3 is shown in Figure 6. The first frame is a pure background image. The traditional frame difference method is consistent with the proposed algorithm. However, the introduction of learning factors in the GMM algorithm, for the long-time static object after moving, will transform it into the background during the learning process, resulting in the missing detection.

GMM algorithm simulates the distribution of the background and the moving target through multiple Gaussian models, and introduces learning factors to dynamically update the background. However, when the target is stationary for a long time, the GMM algorithm changes it into the background area dynamically, which makes the moving target submerge in the background, and then makes the interested moving target extraction fail. The proposed algorithm, which constructs a pure background extraction algorithm, fully considers the characteristics of moving targets, and effectively suppresses the problem of moving targets sub-

merging into the background. Thus, the performance of this algorithm is better than the GMM algorithm.

The traditional algorithms of background reconstruction and moving target extraction are mainly analyzed from the perspective of iteration and fixed background, so there might be incomplete extraction of moving targets or the situation of moving targets in the background, especially when the light intensity changes. The proposed algorithm extracts the moving target from the dynamic perspective, analyzes its target attributes from the inherent attribute perspective to judge whether it is a moving target or a background area, so the algorithm has strong robustness.

In this paper, the proposed algorithm first establishes the pure background according to the visual perception, and then extracts the moving object in order to extract the moving object completely. Type 1: the proposed algorithm takes the last frame as the initial frame of the background, obtains the moving area through the frame difference method, establishes the visual perception model, judges the moving area in the background frame, and updates the background. Type 2 is similar to type 1. The last frame is a pure background image. The moving area is obtained by the frame difference method, and the visual perception model is established to determine that the moving area is in the current frame, and the background is not updated. Type 3: since the background reconstruction and object extraction are two independent steps, the moving object will not be submerged in the background frame because of staying for a long time. It can establish a pure background and extract the moving object completely.

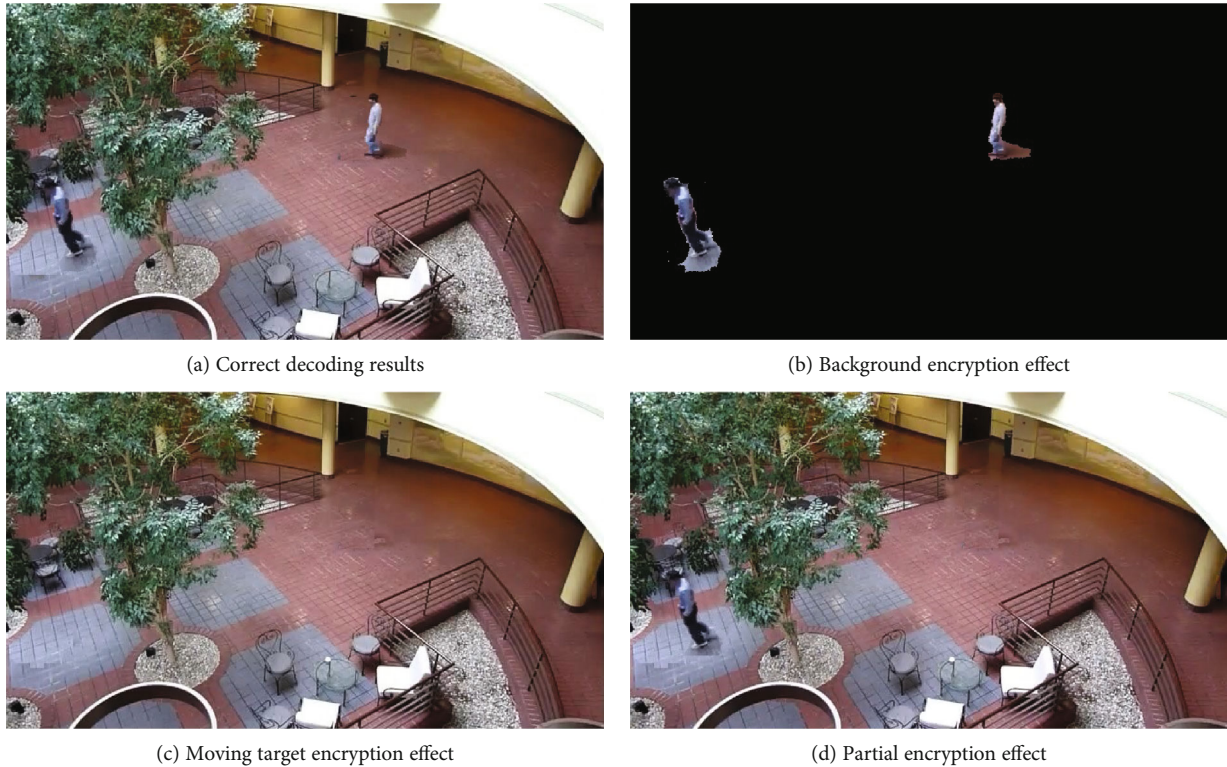


FIGURE 8: Video encryption results.

The number of pixels in the types 1, 2, and 3 of the GMM algorithm and the proposed algorithm are counted to show the superiority of the proposed algorithm. Type 1 is shown in Figure 7(a), since the object is always moving, the GMM algorithm is similar to the proposed algorithm. Type 2 is shown in Figure 7(b), since the first frame contains moving objects, the GMM algorithm gradually learns the background. It is unable to extract the moving objects of the first frame, resulting in the leak detection of moving objects at the beginning. However, the proposed algorithm can effectively solve this kind of problem by reconstructing the background first and then extracting the moving object based on the background model. Type 3 is shown in Figure 7(c), due to the static state of the moving object, the GMM model progressive learning mechanism causes the missing detection when the moving object is submerged in the background at about 500-600 frames. Our algorithm can effectively avoid this situation by building the pure background model. Therefore, the proposed algorithm has strong robustness.

3.2. Data Encryption and Transmission Effect. The proposed algorithm uses video processing methods and introduces the data compression and mapping theory to finally realize the data encryption, as shown in Figure 8(a). The content of the scene cannot be described only by the extraction of the moving target information, as shown in Figure 8(b). The event information cannot be obtained only by the extraction of the background information, as shown in Figure 8(c). Some video content can be shown by the acquisition of some background information and the moving target information, as shown in Figure 8(d). Thus, the video content can be accu-

TABLE 2: Data compression ratio.

Algorithms	Type 1	Type 2	Type 3
MPEG 2	46%	41%	12%
H.264	42%	24%	28%
MPEG 4	50%	43%	17%
OUR	80%	67%	74%

rately reflected only by the acquisition of the video information and location information at the same time. Then, the purpose of video encryption is achieved.

The compression efficiency of the proposed algorithm is proportional to the number of pixels of the moving object in the surveillance video. We analyze the effectiveness of the proposed algorithm from the statistical perspective. The average compression rate of each type of data is shown in Table 2. It can be seen that type 1 has the highest compression rate. Because there is an object that remains static for a long time in the image, type 2 and type 3 have a lower compression rate than type 1.

According to the composition of video surveillance, the proposed algorithm focuses on the data encryption, which is directly proportional to the number and size of moving targets based on the proposed theory.

Based on the uncompressed video, the effect of the mainstream compression algorithms are compared, as shown in Table 2: Both MPEG2 and MPEG4 are modeled to realize the compression from the perspective of motion, and they have a good effect. However, the compression effect is not good for the moving target, which remains stationary for a

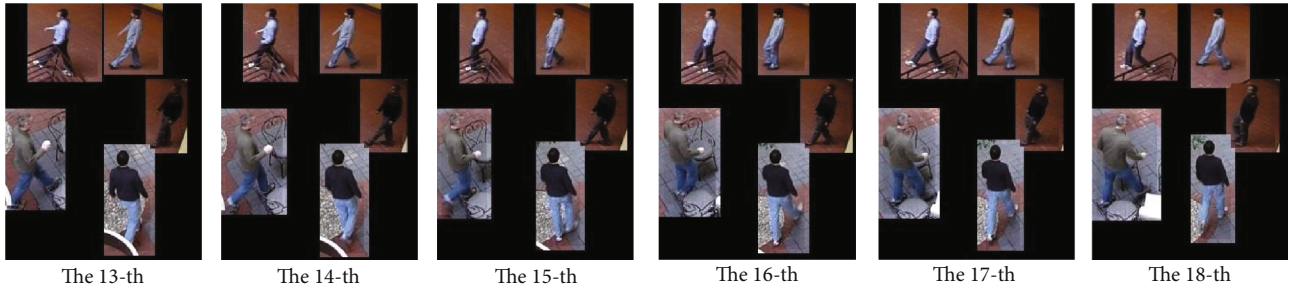


FIGURE 9: Type 1 compression effect images.

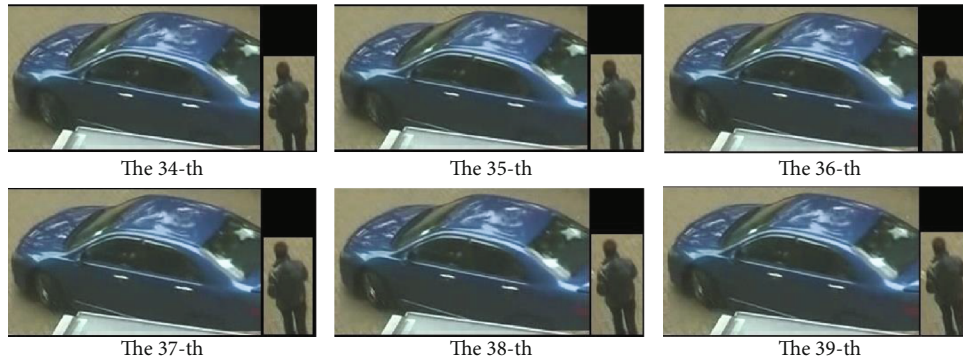


FIGURE 10: Type 2 compression effect images.



FIGURE 11: Type 3 compression effect images.

long time and occupies a large area. H. 264 only applies the same compression strategy for all videos from the perspective of transmission, and the compression effect is relatively average. The proposed algorithm breaks the shackles of the time axis and the space axis for the compression. Compared with the mainstream algorithms, the compression from the two dimensions of time and space has the best effect.

In order to intuitively observe the compressed image effect as shown in Figures 9–11, the proposed algorithm only keeps the changed area, makes full use of the strong correlation of moving objects, divides the independent moving objects into independent areas, greatly reduces the image size and number, and takes advantage of the similarity between moving objects frames to achieve efficient data compression.

3.3. Video Decompression Algorithm Effect. The compressed image is decompressed and reconstructed by the proposed

algorithm according to the mapping relationship, as shown in Figure 12. There is little difference in visual observation. The frame difference between the original image and the restored image is displayed for pixels with a difference greater than 10. For pixels with a difference of more than 10, it can be seen that most of the images appear as scatter noise, which has little impact on the video. However, under the premise of intense light changes, when the object is stationary for a long time, there will be uneven edge distribution and color difference, because of the calculation of the moving object as a whole. It needs further research in the future.

4. Conclusion

In order to solve the problem of surveillance video confidential and efficient transmission, we build a new spatiotemporal model to propose a compression algorithm based on a










	Type 1	Type 2	Type 3
Original image			
Reconstruction image			
Difference image			

FIGURE 12: Image reconstruction comparison.

moving object and background frame. It transforms the compression into the problem of seeking moving object and background. A new data mapping mechanism is built and the compression ratio is more than 60%. It achieves the demand of data transmission confidentially, but for the object color difference caused by a sudden change of the light, it still needs further study.

Data Availability

All used data is within the paper.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work is supported by Light of West China (Grant No. XAB2016B23), the Chinese Academy of Sciences. And the Open Project Program of the State Key Lab of CAD&CG (Grant No. A2026), Zhejiang University.

References

- [1] R. Dey and D. Bhattacharjee, "Single Image De-raining Using GAN for Accurate Video Surveillance," in *Intelligence Enabled Research*, pp. 7–11, Springer, Singapore, 2020.
- [2] A. Goyal, S. B. Anandamurthy, P. Dash et al., "Automatic Border Surveillance Using Machine Learning in Remote Video Surveillance Systems," in *Emerging Trends in Electrical, Communications, and Information Technologies*, pp. 751–760, Springer, Singapore, 2020.
- [3] C. Xiao, L. Wang, M. Zhu, and W. Wang, "A resource-efficient multimedia encryption scheme for embedded video sensing system based on unmanned aircraft," *Journal of Network and Computer Applications*, vol. 59, pp. 117–125, 2016.
- [4] H. Li, K. Wang, X. Liu, Y. Sun, and S. Guo, "A selective privacy-preserving approach for multimedia data," *IEEE Multimedia*, vol. 24, no. 4, pp. 14–25, 2017.
- [5] S. Aljawarneh, M. B. Yassein, and W. A. Talafha, "A multi-threaded programming approach for multimedia big data: encryption system," *Multimedia Tools and Applications*, vol. 77, no. 9, pp. 10997–11016, 2018.
- [6] D. Xu, "Commutative Encryption and Data Hiding in HEVC Video Compression," *IEEE Access*, vol. 7, pp. 66028–66041, 2019.
- [7] N. Khelif, M. B. Amor, F. Kammoun, and N. Masmoudi, "A new evaluation of video encryption security with a perceptual metric," *Journal of Testing and Evaluation*, vol. 48, no. 4, article 20160456, 2020.
- [8] J. Chen, F. Xue, and Y. Kuo, "Distributed compressed video sensing based on key frame secondary reconstruction," *Multimedia Tools and Applications*, vol. 77, no. 12, pp. 14873–14889, 2018.
- [9] T. N. Canh and B. Jeon, "Deep Learning-Based Kronecker Compressive Imaging," in *IEEE International Conference on Consumer Electronic-Asia (ICCE-A)*, 2018.
- [10] A. Adler, D. Boubilil, and M. Zibulevsky, "Block-based compressed sensing of images via deep learning," in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSp)*, pp. 1–6, Luton, UK, 2017.
- [11] K. Xu and F. Ren, "Csvideonet: a real-time end-to-end CSVideoNet: A Real-Time End-to-End Learning Framework for High-Frame-Rate Video Compressive Sensing," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1680–1688, Lake Tahoe, NV, USA, 2018.
- [12] L. Huang, Z. Zhang, S. Wu, and J. Xiao, "Improved video reconstruction basing on single-pixel camera by dual-fiber collecting," in *International Conference in Communications, Signal Processing, and Systems*, pp. 90–97, Springer, Singapore, 2018.
- [13] L. Zhong, S. Wan, and L. Xie, "Adaptive Sampling for Image Compressed Sensing Based on Deep Learning," *Journal of Physics: Conference Series*, vol. 1229, no. 1, article 012016, 2019.
- [14] B. Biswas, S. K. Ghosh, M. Hore, and A. Ghosh, "SIFT-Based Visual Tracking using Optical Flow and Belief Propagation Algorithm," *The Computer Journal*, 2020.

- [15] S. Zheng, X.-P. Zhang, J. Chen, and Y. Kuo, "A new compressed sensing based terminal-to-cloud video transmission system," in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, Sapporo, Japan, Japan, 2019.
- [16] X. Fei, L. Li, H. Cao, J. Miao, and R. Yu, "View's dependency and low-rank background-guided compressed sensing for multi-view image joint reconstruction," *IET Image Processing*, vol. 13, no. 12, pp. 2294–2303, 2019.
- [17] D. M. Rahaman and M. Paul, "Virtual view synthesis for free viewpoint video and multiview video compression using Gaussian mixture modelling," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1190–1201, 2018.
- [18] R. E. Chaudhari and S. B. Dhok, "Fast Quadtree Based Normalized Cross Correlation Method for Fractal Video Compression using FFT," *Journal of Electrical Engineering and Technology*, vol. 11, no. 2, pp. 519–528, 2016.
- [19] H. S. Abbas, M. A. Gregory, and M. W. Austin, "A New Prime Code for Synchronous Optical Code Division Multiple-Access Networks," *Journal of Computer Networks and Communications*, vol. 2018, Article ID 3192520, 11 pages, 2018.
- [20] X. Liu, R. H. Deng, Y. Yang, H. N. Tran, and S. Zhong, "Hybrid privacy-preserving clinical decision support system in fog-cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 825–837, 2018.
- [21] B. Yu, Z. Fu, and S. Liu, "A Novel Three-Layer QR Code Based on Secret Sharing Scheme and Liner Code," *Security and Communication Networks*, vol. 2019, Article ID 7937816, 13 pages, 2019.
- [22] Y. Li, J. Wu, and X. Wang, "Research on moving object extraction method in intelligent traffic video," in *2016 Chinese Control and Decision Conference (CCDC)*, pp. 6059–6063, Yinchuan, China, 2016.
- [23] H. He, Y. Li, and J. Tan, "Relative motion estimation using visual-inertial optical flow," *Autonomous Robots*, vol. 42, no. 3, pp. 615–629, 2018.
- [24] S. S. Sengar and S. Mukhopadhyay, "Detection of moving objects based on enhancement of optical flow," *Optik*, vol. 145, pp. 130–141, 2017.
- [25] X. Ou, P. Yan, W. He et al., "Adaptive GMM and BP neural network hybrid method for moving objects detection in complex scenes," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 2, article 1950004, 2019.
- [26] R. Chavan and S. R. Gengaje, "Multi object detection techniques in video surveillance application," in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, pp. 2011–2015, Chennai, India, 2017.
- [27] C.-H. Yeh, C.-Y. Lin, K. Muchtar, H.-E. Lai, and M.-T. Sun, "Three-pronged compensation and hysteresis thresholding for moving object detection in real-time video surveillance," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 4945–4955, 2017.
- [28] B. Shijila, A. J. Tom, and S. N. George, "Moving object detection by low rank approximation and l_1 -TV regularization on RPCA framework," *Journal of Visual Communication and Image Representation*, vol. 56, pp. 188–200, 2018.
- [29] M. Shang, S. Zeng, and L. Jiang, "A foreground detection algorithm based on improved three-frame difference method and improved Gaussian mixed model," in *Tenth International Conference on Graphics and Image Processing (ICGIP 2018)*, p. 110693T, Chengdu, China, 2019.
- [30] M. S. Zaharin, N. Ibrahim, and T. M. A. T. Dir, "Comparison of human detection using background subtraction and frame difference," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 1, pp. 345–353, 2020.
- [31] J. Guo, J. Wang, R. Bai, Y. Zhang, and Y. Li, "A New Moving Object Detection Method Based on Frame-difference and Background Subtraction," *IOP Conference Series: Materials Science and Engineering*, vol. 242, no. 1, p. 012115, 2017.

Research Article

Location Privacy-Preserving Method Based on Historical Proximity Location

Xueying Guo,¹ Wenming Wang,^{1,2} Haiping Huang ,^{1,3} Qi Li,^{1,3} and Reza Malekian⁴

¹College of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

²School of Computer and Information, Anqing Normal University, Anqing 246011, China

³High Technology Research Key Laboratory of Wireless Sensor Network of Jiangsu Province, Nanjing 210023, China

⁴Department of Computer Science and Media Technology, Malmö University, Malmö 20506, Sweden

Correspondence should be addressed to Haiping Huang; hph@njupt.edu.cn

Received 27 March 2020; Revised 15 June 2020; Accepted 24 June 2020; Published 18 July 2020

Academic Editor: Ximeng Liu

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

With the rapid development of Internet services, mobile communications, and IoT applications, Location-Based Service (LBS) has become an indispensable part in our daily life in recent years. However, when users benefit from LBSs, the collection and analysis of users' location data and trajectory information may jeopardize their privacy. To address this problem, a new privacy-preserving method based on historical proximity locations is proposed. The main idea of this approach is to substitute one existing historical adjacent location around the user for his/her current location and then submit the selected location to the LBS server. This method ensures that the user can obtain location-based services without submitting the real location information to the untrusted LBS server, which can improve the privacy-preserving level while reducing the calculation and communication overhead on the server side. Furthermore, our scheme can not only provide privacy preservation in snapshot queries but also protect trajectory privacy in continuous LBSs. Compared with other location privacy-preserving methods such as k -anonymity and dummy location, our scheme improves the quality of LBS and query efficiency while keeping a satisfactory privacy level.

1. Introduction

With the development of Internet services, mobile communications, and IoT applications, Location-Based Service (LBS) has become one of the popular electronic applications. Users carrying mobile devices loaded with location-based applications, such as Google Maps, Wechat, and Ctrip, are able to send query requests to location service providers (LSPs) and obtain the corresponding service data. With such applications, mobile users can easily obtain information about various Point of Interests (POIs) in the vicinity; for example, users can acquire the bus schedule, the nearest restaurant providing their favorite cuisine, and the recreational facilities from a nearby edge server.

However, since the LSP is potentially untrustworthy, and the submitted queries from users usually include some personal information, such as users' locations and the queried interests, the LSP can easily infer who are doing what in which place, which may jeopardize their privacy. For example, phys-

ical destinations such as medical clinics may indicate a person's health problems. Likewise, regularly staying at certain types of places may be linked directly to one's lifestyles or political associations. Although users may be informed of the policies regarding the collection and distribution of their location data, the execution of these policies is typically beyond the users' control and relies solely on the service providers. Therefore, the privacy of users has not been truly protected and requires further technical attention. Furthermore, LSPs usually need to process a large amount of location service request messages, and the overloaded calculations may cause LSPs to become busy resulting in denial of service.

To address the privacy issue, many technical schemes [1, 2] have been proposed in the literature over recent years. Most of them are based on location perturbation and obfuscation, which employ traditional privacy techniques such as k -anonymity [3, 4]. However, these solutions using k -anonymity have some inherent flaws. First, all mobile users, regardless of whether or not they request LBSs, need to

frequently report their latest locations to the anonymity server. In addition, users without LBSs may not be willing to spend their resources to help others maintain anonymity. Second, excessive location updates from a large number of mobile users also present overwhelming communication and processing bottlenecks on the server side. Third, in addition to the issues mentioned above, another problem is that the area of cloaking regions generated by the existing approaches is highly dependent on the network density. When a user lies in an unpopulated region, its cloaking area may be very large since it needs to contain the user itself and at least $k - 1$ other users. Therefore, these traditional k -anonymity schemes cannot be directly applied to the protection of location privacy due to their inherent flaws.

Trajectory privacy preservation [4] is another challenge in LBSs for the vulnerability of the spatial and temporal information contained in the continuous queries received by the LSP, which may expose users' whereabouts and other private information. It is practically impossible to support anonymity for continuous LBSs using existing techniques such as GM's OnStar services [5]. Continuous LBSs require frequent location updates from their clients. Simply ensuring that each reported location belongs to a cloaking region containing at least k users cannot really achieve the client's k -anonymity protection, and it even significantly increases the computation and communication load of servers. Therefore, how to design a secure and efficient location privacy protection scheme is worth exploring especially in the continuous LBS scenario.

To address the above problems, we propose a new privacy-preserving method based on historical proximity locations. This method ensures that the user can obtain location-based services without submitting the real location information to the untrusted LBS server, which improves the location privacy level and reduces computation and communication load on the server side. In view of the aforementioned issues, the key contributions of this work are summarized as follows:

- (1) In order to avoid the computational overhead of generating pseudolocations on the server side, this paper creatively proposes a scheme that substitutes one existing historical adjacent location around the user for his current location and then submits the selected location to the LBS server
- (2) Historical proximity location query model is adopted to guarantee the location privacy of snapshot queries and continuous queries. In addition, our solution is more difficult for attackers to distinguish the user's true position from historical locations, and at the same time it cannot generate unreasonable positions
- (3) Finally, compared with the existing schemes, performance analysis results show that our proposal can significantly improve the query efficiency while ensuring privacy protection

The remainder of this paper is organized as follows. Related work is reviewed in Section 2. The system model and the proposed privacy-preserving method are introduced

in Section 3. Section 4 presents the experimental results, performance evaluation, and privacy analysis. Finally, we conclude this paper and present future work in Section 5.

2. Related Work

During the past decades, many promising approaches for preserving location privacy in LBSs have been proposed. We roughly divide them into two categories: centralized architecture and noncentralized architecture.

In centralized/edge anonymity server architecture, a centralized entity [6–9] is introduced into the system to protect the location privacy. Under this architecture, k -anonymity is the most popular means used for protecting users' privacy in LBSs. Gruteser and Grunwald [5] originally employed this concept in LBSs. As an extension of the traditional k -anonymity model [10–13], they proposed to reduce the accuracy of users' location information along spatial and/or temporal dimensions for a certain level of anonymity protection. However, all these centralized schemes share some drawbacks: (1) The anonymity server has all the knowledge about users' locations as well as queries, thus it becomes an attractive target for the adversary; so the user's real information will be jeopardized once it is attacked. (2) All users have to continuously send their queries and update their locations to the anonymity server, which causes the anonymizer to be a performance bottleneck and the potential central point of failure for the entire system.

In the noncentralized architectures, users cloak their locations without trusting a trusted third party (TTP). Some approaches, such as obfuscation-based methods [2, 14], cryptographic-based methods [15, 16], and collaboration-based methods [17–19], were proposed to protect the user's privacy. Obfuscation is achieved by adding noise, without revealing the exact location to the LBS servers. For example, Ardagna et al. [2] presented a solution aimed at preserving the location privacy of users by perturbing location information. The main drawback of obfuscation-based methods is that the quality of services (QoS) is degraded because of the low-level accuracy of the query answers. Cryptographic methods are also used to protect privacy data in the LBS; however, they are not practical for mobile devices since they require a powerful computational capability and incur large overhead on the client side. In collaboration-based methods, each user communicates with his peers and collects their location data to generate the cloaking region. The main idea is that, before sending a request to an LSP, the mobile user forms a group with his peers via single-hop communication or multihop routing and generates a cloaking area including k users. Shokri et al. [19] designed a distributed location privacy-preserving algorithm for a collaborative group, called MobiCrowd, which allows users to answer LBS queries from neighboring peers so that querying users can protect their location privacy from the LSP. These approaches focus mainly on snapshot queries, and the problem of protecting location privacy in the continuous LBSs is not considered in the TTP-free methods. With the rise of edge computing, Wang et al. [20] proposed an edge-based model for data collection, in which the raw data from wireless sensor

networks (WSNs) is differentially processed by algorithms on edge servers for privacy computing. To avoid potential information leakage and usage, the user's exact location should not be exposed to the edge node. Tian et al. [21] proposed a stochastic location privacy protection scheme for edge computing, in which the geographical distribution of surrounding users is obtained by analyzing the proposed long-term density map and short-term density map. This scheme is practicable for the real scenario when the edge computing server is honest but curious.

Furthermore, in a few privacy-preserving techniques, an attempt was made to use the TTP model for continuous LBSs [22–24]. Zhang et al. [22] proposed an algorithm for k -anonymity trajectory in LBSs, the main idea of which is to continuously expand an initial cloaked area to include at least the same k users. This means that while a request for an LBS is in progress, no grouped user who participated in the original anonymity set of the requestor is allowed to leave the group, since this action would jeopardize the privacy of the requestor. Xu and Cai [24] exploited historical locations to construct the k -anonymity trajectory and then presented algorithms for spatial cloaking. However, when a user moves on the cloaked path, the LBS can still easily identify the user's actual location if no other user exists on that path.

To address the above limitations, we propose a new privacy-preserving method based on historical proximity locations to protect location privacy in both snapshot queries and continuous queries.

3. System Overview

3.1. Preliminaries

Definition 1. The requested message Q submitted by the user to LSP can be expressed as a five tuple:

$$Q = \{id, loc, t, qry, r\}, \quad (1)$$

where id represents the user's identity information; $loc = \{lx, ly\}$ is the user's location, which can be directly obtained from a Global Positioning System (GPS) or using other positioning devices; t denotes the time at which the user sends the request; qry represents the query content the user wants to submit; and r represents the user's query radius, and naturally, the corresponding query area is πr^2 .

Definition 2. d_{\min} denotes the minimum distance allowed between the user's current location and the historical proximity location selected to be reported to the LSP. This limited distance prevents the selected historical proximity location from being too close to the user's current one to better protect the location privacy. Likewise, in order to guarantee the query quality, d_{\max} represents the maximum distance between the user's current location and the selected historical proximity location.

Definition 3. W_{true} represents the set of POIs the user can obtain under ideal conditions; W is the set of POIs returned

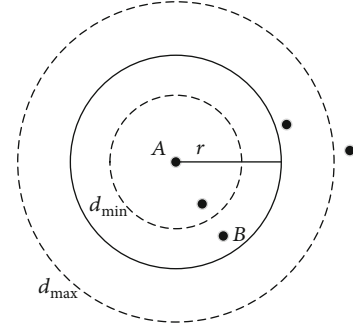


FIGURE 1: Location sampling phase.

by LSP searching according to the user's submitted locations, query content, and query radius.

Definition 4. $P = W_{\text{true}}/W$, which represents the query quality, is the ratio of the number of POIs that the user can obtain under ideal conditions to that of POIs the user receives from the LSP.

3.2. Location Privacy Protection Model. Similar to existing work [25, 26], our system lets mobile users achieve LBSs through an anonymity server, which is considered as a TTP. However, the difference between our centralized architecture and the existing ones is that it can effectively reduce the computing and communication load based on the adopted privacy protection method.

A database that stores a large number of historical proximity locations is essential for the TTP providing privacy service in our model, and the specific characteristics of the database are given as follows:

- (1) Initially, the database may be empty and the users can obtain the location service with k -anonymity protection, during which mobile users report their locations periodically to the TTP, and the k positions utilized in the anonymity process will be subsequently added to the database as historical proximity locations. Unlike existing techniques, such a periodic location update is no longer needed after the initial phase, which may last only a short time period. More location data can be obtained with more and more mobile users participating in the requests of LBSs
- (2) After the initial phase, there are enough historical locations recorded in the database. As shown in Figure 1, suppose that a user is requesting location services at location A, if there are a certain number of historical proximity locations existing in the database that satisfy $d_{\min} < d < d_{\max}$, where d represents the distance between the historical location and the user's current location. In this case, the TTP will select the nearest historical location substituting the current location A and send it to the LSP. A will be subsequently added to the database as a historical location after the query process. However, if there are no historical proximity locations in the database

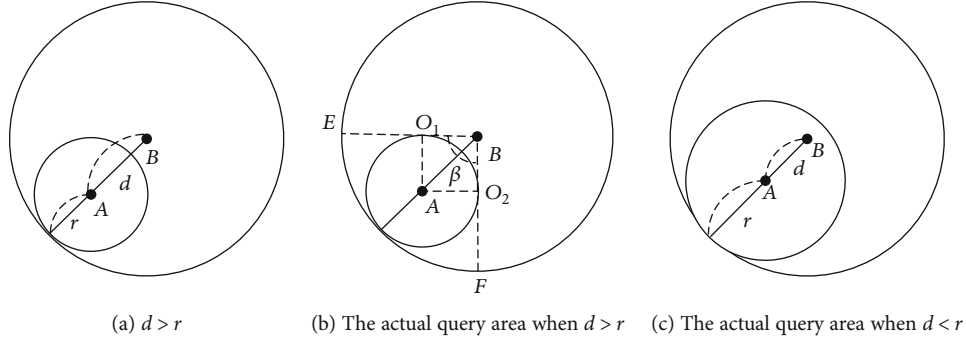


FIGURE 2: The actual query area.

that satisfy $d_{\min} < d < d_{\max}$, the k -anonymity technique will be activated to provide privacy protection services for the user

Obviously, there will be a continuous increase in the number of historical proximity locations recorded in the database, and under this circumstance, the k -anonymity protection is no longer frequently needed.

Furthermore, for efficient retrieval of location data, we index the database using a simple grid-based approach. The entire domain is recursively partitioned into cells in a quad-tree style. Unless a cell has been already at its minimal size (our implementation sets each cell to be at least 200×200 meter²), it is split if the number of locations inside it exceeds a predetermined threshold. Thus, given a cell corresponding to the user's current location, we can effectively retrieve the location data and obtain historical proximity locations.

3.3. Privacy Preservation in Snapshot Queries

3.3.1. Query Area. As shown in Figure 2, the user is located at position A , the query radius is r , the nearest historical proximity location of point A is B , and d is the distance between A and B ($d_{\max} > d > d_{\min}$).

- (1) As shown in Figure 2(a), when $d > r$ is satisfied, a circle is generated with point B as the center and $d + r$ as the radius. Draw two tangent lines (BE and BF) to the circle A via point B with O_1 and O_2 as the tangent points. Wherein, $\angle EBF = \beta$ (denoted in radians) is shown in Figure 2(b). To cover all the possible target positions, the fan EBF is enough as the effective query region, while the actual query region is the whole area of circle B and the area of the fan EBF can be computed as

$$S = S_{EBF} = \frac{\beta R^2}{2} = \frac{2 * \sin^{-1}(r/d)(d+r)^2}{2} = (d+r)^2 * \sin^{-1} \frac{r}{d}. \quad (2)$$

- (2) As shown in Figure 2(c), when $d < r$ is satisfied, if the user wants to query all the target positions, we regard the entire circle which is centered on B as both of the

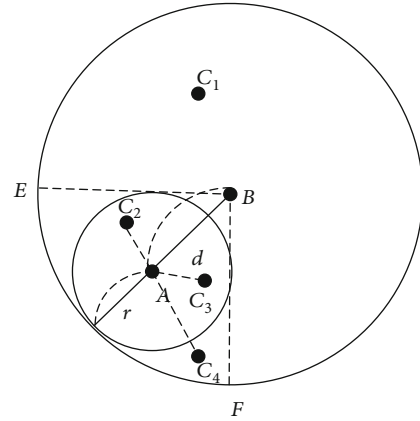


FIGURE 3: Filtering of query results.

effective query region and the actual query region, where $R = d + r$ is the radius, and the area of the query region is $S = \pi(d + r)^2$.

3.3.2. Query Process and Filtering of Query Results. The LSP cannot directly search the irregular area such as the sector area mentioned above during the process of LBS. However, it is feasible to first filter the query results on the TTP side and then filter the results on the client side, which can efficiently reduce the overhead of mobile devices carried by the users. As shown in Figure 3, when $d > r$ ($d_{\max} > d > d_{\min}$) is satisfied, the user at location A , for example, is searching for gas stations nearby with the query radius r . The specific query and filtering process is as follows. Once receiving the request from the user located at A , the TTP will search the database and deliver the information of location B , which is selected carefully as a historical proximity position of A , to the LSP. And then the LSP will search the entire circle B with $d + r$ as the radius, i.e., the actual query region, for target positions meeting the request. After that, the messages related to the gas stations C_1, C_2, C_3 , and C_4 will be returned to the TTP from the LSP as the results. Then, the TTP will filter out C_1 which is out of the user's query area. Finally, the user's mobile device will calculate the distance (d_2, d_3 , and d_4) from location A to the remaining gas station candidates C_2, C_3 , and C_4 , respectively, with the help of a map installed before. Compare each d_i with the query radius

r ; if it is smaller than r , the corresponding information will be retained, otherwise it will be deleted, so C_4 will be filtered out as a result. Ultimately, the location information of gas stations C_2 and C_3 will be sent to the user.

When $d < r$ is satisfied, the process is similar, which is not repeated here.

It is worth noting that due to the indirect query method in our scheme, the error of the distance d between A and B will also lead to the error of the actual query radius $d + r$, which may result in the actual query area being too large or too small, possibly accompanied with a declined quality of services.

3.4. Privacy Preservation in Continuous Queries. Existing techniques mostly focus on snapshot queries. However, privacy preservation in continuous LBSs is more challenging than that in snapshot queries because adversaries could use the spatial and temporal correlations on the user trajectory to infer the user's private information. To deal with the concern, a privacy-preserving method for continuous LBSs based on historical adjacent locations is described in this section.

3.4.1. Average Query Error. As mentioned above, in snapshot queries, the error of the distance d between the user's current location and the reported location, which is selected from the historical proximity locations in the database by the TTP, may bring about a decline in the quality of queries. Similarly, it is the same in privacy preservation scenarios of continuous LBSs queries. We give a formal definition of the average error degree in continuous LBSs as follows, where d_i is the distance between A_i and B_i , $d_{\max} > d_i > d_{\min}$.

Definition 5. Given a trajectory $T = \{A_0, A_1, \dots, A_n\}$, which is generated by the user over a period of time, where A_i represents the user location at the time point i ; in response, the TTP will compute a new trajectory $T' = \{B_0, B_1, \dots, B_n\}$ based on T using historical proximity locations, where B_i represents the historical proximity location of A_i at the time point i . The average query error can be defined as

$$\bar{d} = \frac{\sum_{i=0}^n d_i}{n}, \quad (3)$$

Obviously, the smaller \bar{d} is, the smaller the error degree will be. For the quality of queries, \bar{d} in the query process needs to be as small as possible. Therefore, in the process of the user's moving on the trajectory, it is better to select the nearest historical proximity location B_i substituting the corresponding A_i when sending it to the LSP.

3.4.2. Trajectory Privacy-Preserving Algorithm. If there are enough historical proximity locations around the user's trajectory T , it is easy to find the corresponding B_i for each A_i , and B_i will not coincide with any B_j , where $0 \leq i \leq j \leq n$.

However, if the historical proximity locations around the trajectory T are sparse, there is a certain possibility that B_i and B_j coincide with each other. As shown in Figure 4, the directed lines denote a trajectory formed by the user over a

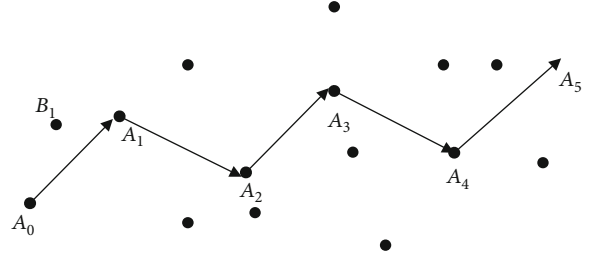


FIGURE 4: B_1 is selected as the historical location for both A_0 and A_1 .

period of time, and the solid nodes nearby denote the existing historical proximity locations. Since B_1 is both the nearest historical proximity location of A_0 and that of A_1 , when the user is at the 0th time point and the 1st time point, B_1 will be selected and sent to the LSP for query results on behalf of A_0 as well as A_1 , resulting in the same selection of historical proximity locations at different time points, i.e., $B_0 = B_1$. In this case, once the LSP receives the same location B_i at different time points, it will be easy to infer that the user is wandering in the vicinity of B_i during this period of time, which actually leaks the user's privacy.

To solve this problem, we make further constraints and give Definition 6.

Definition 6. Given a trajectory $T = \{A_0, A_1, \dots, A_n\}$, which is generated by the user over a period of time, where A_i represents the location of the user at the time point i ; there is a new trajectory $T' = \{B_0, B_1, \dots, B_n\}$ as the historical proximity trajectory (HPT) of T , where $0 \leq i < j \leq n$, $B_i \neq B_j$, and B_i represents the corresponding historical proximity location selected for location A_i .

Therefore, aiming at the problem for the privacy preservation of continuous LBSs, the key to our solution is how to find the corresponding historical proximity trajectory (HPT) T' for the user's trajectory T while guaranteeing the minimum value of \bar{d} on the premise of satisfying both Definition 5 and Definition 6. The following is the specific solution description for this problem.

Given a trajectory $T = \{A_0, A_1, \dots, A_n\}$ of the user, let $T'' = \{C_0, C_1, \dots, C_m\}$ be an ordered set of historical proximity locations along the direction of trajectory T , satisfying $d_{\max} > d_k > d_{\min}$ (d_k is the distance from A_k to any location C_{k+i} among $C_k \sim C_{k+m-n}$) and $m \geq n$, where m denotes the number of historical proximity locations around the trajectory T , and n represents the number of locations the user left on trajectory T . Then, the minimum sum of error degree between the historical proximity trajectory T' and the user's trajectory T is defined as $D(n, m) = n * \bar{d}$. If C_m is selected as the historical proximity location of A_n and sent to the TTP, then the solution for getting the minimum value of $D(n-1, m-1)$ is certainly contained in the solution for getting that of $D(n, m)$. If C_m is not selected to substitute A_n , then the optimal solution for getting the minimum value of $D(n, m)$ is bound to contain the solution

Input: T, T''

Output: Array B, is used to record the locations selected from T'' and reported to LSP as historical proximity locations of user's trajectory T

```

for  $i = 0$  to  $n$  do
   $D[i][i] = D[i-1][i-1] + d[i][i]$ 
   $B[i] = i$ 
for  $i = 0$  to  $n$  do
  for  $j = i + 1$  to  $m$ 
    if ( $D[i-1][j-1] + d[i][j] > D[i][j-1]$ )
       $D[i][j] = D[i][j-1]$ 
       $B[i] = j - 1$ 
    else
       $D[i][j] = D[i-1][j-1] + d[i][j]$ 
       $B[i] = j$ 
return B

```

ALGORITHM 1: selectHistoryLocation(T, T'').

for getting that of $D(n, m-1)$. Therefore, the recursive relationship can be denoted as follows:

$$D[n][m] = \begin{cases} \sum_{i=0}^n d[n][m], & n = m, \\ \min \{D[n-1][m-1] + [n][m], D[n][m-1]\}, & n < m, \end{cases} \quad (4)$$

where $d[n][m]$ represents the distance between A_n and B_m .

The pseudocode of the above procedure is given in Algorithm 1.

In Algorithm 1, array B holds the subscripts of the selected locations on T'' , and the complexity of the algorithm is $O(n^3)$. Algorithm 2 is used to get the historical proximity trajectory T' that guarantees the minimum value of \bar{d} .

Besides, there is still a special situation needing a discussion. It is likely that the quantity of the historical proximity locations recorded in the database is not enough for the algorithm we proposed. As is shown in Figure 5, when m is much smaller than n , no matter how it is selected, it will occur that one historical proximity location is selected two or more times on the user's trajectory. Considering the peculiarity of this problem, we propose to employ a symmetry mechanism to generate dummy locations in our scheme, and the specific procedure is described as follows.

As is shown in Figure 6, when there are no other historical locations available except one existing historical proximity location B_k (for example B_1) of location A_i (for example A_1), it connects B_k to A_i and extends the connecting line to point V_j (for example V_1), making $B_k A_i = V_j A_i$, where V_j is the dummy location generated as a historical adjacent location of A_i by symmetry. However, it is possible that the dummy location generated by symmetry is unreasonable (for example, the dummy location is in a lake), so some adjustment is necessary. As shown in Figure 7, suppose that the generated dummy location V_1 is unreasonable, and the TTP will rotate V_1 and adjust the distance from V_1 to A_1 to make it meet the rationality requirements, and finally a rea-

Input: T''
Output: T'
 for $i = 0$ to n do
 $T'[i] = T''[B[i]]$
 return T'

ALGORITHM 2: getHistoryTrack(T').

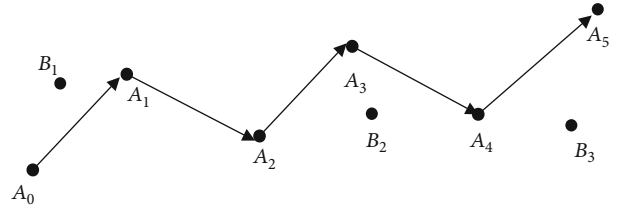


FIGURE 5: Sparse historical proximity locations.

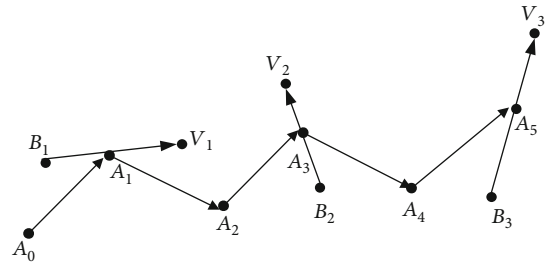


FIGURE 6: Generating dummy locations by symmetry.

sonable dummy location V'_1 will be obtained as the historical proximity location of A_1 .

Besides, there still exists a small possibility of $n > 2m$, in this case the number of historical proximity locations is smaller than n (the number of locations on trajectory T), even if the number of historical locations is expanded from m to $2m$ with the aid of the symmetry mechanism. To deal with this issue, we can activate the k -anonymity

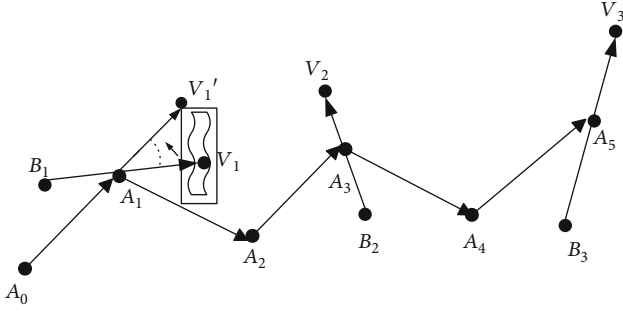


FIGURE 7: Generating dummy locations by symmetry and rotation.

technology to protect location privacy and add the user's locations to the database.

4. Experiment and Analysis

In this section, the experimental evaluation of the feasibility and efficiency of our proposed method under various parameter settings will be presented. Firstly, we analyze the influence of several parameters on the average query error \bar{d} . Secondly, we compare our method with other location privacy-preserving techniques in terms of query efficiency, query quality, and anonymity degree. The experimental region is within 10 square kilometers of the Sanpailou Campus of Nanjing University of Posts and Telecommunications. The data utilized in the experiments are captured by the coordinate pickup tool provided by Google. Our experiments are implemented with the Java Development Kit- (JDK-) 1.7 and Eclipse Integrated Development Environment (IDE), running on a local machine with an Intel Core-i5 2.8 GHz, 8 GB RAM, and Microsoft Windows 7 OS.

4.1. Influence Factors of the Average Query Error \bar{d} . Within the range of the experimental region, 10 coordinate points are generated randomly to construct a trajectory T of the user, i.e., let $n = 10$. And then 20~40 locations from the database are selected as historical adjacent positions around the user's trajectory generated before.

d_{\min} is set by the user, and a smaller d_{\min} has more probability to be taken to ensure the quality of services in a densely populated area; on the contrary, in a sparsely populated area, a larger d_{\min} means better privacy level. As shown in Figure 8, it can be seen that \bar{d} increases with the increase of d_{\min} : when selecting historical adjacent locations, it is necessary to consider whether the distance d from the user to the historical adjacent location is larger than d_{\min} , so as to exclude some positions that are too close to the user. The more there are historical proximity locations, the smaller d will be, and this results in a smaller average error degree \bar{d} . Besides, \bar{d} approaches d_{\min} infinitely as m approaches infinity.

d_{\max} is also set by the user, and usually it cannot be set too small. Since d_{\max} is the maximum distance between the user's current location and the historical proximity location reported to the LSP, a d_{\max} that is too small will filter out most historical adjacent locations, reducing the privacy protection level. As shown in Figure 9, when m takes the value

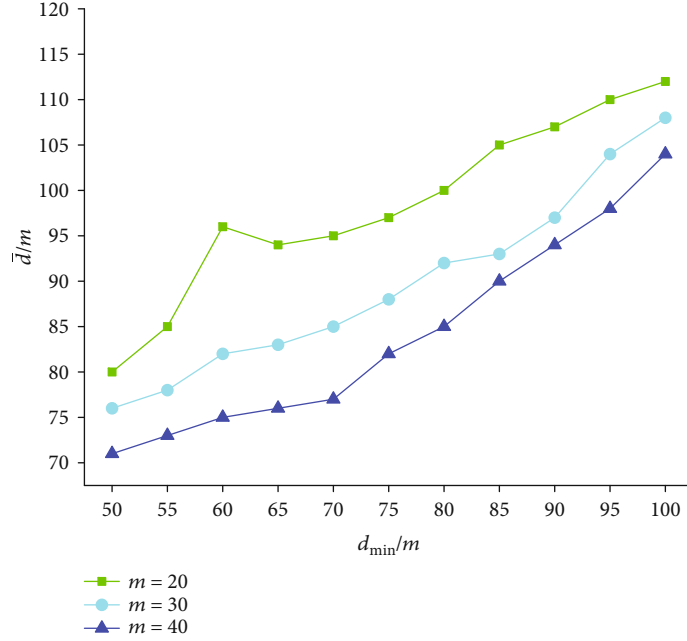
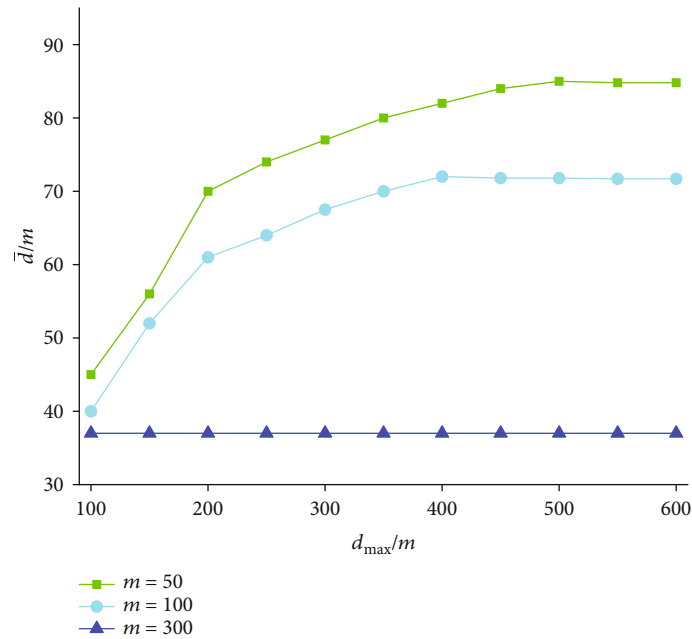
of 50 and 100, respectively, \bar{d} increases as d_{\max} grows in the initial phase. This is because when m and d_{\max} are both small, the number of the screened historical locations m' is smaller than n , and dummy locations will be generated as historical adjacent locations by the symmetry mechanism. Therefore, there is more probability of selecting the nearest historical locations, leading to a smaller \bar{d} . However, when d_{\max} gradually grows, m' will also increase as the screening conditions for historical locations are relaxed, so the number of historical locations generated by the symmetry mechanism will decrease, accompanied with an increase of \bar{d} . Until there is no need for generating symmetrical historical locations, d_{\max} will no longer affect the historical locations selected. When $m = 300$ and $d_{\max} = 100$, the n locations closest to the trajectory T selected from the m historical points are not screened out, so \bar{d} remains constant as d_{\max} increases.

We have discussed the influence of historical adjacent location parameter selection on \bar{d} . The experimental results clearly show the specific effects of different values of d_{\min} and d_{\max} on \bar{d} . Therefore, in practical application scenarios, the values of parameters d_{\min} and d_{\max} should be selected according to specific requirements and allowable errors.

4.2. Performance Comparison

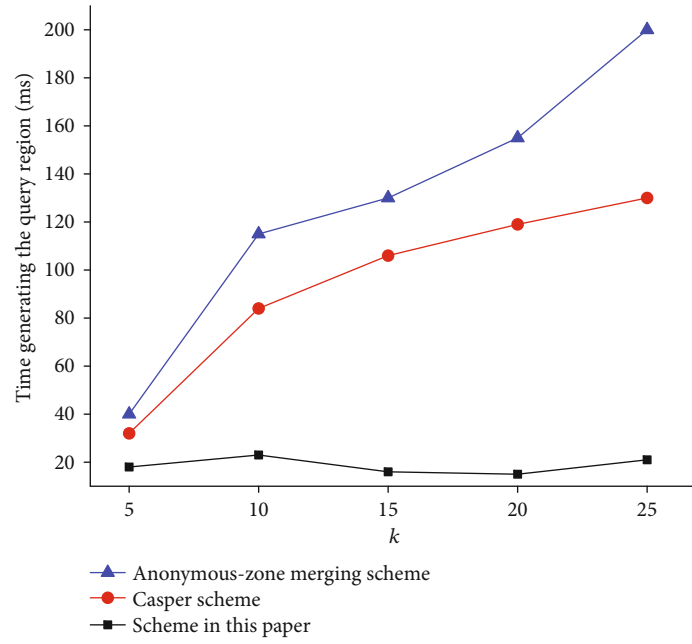
4.2.1. Performance Comparison under Snapshot Queries. In our experiments, coordinate data of 500 target positions such as hotels, hospitals, and gas stations were captured, and 5000 coordinate points were randomly selected as historical adjacent positions as well as other users' positions required when using k -anonymity and were stored in the database.

(1) Query Efficiency. The query efficiency is usually synthetically evaluated with the total time cost that contains TTP spending on generating the actual query area and the LSP spending on replying to the requested query. As shown in Figure 10(a), the Casper scheme in [25] and the anonymous-zone merging scheme in [17] generate the query region using k -anonymity, so the query domain generation time is the sum of the time of the database searching other $k - 1$ users around and that of constructing the anonymous domain containing k users. However, in most cases, the query domain generation time of our solution is almost the time to search for historical adjacent locations in the database, which has nothing to do with k . Therefore, the time to generate the query region in our scheme is relatively less and does not increase linearly with the increase of k . As shown in Figure 10(b), when the same query radius $r = 300$ m is taken and d is set to 75 m, the query region in our scheme is independent of k and its area does not exceed $\pi (d + r)^2$; the area of query region generated by the Casper scheme is theoretically no less than $k\pi r^2$, which is larger than those of our scheme and the anonymous-zone merging scheme; besides, the query areas of the two compared schemes increase significantly with the increase of k . Sufficient historical locations will ensure a smaller d in our scheme, and thus guarantee a smaller query area. Figure 10(c) illustrates that the query processing time is

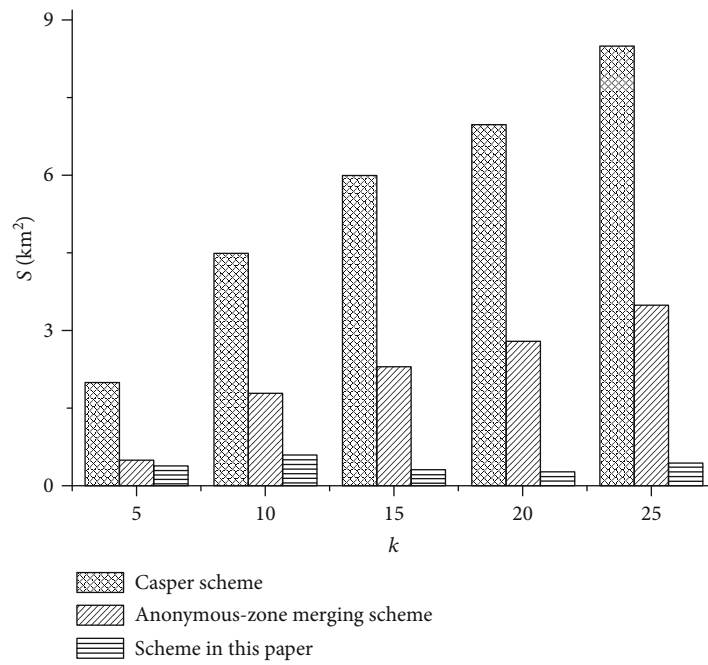
FIGURE 8: Impact of m and d_{\min} on \bar{d} .FIGURE 9: Impact of m and d_{\max} on \bar{d} .

positively correlated with the query area. In order to facilitate the comparison between our scheme and the other two k -anonymity schemes, a smaller anonymity degree $k=5$ is taken and d is set to 75 m. As shown in Figure 10(d), both the query area and the query processing time gradually increase when the radius grows. However, compared with the Casper scheme and the anonymous-zone merging scheme, the query area generated by our scheme is relatively small, which results in a shorter query processing time.

(2) *Query Quality.* Evaluation of the query quality is based on the ratio P of the number of POIs that the user can obtain in theory to that of positions returned by the LSP when the user requests with the query radius r , i.e., $P = W_{\text{true}}/W$, as explained in the previous study. In the experiment, we randomly select 20 points as the positions where the user can send the query. We vary the value of r , repeat the experiment 20 times, and then take the average value of P as the analysis object. Furthermore, we also compare our scheme with the



(a)



(b)

FIGURE 10: Continued.

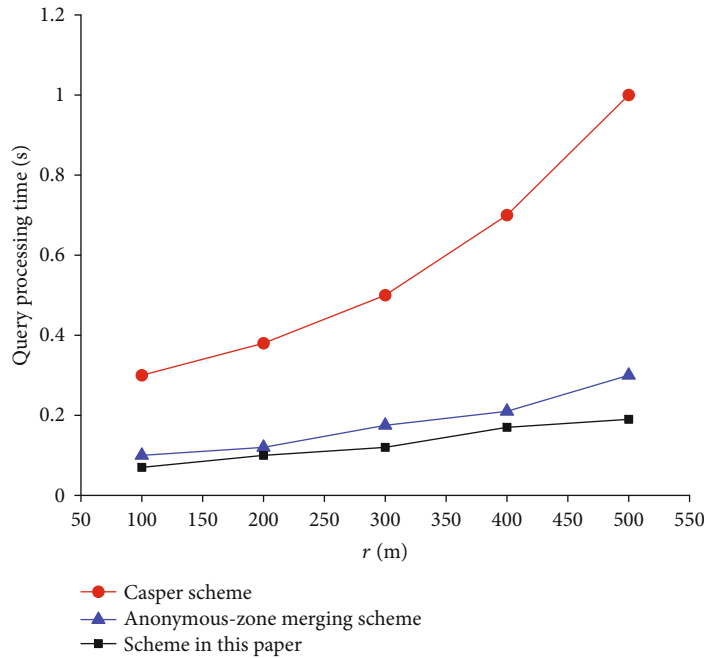
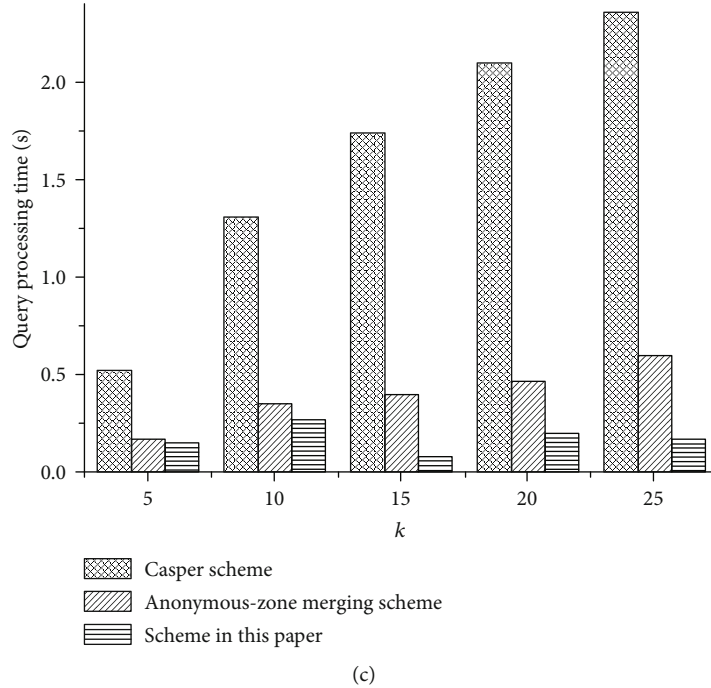


FIGURE 10: Query efficiency comparison. (a) Time comparison of generating the query area. (b) Query area comparison. (c) Query processing time comparison. (d) Impact of r on query processing time.

enhanced pseudonym selection scheme in [26] besides the other two schemes mentioned before. As shown in Figure 11, our scheme maintains satisfactory query quality and stability with the increase of the query radius. In contrast, as for the Casper scheme and the anonymous-zone merging scheme, the query area increases significantly as r becomes larger, which indicates that a great number of POIs cross the user's query area, resulting in the decline of query quality. In addition, since the enhanced pseudonym selection scheme

cannot flexibly adjust the query region to cover all the target positions, it is difficult for it to guarantee high query quality.

Experimental results show that, our scheme can effectively improve the query efficiency while guaranteeing satisfactory query quality in snapshot queries.

4.2.2. Performance Comparison under Continuous Queries. In the experiment, we select \bar{d} , defined as the average query

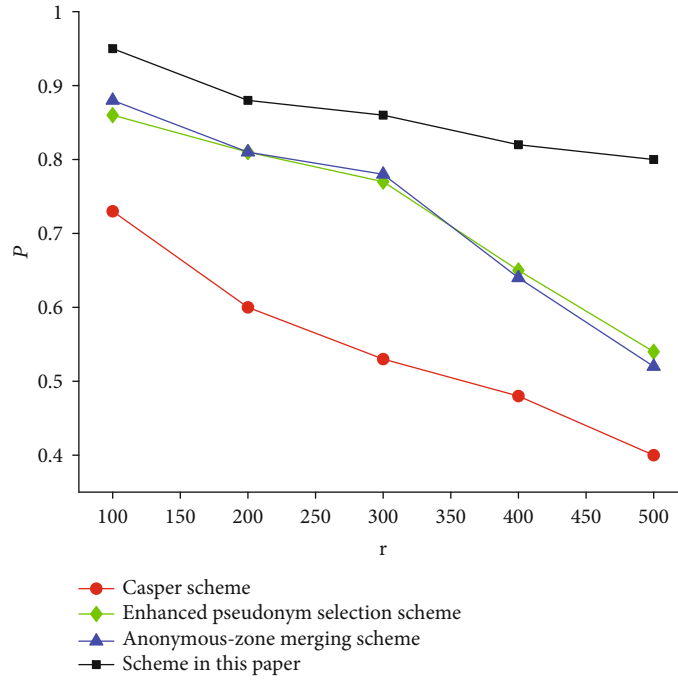


FIGURE 11: Query quality comparison.

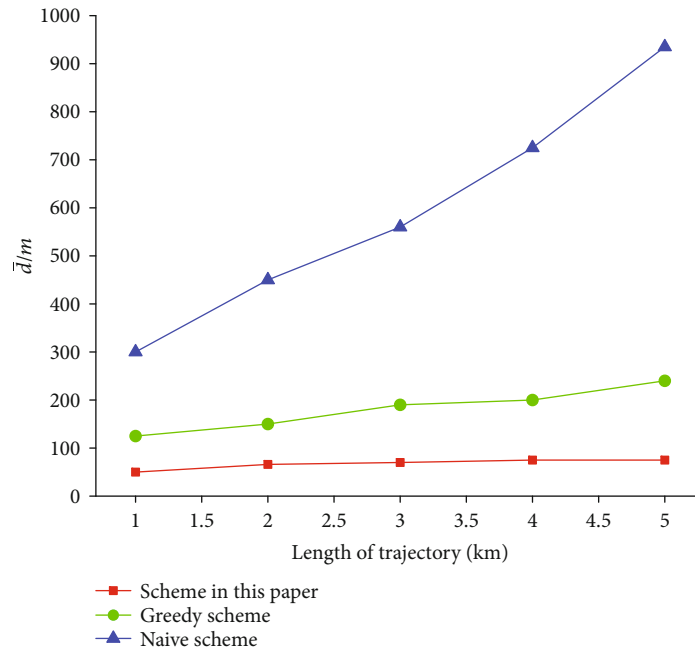
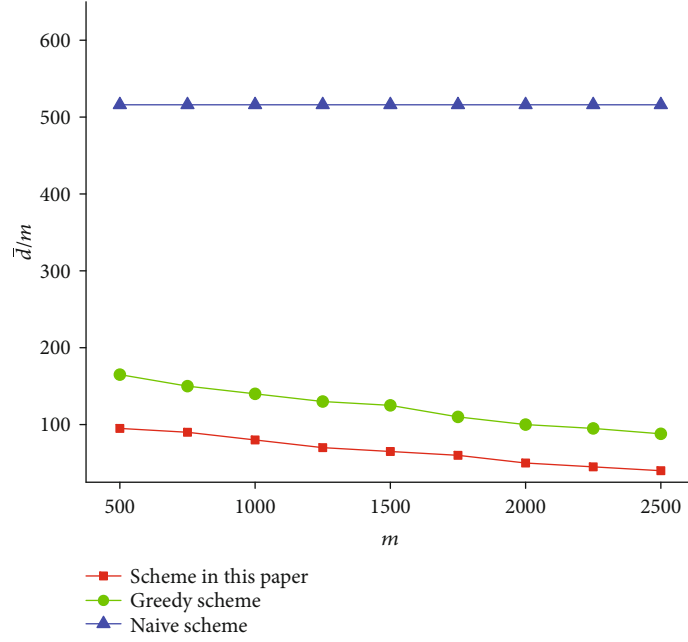
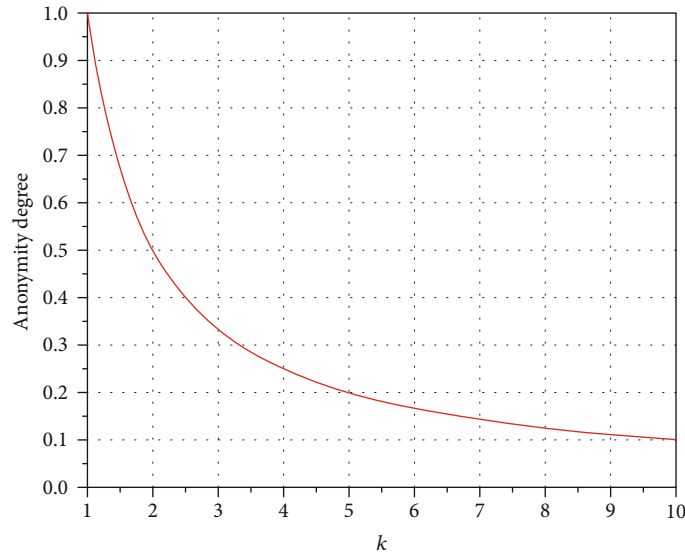


FIGURE 12: Impact of the length of trajectory on \bar{d} .

error for a user’s trajectory in continuous LBSs, as our performance evaluation metric. And obviously, the smaller \bar{d} is, the better the quality of services will be in continuous LBSs. We compare our scheme with two other existing schemes, the Native scheme in [22] and the Greedy scheme in [24], which are both extensions of the k -anonymity method. Within the experimental region, the length of the user’s trajectory was set to 1-5 km, and 500-2500 coordinate points were captured

within the radius of 200 m around the trajectory as historical proximity locations and added to the database.

We set $m = 1000$, and the experiment results are shown in Figure 12. In the Native scheme, the cloaking area will become increasingly large since the traditional trajectory k -anonymity method expands an initial cloaking region to cover at least the same k users who may move in different directions, resulting in a sharp increase of \bar{d} . Moreover, the

FIGURE 13: Impact of m on \bar{d} .FIGURE 14: Anonymity degree about k .

query sequence is consistent with the user's movement direction, which may provide some valuable information for the adversary to infer the user's trajectory. In the Greedy scheme, the Greedy algorithm is utilized to verify that each node on the candidate $k - 1$ trajectories is as close to the user's trajectory as possible; however, since a complete historical trajectory will be finally selected from the $k - 1$ candidates, it cannot guarantee that each position on the selected historical trajectory is the nearest point for each node on the user's real trajectory.

The following is the analysis of the impact of m (the number of historical proximity locations) on \bar{d} of the three schemes, and the length of the user's trajectory is set to 3 km. As shown in Figure 13, for the Native scheme, m has

no effect on \bar{d} since the generated cloaking area is only relevant to the current locations of the other users. However, \bar{d} declines with the increase of m for the Greedy scheme and our scheme, since historical trajectories and historical proximity positions are utilized in the two schemes, respectively.

4.3. Privacy Analysis. In this section, we will evaluate the privacy degree of our solution by comparing it with the k -anonymity and dummy location technology.

In the process of k -anonymity protection, the LSP receives the locations of k users involved with the service requestor, so the probability of identifying the user's real location is $1/k$. As shown in Figure 14, the larger the value

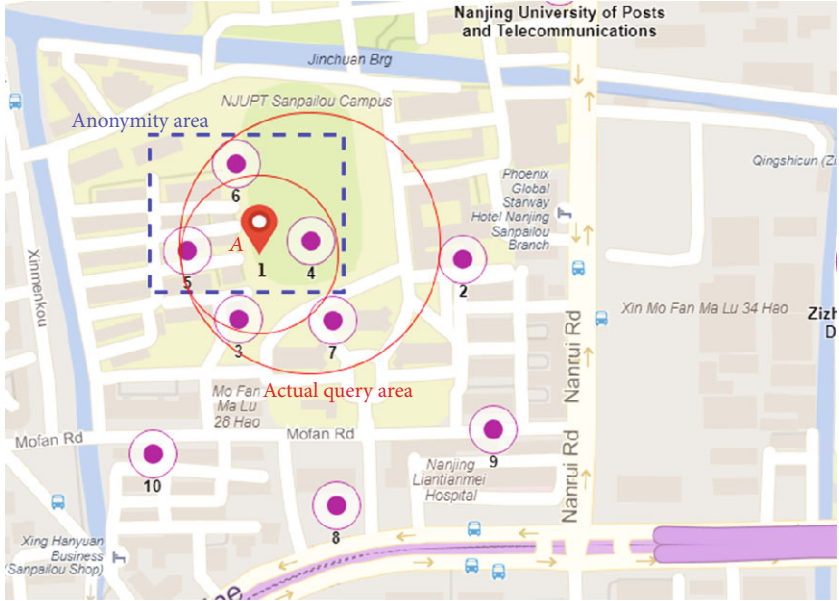
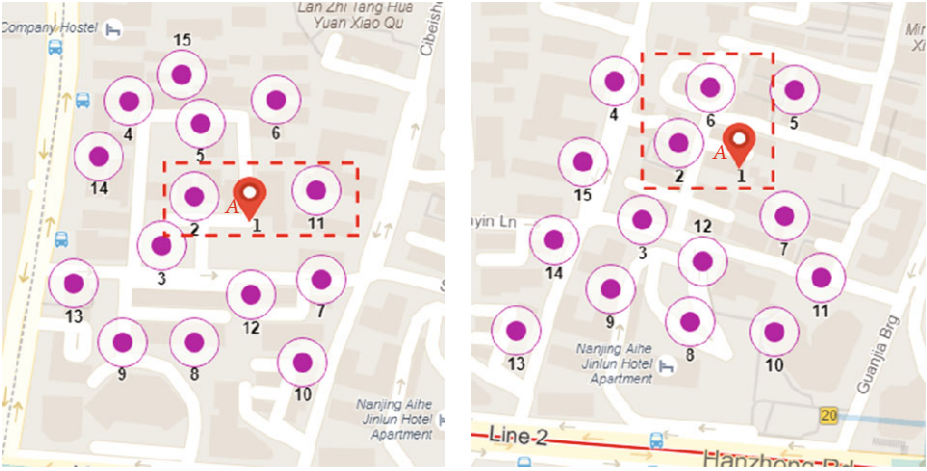
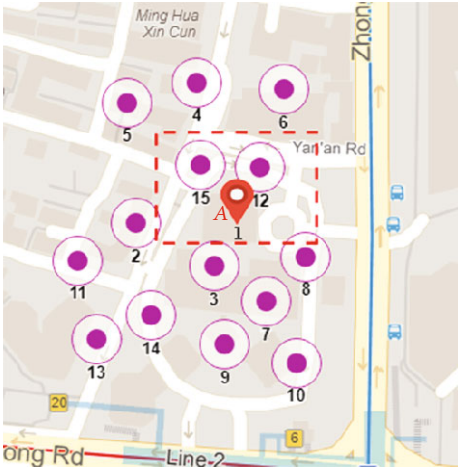


FIGURE 15: Privacy analysis in snapshot queries.



(a) At time point t_1

(b) At time point t_2



(c) At time point t_3

FIGURE 16: Privacy analysis in continuous queries.

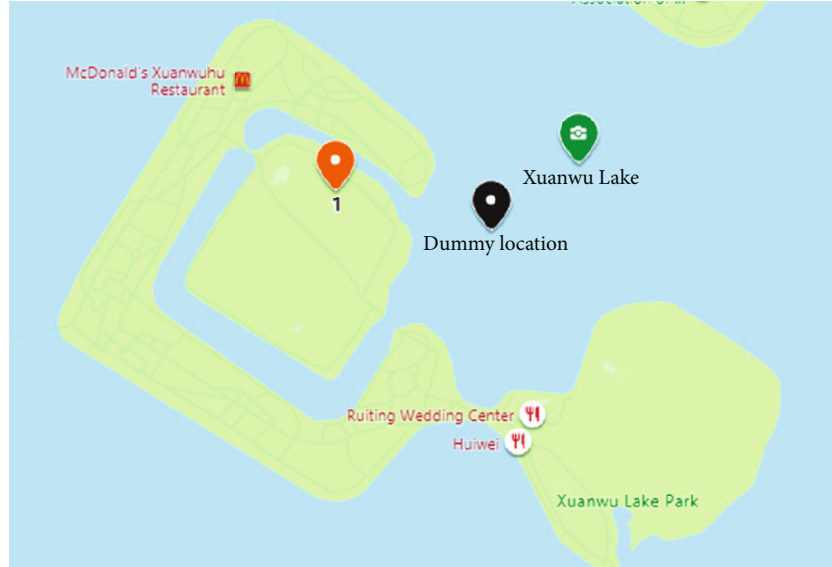


FIGURE 17: Unreasonable locations generated by dummy location.

of k , the higher the privacy degree will be; however, it will also lead to a decrease in query efficiency and quality. As shown in Figure 15, a user A , located at position 1 in the Sanpailou Campus of Nanjing University of Posts and Telecommunications, wants to request the location service together with 9 other users in the vicinity who also request services. Suppose that user A obtains location service through k -anonymity with a cloaking area containing users in positions 4, 5, and 6, the probability that the adversary recognizes user U will be $1/4$. However, if user A adopts the scheme as described in this paper, point 4 will be treated as a historical adjacent position to be queried. By the description of the proposed scheme, the actual query area covers a total of 6 points including user A and points 1, 3, 4, 5, 6, and 7, which is denoted by the big red circle in Figure 15. Therefore, the probability of identifying user A is only $1/6$.

The advantages of our proposal will become more obvious in continuous queries under densely populated areas. We take Xinjiekou, the commercial center of Nanjing City, as an example. As shown in Figure 16, user A sends a service request with $k=3$, the anonymous set of which is $\{1, 2, 11\}$ at the initial time t_1 ; while that updates to $\{1, 2, 6\}$ and $\{1, 12, 15\}$ separately at t_2 and t_3 . At each moment, the probability of identifying user A is $1/3$. However, if an attacker obtains the user's anonymous sets at the three moments and then performs an intersection operation, then the true identity of A can be obtained. In our proposed scheme, point 2 is selected as a historical adjacent position at time t_1 . At time t_2 , considering that point 2 remains closest to A and to block the attacker from speculating that A is located near point 2, we chose point 7 as the historical adjacent position according to the historical proximity selection method described in the scheme ($B_i \neq B_j$). Due to population density, there are more historical adjacent locations around the user, so the possibility of the users' real identity being exposed is lower.

In addition, it is more reasonable to utilize historical proximity locations instead of the randomly generated loca-

tions by the traditional dummy location technology. As shown in Figure 17, when the user requests LBSs at position 1 with the dummy location technology, it is possible that the pseudonym-location generated is in the lake. In this case the adversary can easily identify it with background knowledge and filter out other unreasonable locations, which definitely decreases the privacy degree. Fortunately, using the historical proximity locations proposed in this paper can address this problem.

Compared with most of the existing methods such as k -anonymity and dummy location, which have to report the user's true position to the service providers, our approach submits the historical proximity location to substitute the user's current location, which improves the privacy level.

5. Conclusion

This paper proposed a solution for location privacy protection in both snapshot queries and continuous queries. With historical proximity locations around the user submitted to the LSP for query instead of the true location, the user's location information is completely anonymous from the LSP in the whole process of requesting the LBSs, and thus a high privacy level is achieved. Compared with k -anonymity and the enhanced pseudonym selection scheme, our scheme is simple and feasible, and can achieve better query efficiency and higher query quality. Our proposal also provides a new solution for dealing with the problem of maintaining the equilibrium among the privacy level, query efficiency, and quality of services. However, in continuous queries, we have not succeeded in achieving sufficient anonymity protection level for a user's movement trajectory. It is still possible for an attacker to infer the general direction of the user's movement by analyzing the changes of the user's query range recorded in the LSP. This is a difficulty in the current techniques of trajectory privacy preservation, and it is also the focus of future research work.

Data Availability

No data were used to support this study.

Disclosure

X. Guo and W. Wang are co-first authors.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

X. Guo and W. Wang contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant number 61672297), the Key Research and Development Program of Jiangsu Province (grant number BE2017742), the Postgraduate Research & Practice Innovation Program of Jiangsu Province (grant number KYCX19_0908), and the Key Project on Anhui Provincial Natural Science Study by Colleges and Universities (grant numbers KJ2019A0579 and KJ2019A0554).

References

- [1] C. Yin, J. Xi, R. Sun, and J. Wang, "Location privacy protection based on differential privacy strategy for big data in industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3628–3636, 2017.
- [2] C. Ardagna, M. Cremonini, S. D. C. di Vimercati, and P. Samarati, "An obfuscation-based approach for protecting location privacy," *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 1, pp. 13–27, 2011.
- [3] S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatial K-anonymity driven privacy enhancement scheme in continuous location-based services," *Future Generation Computer Systems*, vol. 94, pp. 40–50, 2019.
- [4] Y. Dong and D. Pi, "Novel privacy-preserving algorithm based on frequent path for trajectory data publishing," *Knowledge-Based Systems*, vol. 148, pp. 55–65, 2018.
- [5] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st international conference on Mobile systems, applications and services*, pp. 31–42, 2003.
- [6] Y. Zhang and Q. Zhang, "A k-anonymous location privacy protection method of dummy based on geographical semantics," *International Journal of Network Security*, vol. 21, no. 6, pp. 937–946, 2019.
- [7] J. Wang, J. Luo, X. Liu, Y. Li, and S. Liu, "Improved Kalman filter based differentially private streaming data release in cognitive computing," *Future Generation Computer Systems*, vol. 98, pp. 541–549, 2019.
- [8] T. Nakagawa and H. Arai, "Personalized anonymization for set-valued data by partial suppression," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 1003–1010, IEEE, 2017.
- [9] G. Ghinita, K. Zhao, D. Papadias, and P. Kalnis, "A reciprocal framework for spatial K-anonymity," *Information Systems*, vol. 35, no. 3, pp. 299–314, 2010.
- [10] W. Zhang, M. Li, R. Tandon, and H. Li, "Online location trace privacy: an information theoretic approach," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 235–250, 2019.
- [11] G. Li, L. Li, J. Li, and Y. Li, "Network Voronoi diagram on uncertain objects for nearest neighbor queries," *Information Sciences*, vol. 301, pp. 241–261, 2015.
- [12] S. Tang, S. Liu, X. Huang, and Z. Liu, "Privacy-preserving location-based service protocols with flexible access," *International Journal of Computational Science and Engineering*, vol. 20, no. 3, pp. 412–423, 2019.
- [13] J. Ren, Y. Zhang, K. Zhang, and X. Shen, "Adaptive and channel-aware detection of selective forwarding attacks in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3718–3731, 2016.
- [14] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *International conference on pervasive computing*, pp. 152–170, Springer, Berlin, Heidelberg, 2005.
- [15] X. Li, E. Wang, W. Yang, and J. Ma, "DALP: a demand-aware location privacy protection scheme in continuous location-based services," *Concurrency and Computation: Practice and Experience*, vol. 28, no. 4, pp. 1219–1236, 2016.
- [16] J. Zhang, X. Zhao, and C. Ji, "A novel authenticated encryption scheme and its extension," *Information Sciences*, vol. 317, pp. 196–201, 2015.
- [17] T. Peng, Q. Liu, D. Meng, and G. Wang, "Collaborative trajectory privacy preserving scheme in location-based services," *Information Sciences*, vol. 387, pp. 165–179, 2017.
- [18] B. Wang, L. Zhang, and G. Zhang, "A fine granularity based user collaboration algorithm for location privacy protection," *PloS one*, vol. 14, no. 7, p. e0220278, 2019.
- [19] R. Shokri, G. Theodorakopoulos, P. Papadimitratos, E. Kazemi, and J.-P. Hubaux, "Hiding in the mobile crowd: location privacy through collaboration," *IEEE transactions on dependable and secure computing*, vol. 11, no. 3, pp. 266–279, 2014.
- [20] T. Wang, Y. Mei, W. Jia, X. Zheng, G. Wang, and M. Xie, "Edge-based differential privacy computing for sensor-cloud systems," *Journal of Parallel and Distributed Computing*, vol. 136, pp. 75–85, 2020.
- [21] Y. Tian, B. Song, M. Al Rodhaan et al., "A stochastic location privacy protection scheme for edge computing," *Mathematical Biosciences and Engineering*, vol. 17, no. 3, pp. 2636–2649, 2020.
- [22] S. Zhang, G. Wang, M. Z. A. Bhuiyan, and Q. Liu, "A dual privacy preserving scheme in continuous location-based services," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4191–4200, 2018.
- [23] S. Zhang, G. Wang, Q. Liu, and J. H. Abawayj, "A trajectory privacy-preserving scheme based on query exchange in mobile social networks," *Soft Computing*, vol. 22, no. 18, pp. 6121–6133, 2018.
- [24] T. Xu and Y. Cai, "Exploring historical location data for anonymity preservation in location-based services," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, pp. 1220–1228, Phoenix, AZ, USA, 2008.

- [25] C.-Y. Chow, M. F. Mokbel, and W. G. Aref, "Casper*," *ACM Transactions on Database Systems*, vol. 34, no. 4, pp. 1–48, 2009.
- [26] X. Li, M. Miao, H. Liu, J. Ma, and K. C. Li, "An incentive mechanism for K-anonymity in LBS privacy protection based on credit mechanism," *Soft Computing*, vol. 21, no. 14, pp. 3907–3917, 2017.

Research Article

An Efficient Searchable Public-Key Authenticated Encryption for Cloud-Assisted Medical Internet of Things

Tianyu Chi,¹ Baodong Qin ,^{1,2} and Dong Zheng^{1,3}

¹School of Cyberspace Security, Xi'an University of Posts & Telecommunications, Xi'an, Shaanxi, China

²State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China

³Westone Cryptologic Research Center, Beijing, China

Correspondence should be addressed to Baodong Qin; qinbaodong@foxmail.com

Received 27 March 2020; Revised 13 June 2020; Accepted 17 June 2020; Published 14 July 2020

Academic Editor: Huaqun Wang

Copyright © 2020 Tianyu Chi 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 recent years, it has become popular to upload patients' medical data to a third-party cloud server (TCS) for storage through medical Internet of things. It can reduce the local maintenance burden of the medical data and importantly improve accuracy in the medical treatment. As remote TCS cannot be fully trusted, medical data should be encrypted before uploading, to protect patients' privacy. However, encryption makes search capabilities difficult for patients and doctors. To address this issue, Huang et al. recently put forward the notion of Public-key Authenticated Encryption with Keyword Search (PAEKS) against inside keyword guessing attacks. However, the existing PAEKS schemes rely on time-consuming computation of parings. Moreover, some PAEKS schemes still have security issues in a multiuser setting. In this paper, we propose a new and efficient PAEKS scheme, which uses the idea of Diffie-Hellman key agreement to generate a shared secret key between each sender and receiver. The shared key will be used to encrypt keywords by the sender and to generate search trapdoors by the receiver. We prove that our scheme is semantically secure against inside keyword guessing attacks in a multiuser setting, under the oracle Diffie-Hellman assumption. Experimental results demonstrate that our PAEKS scheme is more efficient than that of previous ones, especially in terms of keyword searching time.

1. Introduction

In today's society, almost all medical service providers will use some form of electronic medical record system [1]. Specifically, medical Internet of things (MIoT) has become a new technology to gather data from patients by small wearable devices or implantable sensors. With the increasing number of medical data, the burden of hospital storage equipment is heavy, and it needs a professional person to maintain. If the hardware storage device is damaged and data is lost due to other force majeure factors, it will lead to very serious consequences. The most important way to solve this problem is to upload the data to the third-party cloud server (TCS). However, after the data is uploaded to the TCS, the patient's privacy will not be guaranteed. Once the cloud server managers or external malicious attackers steal the data, it will cause data leakage and other problems [2].

In order to solve the problem of data security, the best way is to encrypt the data and then upload the result to TCS. But when medical service providers want to retrieve the electronic medical records of patients, it becomes more difficult. First, doctors need to download all encrypted data to a local server and then decrypt it locally. After that, they can search for the desired results in the plaintext medical data. However, this process is very cumbersome and impractical for most applications. Due to the powerful cloud computing, medical institutions hope that the cloud server can complete the retrieval function instead of doing it themselves. But if the key is sent to the cloud server, the patient's private data still has the risk of exposure.

To address the above security issues, the conception of (symmetric-key) searchable encryption (SE) was proposed by Song et al. [3]. It is a powerful technology that allows the cloud server to search on encrypted data using some search

trapdoors generated by the local data users. In 2004, Boneh et al. [4] proposed a public-key version of SE, namely, Public-key Encryption with Keyword Search (PEKS). This scheme embeds keywords in public-key encryption and is very suitable for scenarios of a multiuser data sharing setting, e.g., medical data sharing. There are three parties in the PEKS scheme: cloud server, data sender, and data receiver. The sender (e.g., patient) has a lot of privacy files F_i and wants to share them with the receiver (e.g., doctor). First, the sender extracts the keyword w_i from each file F_i , encrypts the keyword with the PEKS scheme, and then encrypts each file with other encryption schemes (not necessarily the same as the PEKS scheme). Let the keyword cipher text be C_{w_i} . The sender uploads all cipher texts to the TCS. In order to search whether there is a document containing the keyword w in the encrypted document, the receiver generates a search trapdoor T_w of the keyword w and sends the trapdoor to the cloud server. After the server receives T_w , it checks whether each keyword cipher text matches with the search trapdoor. If so, it indicates that the corresponding encrypted document must contain the desired keyword. After that, the results are returned to the receiver, and the receiver can get the required plaintext data by decrypting the encrypted documents.

As mentioned in Figure 1, we will apply searchable encryption to telemedicine services, where the patient is the sender and the medical service provider is the receiver. Each patient can encrypt and upload their own electronic medical record to the cloud server. When the patient wants to see a doctor remotely, the doctor can retrieve the medical data information related to some disease on the third-party cloud server according to the keyword information of the patient. In this process, doctors will only get data related to a certain disease and will not expose other information (such as name) of patients.

However, PEKS inherently has a disadvantage to resist against keyword guessing attacks (KGA). Ideally, a keyword space can be considered infinite. In practice, however, this is not the case. In real life, users often use a limited number of keywords because of their living habits, which leads to the transformation of the original polynomial space into an affixed and low-entropy space. In this case, the adversary can guess the keywords contained in the searching trapdoor as follows: First, the adversary guesses all the keyword spaces of the user and then generates keyword cipher text one by one. The adversary checks the trapdoor requested by the user one by one with keyword cipher texts generated by itself. If there is coincidentally the same situation, the adversary can obtain the keyword information retrieved by the user, thus exposing the privacy of the user. This kind of attack can be easily mounted by the cloud server, as the cloud server has users' searching trapdoors. Such attack is often called inside keyword guessing attacks (IKGA).

To resist against KGA is very challenging. Recently, many methods [5–11] were proposed to prevent KGA on PEKS schemes; however, most of them were later proven insecure [12–15]. In 2017, Huang and Li [16] proposed a new primitive, namely, Public-key Authenticated Encryption with Keyword Search (PAEKS), to solve the problem of inside KGA. In PAEKS, the data sender not only encrypts a key-

word but also authenticates it, so that a search trapdoor can only match with the corresponding data sender. PAEKS is also applicable to cloud-assisted MIoT, as in general, the doctor just searches on a designated patient's medical data. However, the proposed concrete PAEKS scheme still has some security issues [17–19]. In particular, Noroozi and Eslami [18] pointed out that it cannot handle multiuser settings and provided an improvement security model for PAEKS in a multiuser setting.

1.1. Our Contribution. In this paper, we research on new and efficient construction of PAEKS schemes in a multiuser setting for cloud-assisted MIoT. Our main contributions are as follows:

- (i) We observe that in PAEKS, both the data sender and data receiver hold a pair of public/secret keys. If they can compute a shared key without any interaction, then the shared key can be viewed as a secret key of a symmetric searchable encryption scheme. Inspired by this, we propose an efficient PAEKS scheme, which involves the (noninteractive) Diffie-Hellman key exchange scheme to compute the shared key and Song et al.'s SSE scheme to encrypt keywords. It removes the usage of time-consuming operation of pairing in previous PAEKS schemes
- (ii) We show that our scheme is semantically secure against IKGA in a multiuser setting under the oracle Diffie-Hellman assumption [20]. Specially, it satisfies both cipher text indistinguishability and trapdoor indistinguishability
- (iii) We compare our scheme with some related PAEKS scheme in terms of security and computation efficiency and also do some experiments to demonstrate the efficiency of our schemes for protecting the privacy of cloud-assisted MIoT data. Experiment results show that our scheme is more efficient than that of previous ones, especially in terms of keyword searching time

1.2. Paper Organization. In the next section, we will briefly introduce some cryptographic primitives. Our main construction of the PAEKS scheme and its security proof are given in Section 3. In Section 4, we compare the efficiency of our scheme with that of other related PAEKS schemes. Finally, we summarize the paper in Section 5.

2. Preliminaries

In this section, we recall some basic conceptions of cryptographic primitives that will be used in this paper, including cyclic group, hardness assumption, pseudorandom functions, syntax of PAEKS, and its security model.

2.1. Cyclic Group. Let G be a group with order p . We say that G is a cyclic group, if the group G can be generated by a single element $g \in G$. That is, every element $h \in G$ has the form $h = g^x$ for some exponent $x \in \mathbb{Z}_p$. We call g to be a

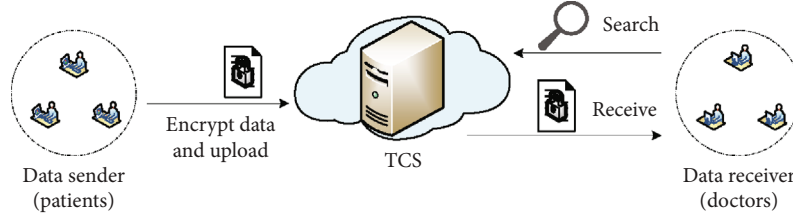


FIGURE 1: Telemedicine service based on searchable encryption.

generator of the group. In our scheme, we use a cyclic group with a prime order; i.e., p is a prime. In this case, any group element except the identity will be a generator.

2.2. Oracle Diffie-Hellman (ODH) Problem [20]. Let G be a cyclic group with prime order p and a generator g . Let H be a hash function from G to some n -bit length space $\{0, 1\}^n$. The ODH problem states that given a tuple (g, g^x, g^y, T) and an oracle $\mathcal{O}_{g^x}(\cdot)$, to decide whether T is $H(g^{xy})$ or a random string from $\{0, 1\}^n$, here, x and y are randomly chosen from \mathbb{Z}_p , and the oracle returns $H(h^x)$ for each $h \in G$. Let \mathcal{A} be any probabilistic polynomial time (PPT) algorithm. We say that \mathcal{A} breaks the ODH problem over group G and H with advantage at most ϵ_{odh} , if

$$\left| \Pr \left[\mathcal{A}^{\mathcal{O}_{g^x}(\cdot)}(g, g^x, g^y, H(g^{xy})) = 1 \right] - \Pr \left[\mathcal{A}^{\mathcal{O}_{g^x}(\cdot)}(g, g^x, g^y, T) = 1 \right] \right| < \epsilon_{\text{odh}}. \quad (1)$$

Definition 1 (ODH assumption). We say that the ODH assumption holds over group G and H , if for any PPT algorithm \mathcal{A} , its advantage ϵ_{odh} in solving the ODH problem is negligible in κ (the bit length of p).

2.3. Pseudorandom Functions (PRFs). A pseudorandom function is a family of functions such that for a random choice from the family, its input/output behavior is computationally indistinguishable from that of a random function. A formal definition of PRFs is given below.

Definition 2 (PRFs). Let $f : \mathcal{K} \times \mathcal{X} \rightarrow \mathcal{Y}$ be a family of functions indexed with key space \mathcal{K} from \mathcal{X} to \mathcal{Y} . We say that f is an ϵ_f -securePRF_s if

- (1) Given a key $k \in \mathcal{K}$ and an input $x \in \mathcal{X}$, there is an efficient algorithm to compute the output $y = f_k(x)$
- (2) For any PPT algorithm \mathcal{A} that makes at most polynomial number of oracle queries, the following advantage is at most ϵ_f :

where $F = \{f : \mathcal{X} \rightarrow \mathcal{Y}\}$ and the oracles are given an input $x \in \mathcal{X}$ and output the corresponding image of the function.

$$\left| \Pr \left[\mathcal{A}(\mathcal{O}_{g^x}(\cdot)) = 1 : k \leftarrow \mathcal{K} \right] - \Pr \left[\mathcal{A}(\mathcal{O}_{g^x}(\cdot)) = 1 : f \leftarrow F \right] \right| \leq \epsilon_f, \quad (2)$$

The above definition indicates that, given any polynomial number of valid input/output pairs $(x_i, f_k(x_i))$, no PPT adversary can predicate $f_k(x)$ for a new and distinct input x . Specifically, $f_k(x)$ is computationally indistinguishable from a random $y \in \mathcal{Y}$.

2.4. PAEKS and Security Model. The notion of Public-key Authenticated Encryption with Keyword Search (PAEKS) was first proposed in [16] to protect the privacy of a keyword against inside keyword guessing attacks. It involves the public/secret key pair into the cipher text to prevent keyword guessing attacks by the insider server. We first recall its definition.

Definition 3 (syntax of PAEKS). A PAEKS scheme consists of the following six PPT algorithms:

- (i) *Setup* (λ). This is the global parameter generation algorithm. It takes the security parameter λ as input and outputs global system parameter Param
- (ii) *KeyGen_S* (Param). This is the sender's key generation algorithm. It takes the global system parameter Param as input and outputs a public/secret key pair (pk_S, sk_S)
- (iii) *KeyGen_R* (Param). This is the receiver's key generation algorithm. It takes the global system parameter Param as input and outputs a public/secret key pair (pk_R, sk_R)
- (iv) *PAEKS* (sk_S, pk_R, w). This is the keyword encryption algorithm performed by the sender. It takes the sender's secret key sk_S , the receiver's public key pk_R , and a keyword w as input and outputs a PAEKS cipher text C of the keyword w
- (v) *Trapdoor* (sk_R, pk_S, w). This is the trapdoor generation algorithm performed by the receiver. It takes the receiver's secret key sk_R , the sender's public key pk_S , and a keyword w as input and outputs a trapdoor T_w
- (vi) *Test* (T_w, C, pk_S, pk_R). This is the test algorithm performed by the cloud server. It takes a trapdoor T_w , a PAEKS cipher text C , the sender's public key pk_S , and the receiver's public key pk_R as input and outputs 1 if C and T_w contain the same keyword and 0 otherwise

Next, we recall the improved security model for PAEKS in a multiuser setting by Noroozi and Eslami [18]. It includes trapdoor indistinguishability (TI) and cipher text indistinguishability (CI). Both of them are described through games played between an adversary \mathcal{A} and the challenger \mathcal{C} .

Definition 4 (TI security game). The TI security game is described as follows:

- (i) *Initialization*. Given a security parameter λ , the challenger generates the global system parameter. Then, the challenger generates the receiver's public/secret keys (pk_R, sk_R) and the sender's public/secret keys (pk_S, sk_S) . It executes the adversary \mathcal{A} on input $(Param, pk_S, pk_R)$
- (ii) *Phase 1*. The adversary \mathcal{A} is permitted to adaptively query the following two oracles polynomial times:
 - (a) *Cipher Text Oracle* $\mathcal{O}_C(w, pk)$. Given a keyword w and a public key pk , the challenger computes the cipher text C by running the algorithm PAEKS (sk_S, pk, w) and returns the cipher text to \mathcal{A}
 - (b) *Trapdoor Oracle* $\mathcal{O}_T(w, pk)$. Given a keyword w and a public key pk , the challenger computes the trapdoor T_w by running the algorithm trapdoor (sk_R, pk, w) and returns the trapdoor to \mathcal{A}
- (iii) *Challenge*. When phase 1 ends, the adversary \mathcal{A} outputs two challenge keywords w_0^* and w_1^* , which have not been queried to the oracles $\mathcal{O}_C(\cdot, pk_R)$ and $\mathcal{O}_T(\cdot, pk_S)$ before. Now, the challenger chooses a random bit $b \in \{0, 1\}$, computes the $T_{w_b^*} \leftarrow \text{trapdoor}(sk_R, pk_S, w_b^*)$, and returns it to the adversary \mathcal{A}
- (iv) *Phase 2*. In this phase, the adversary \mathcal{A} can continue to access the oracles, with the restriction that neither w_0^* nor w_1^* could be queried to the oracles $\mathcal{O}_C(\cdot, pk_R)$ and $\mathcal{O}_T(\cdot, pk_S)$
- (v) *Guessing*. Finally, the adversary \mathcal{A} outputs a bit b' as the guess of b . If $b' = b$, we say that \mathcal{A} wins the game

We define \mathcal{A} 's advantage in breaking the TI security of PAEKS as

$$\text{Adv}_{\mathcal{A}}^{\text{TI}}(\lambda) = \left| \Pr [b' = b] - \frac{1}{2} \right|. \quad (3)$$

Definition 5 (CI security game). Similarly, the CI security game can be described as follows:

- (i) *Initialization*. Given a security parameter λ , the challenger generates the global system parameter $Param$.

Then, the challenger generates the receiver's public/secret keys (pk_R, sk_R) and the sender's public/secret keys (pk_S, sk_S) . It executes the adversary \mathcal{A} on input $(Param, pk_S, pk_R)$

- (ii) *Phase 1*. The adversary \mathcal{A} is allowed to adaptively query the following two oracles polynomial times:
 - (a) *Cipher Text Oracle* $\mathcal{O}_C(w, pk)$. Given a keyword w and a public key pk , the challenger computes the cipher text C by running the algorithm PAEKS (sk_S, pk_R, w) and returns the cipher text to \mathcal{A}
 - (b) *Trapdoor Oracle* $\mathcal{O}_T(w, pk)$. Given a keyword w and a public key pk , the challenger computes the trapdoor T_w by running the algorithm trapdoor (sk_R, pk_S, w) and returns the trapdoor to \mathcal{A}
- (iii) *Challenge*. When phase 1 ends, the adversary \mathcal{A} outputs two challenge keywords w_0^* and w_1^* , which have not been queried to the oracles $\mathcal{O}_C(\cdot, pk_R)$ and $\mathcal{O}_T(\cdot, pk_S)$ before. Now, the challenger chooses a random bit $b \in \{0, 1\}$, computes the $C_{w_b^*} \leftarrow \text{PAEKS}(sk_S, pk_R, w_b^*)$, and returns it to the adversary \mathcal{A}
- (iv) *Phase 2*. In this phase, the adversary \mathcal{A} can continue to access the oracles, with the restriction that neither w_0^* nor w_1^* could be queried to the oracles $\mathcal{O}_C(\cdot, pk_R)$ and $\mathcal{O}_T(\cdot, pk_S)$
- (v) *Guessing*. Finally, the adversary \mathcal{A} outputs a bit b' as the guess of b . If $b' = b$, we say that \mathcal{A} wins the game

We define \mathcal{A} 's advantage in breaking the CI security of PAEKS as

$$\text{Adv}_{\mathcal{A}}^{\text{CI}}(\lambda) = \left| \Pr [b' = b] - \frac{1}{2} \right|. \quad (4)$$

If for any PPT adversary \mathcal{A} , both $\text{Adv}_{\mathcal{A}}^{\text{TI}}(\lambda)$ and $\text{Adv}_{\mathcal{A}}^{\text{CI}}(\lambda)$ are negligible in the security parameter λ ; we say that the PAEKS is semantically secure against inside keyword guessing attacks.

3. Our PAEKS Scheme

In this section, we introduce a PAEKS scheme for an electronic medical record system. The system framework is given in Figure 2.

3.1. The Construction. Our PAEKS scheme is described as follows:

- (i) *Setup* (λ). Select a cyclic group G with prime order p and a random generator g of G . Select three pseudorandom functions: $E : \mathcal{K}' \times \mathcal{W} \rightarrow \{0, 1\}^n$, $f : \mathcal{K}'' \times$

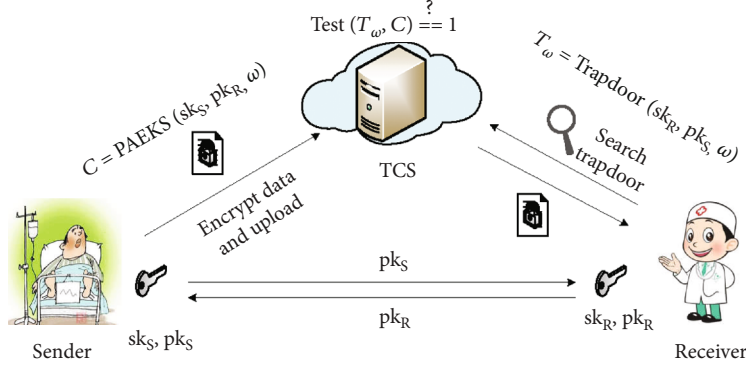


FIGURE 2: Our PAEKS system framework.

$\{0, 1\}^n \rightarrow \mathcal{K}$, and $F : \mathcal{K}' \times \{0, 1\}^{m-n} \rightarrow \{0, 1\}^m$, where \mathcal{K}' , \mathcal{K}'' , and \mathcal{K} are the key spaces of the three PRFs, respectively, and \mathcal{W} is the keyword space. Let H be a hash function, defined as $H : G \rightarrow \mathcal{K}' \times \mathcal{K}''$. Finally, return $\text{Param} = (G, g, p, E, f, F)$

- (ii) $\text{KeyGen}_S(\text{Param})$. Randomly select $x \leftarrow Z_p$, and set $\text{pk}_S := g^x$ and $\text{sk}_S := x$. Return sk_S and pk_S
- (iii) $\text{KeyGen}_R(\text{Param})$. Randomly select $y \leftarrow Z_p$, and set $\text{pk}_R := g^y$ and $\text{sk}_R := y$. Return sk_R and pk_R
- (iv) $\text{PAEKS}(\text{sk}_S, \text{pk}_R, w)$. To encrypt a keyword $w \in \mathcal{W}$, do the following:
 - (a) Compute the keys $k' \| k'' = H(\text{pk}_R^{\text{sk}_S}) = H(g^{xy})$
 - (b) Compute $X = E_{k'}(w)$ and $k = f_{k''}(X)$
 - (c) Select a random string $S \in \{0, 1\}^{n-m}$ and set $U = S \| F_k(S)$
 - (d) Set $C = X \oplus U$
 - (e) Finally, return C
- (v) $\text{Trapdoor}(\text{sk}_R, \text{pk}_S, w)$. Compute $k' \| k'' = H(\text{pk}_R^{\text{sk}_S}) = H(g^{xy})$, $X = E_{k'}(w)$ and $k = f_{k''}(X)$. Return the trapdoor $T_w = X \| k''$
- (vi) $\text{Test}(T_w, C)$. Compute $U = C \oplus X$ and parse it as $S \| T$. If $T = F_k(S)$ holds, return 1; otherwise, return 0

3.1.1. Correctness. Let the receiver's key pair be $(\text{pk}_R, \text{sk}_R) = (g^y, y)$ and the sender's key pair be $(\text{pk}_S, \text{sk}_S) = (g^x, x)$. Then, the key $k' \| k'' = H(g^{xy})$ can be generated by each other. Let C be a cipher text of keyword w generated by the sender and T_w be the corresponding search trapdoor generated by the receiver. According to the keyword encryption algorithm, there must exist two strings X and U and a random string $S \in \{0, 1\}^{n-m}$ such that $C = X \oplus U$, $X = E_{k'}(w)$, and $U = S \| F_k(S)$, where $k = f_{k''}(X)$. For a right trapdoor of

keyword w , it should be in the form $T_w = X \| k$, where $X = E_{k'}(w)$ and $k = f_{k''}(X)$. So, $X \oplus C = U$. Let S be the first $n - m$ bits of U and T be the last m bits. Clearly, $T = F_k(S)$ will hold. Thus, for the same keyword, the cipher text will match with the corresponding trapdoor.

Fixing a cipher text C' of a distinct keyword $w' \neq w$, we have $C' = X' \oplus U'$, for some $X' = E_{k'}(w')$. Since E is a pseudorandom function, then $X = E_{k'}(w)$ is a random string over $\{0, 1\}^n$ with probability at least $1 - \epsilon_E$. In this case, $X \oplus C'$ will be a random string. Since F is also a pseudorandom function, for a random string $S' \| T'$, the equation $F_k(S') = T'$ holds with probability at most $\epsilon_E + (1/2^n)$. Thus, the cipher text C' matches with the search trapdoor with a negligible probability. So our PAEKS scheme satisfies the correctness.

3.2. Security Proof. In this section, we prove that our PAEKS scheme satisfies both trapdoor indistinguishability and cipher text indistinguishability. Its trapdoor indistinguishability follows from the theorem below.

Theorem 6. *If the oracle Diffie-Hellman assumption holds and E is a pseudorandom function, then our PAEKS scheme achieves trapdoor indistinguishability. Specifically, for any PPT adversary \mathcal{A} , we have*

$$\text{Adv}_{\mathcal{A}}^{\text{TI}}(\lambda) \leq \epsilon_{\text{odh}} + \epsilon_E, \quad (5)$$

where ϵ_{odh} and ϵ_E are the advantages to break the ODH assumption and the pseudorandomness of the PRF E .

Proof. Let \mathcal{A} be any PPT adversary that aims to break the security of trapdoor indistinguishability of our PAEKS scheme. We prove Theorem 6 by a sequence of games. Let Suc_i denote the event that \mathcal{A} succeeds (i.e., $b' = b$) in the i -th game.

Game 0. This is the original trapdoor in a distinguishability game as defined in Definition 4. In this game, the challenger generates two public/secret key pairs $(\text{pk}_S, \text{sk}_S)$ and $(\text{pk}_R, \text{sk}_R)$ for the sender and the receiver, respectively, and gives the public keys to \mathcal{A} . In addition, the adversary can adaptively issue queries to the trapdoor oracle $\mathcal{O}_T(\cdot, \cdot)$ and cipher text oracle $\mathcal{O}_C(\cdot, \cdot)$ with any keyword $w \in \mathcal{W}$ and public key pk . But, for the two challenge keywords w_0^* and w_1^* ,

the adversary cannot submit them to the oracles $\mathcal{O}_T(\cdot, \text{pk}_S)$ and $\mathcal{O}_C(\cdot, \text{pk}_R)$. Let $T_{w_b^*}$ denote the challenge trapdoor of w_b^* , where $b \leftarrow \{0, 1\}$. Let b' denote the guess of b by \mathcal{A} . So, \mathcal{A} 's advantage in this game is

$$\text{Adv}_{\mathcal{A}}^{\text{PI}}(\lambda) = \left| \Pr [\text{Suc}_0] - \frac{1}{2} \right|, \quad (6)$$

Game 1. This game is the same as the previous game with the exception of $k' \| k''$ being sampled from $\mathcal{K}' \times \mathcal{K}''$ uniformly at random. Recall that in the previous game, the challenger computes $k' \| k''$ by $H(\text{pk}_S^{\text{sk}_R})$ (namely, $H(\text{pk}_R^{\text{sk}_S})$) according to the keyword encryption algorithm (namely, trapdoor generation algorithm). We now prove that

$$|\Pr [\text{Suc}_1] - \Pr [\text{Suc}_0]| \leq \epsilon_{\text{odh}}, \quad (7)$$

Given an instance of the oracle Diffie-Hellman problem (g, g^x, g^y, K) , where $K = H(g^{xy})$ or K is a random string from $\mathcal{K}' \times \mathcal{K}''$, we construct an algorithm \mathcal{B} to solve it using \mathcal{A} as a subroutine. \mathcal{B} sets $\text{pk}_S := g^x$ and $\text{pk}_R := g^y$ and gives them to \mathcal{A} . The corresponding secret keys are implicitly set to be x and y , respectively. In addition, \mathcal{B} chooses the other system parameters, including E, f, F , by itself. Parse K as $k' \| k''$. When \mathcal{A} issues queries to the oracles $\mathcal{O}_T(\cdot, \text{pk})$ and $\mathcal{O}_C(\cdot, \text{pk})$ with $\text{pk} \notin \{\text{pk}_S, \text{pk}_R\}$, \mathcal{B} involves the oracle $\mathcal{O}_{g^y}(\text{pk})$ or $\mathcal{O}_{g^x}(\text{pk})$ to obtain the shared key $k' \| k''$. When \mathcal{A} issues queries to the oracles $\mathcal{O}_T(\cdot, \text{pk}_S)$ and $\mathcal{O}_C(\cdot, \text{pk}_R)$, \mathcal{B} uses $k' \| k''$ to generate cipher texts and trapdoors. For example, for a keyword w , \mathcal{B} computes the cipher text C as follows:

- (1) Compute $X = E_{k'}(w)$ and $k = f_{k''}(X)$
- (2) Select a random string $S \in \{0, 1\}^{n-m}$ and set $U = S \| F_k(S)$
- (3) Set $C = X \oplus U$

Given two challenge keywords w_0^* and w_1^* , \mathcal{B} computes the challenge trapdoor $T_{w_b^*}$ as follows:

- (1) Choose a random bit $b \in \{0, 1\}$
- (2) Compute $X^* = E_{k'}(w_b^*)$ and $k^* = f_{k''}(X^*)$
- (3) Set $T_{w_b^*} = X^* \| k^*$

Finally, \mathcal{A} outputs a bit b' as a guess of b . If $b' = b$, \mathcal{B} outputs 1; otherwise, \mathcal{B} outputs 0.

Clearly, if $K = H(g^{xy})$, the above game is identical to Game 0. Otherwise, it is identical to Game 1. So,

$$\begin{aligned} \Pr [\mathcal{B}(g^x, g^y, H(g^{xy})) = 1 : x, y \leftarrow \mathbb{Z}_p] &= \Pr [\text{Suc}_0], \\ \Pr [\mathcal{B}(g^x, g^y, K) = 1 : x, y \leftarrow \mathbb{Z}_p, K \leftarrow \mathcal{K}' \times \mathcal{K}''] &= \Pr [\text{Suc}_1]. \end{aligned} \quad (8)$$

This proves the result of Equation (7).

Game 2. This game is identical to the previous game with the exception of X^* being sampled randomly from $\{0, 1\}^n$. Assuming that E is a pseudorandom function, we have

$$|\Pr [\text{Suc}_2] - \Pr [\text{Suc}_1]| \leq \epsilon_E. \quad (9)$$

We now prove Equation (9). Given a challenge pseudorandom function E , we construct an algorithm \mathcal{B} to break its pseudorandomness using \mathcal{A} as a subroutine. \mathcal{B} chooses the system parameter, the sender and receiver's public/secret key pairs, as in the previous game, with the exception of E being provided by its own challenger. Specifically, the random string $k' \in \mathcal{K}'$ is chosen by \mathcal{B} itself, but k' is implicitly defined by the secret key of the challenge pseudorandom function E . Next, we show how \mathcal{B} answers \mathcal{A} 's queries of cipher texts and trapdoors with (w, pk_R) or (w, pk_S) , respectively. For a keyword w , \mathcal{B} computes its cipher text as follows:

- (1) Query the challenger of E with w to obtain the result $X = E_{k'}(w)$
- (2) Compute $k = f_{k''}(X)$
- (3) Select a random string $S \in \mathcal{S}$ and compute $U = S \| F_k(S)$
- (4) Set $C = X \oplus U$

\mathcal{B} computes its trapdoor as follows:

- (1) Submit w to its own challenger to obtain the result $X = E_{k'}(w)$
- (2) Compute $k = f_{k''}(X)$
- (3) Set $T_w = X \| k$

When \mathcal{A} submits two challenge keywords w_0^* and w_1^* , \mathcal{B} picks a random bit b and sends w_b^* to the oracle of PRF E for challenging. The PRF challenger will return the challenge PRF value X^* to \mathcal{B} , which may be $E_{k'}(w_b^*)$ or a random value. \mathcal{B} then computes $k^* = f_{k''}(X^*)$ and returns $T_{w_b^*} = X^* \| k^*$ to the adversary. Finally, \mathcal{A} outputs a guess bit b' . If $b' = b$, \mathcal{B} outputs 1; otherwise, it outputs 0.

From the above analysis, it is clear that if $X^* = E_{k'}(w_b^*)$, \mathcal{B} actually simulates an environment of Game 1 for the adversary \mathcal{A} . If X^* is random, the simulated environment is identical to Game 2. Thus, if \mathcal{A} 's success probability between Game 1 and Game 2 has difference $\Pr [\text{Suc}_2] - \Pr [\text{Suc}_1]$, then \mathcal{B} can distinguish $X^* = E_{k'}(w_b^*)$ from a random one with the same advantage. This computes the proof of Equation (9).

Note that in Game 2, the challenge trapdoor is independent of the two challenge keywords. So, the adversary has no success advantage in this game, i.e.,

$$\Pr [\text{Suc}_2] = \frac{1}{2}. \quad (10)$$

TABLE 1: Efficiency comparison.

Schemes	Encryption	Trapdoor	Test	IKGA	Multiusers
Boneh et al. [4]	$3E + H_1 + H_2 + P$	$E + H_1$	P	No	Yes
Huang and Li [16]	$3E + H_1$	$E + H_1 + P$	$2P$	Yes	No
Noroozi and Eslami [18]	$3E + H_1$	$E + H_1 + P$	$2P$	Yes	Yes
Qin et al. [19]	$3E + H_1 + H_2 + P$	$2E + H_1$	P	Yes	Unknown
Ours	$E + H_2 + 3F$	$E + H_2 + 2F$	F	Yes	Yes

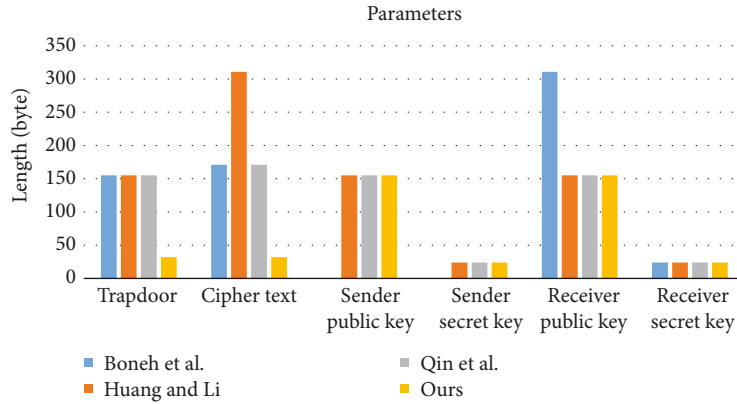


FIGURE 3: Comparison of parameter sizes.

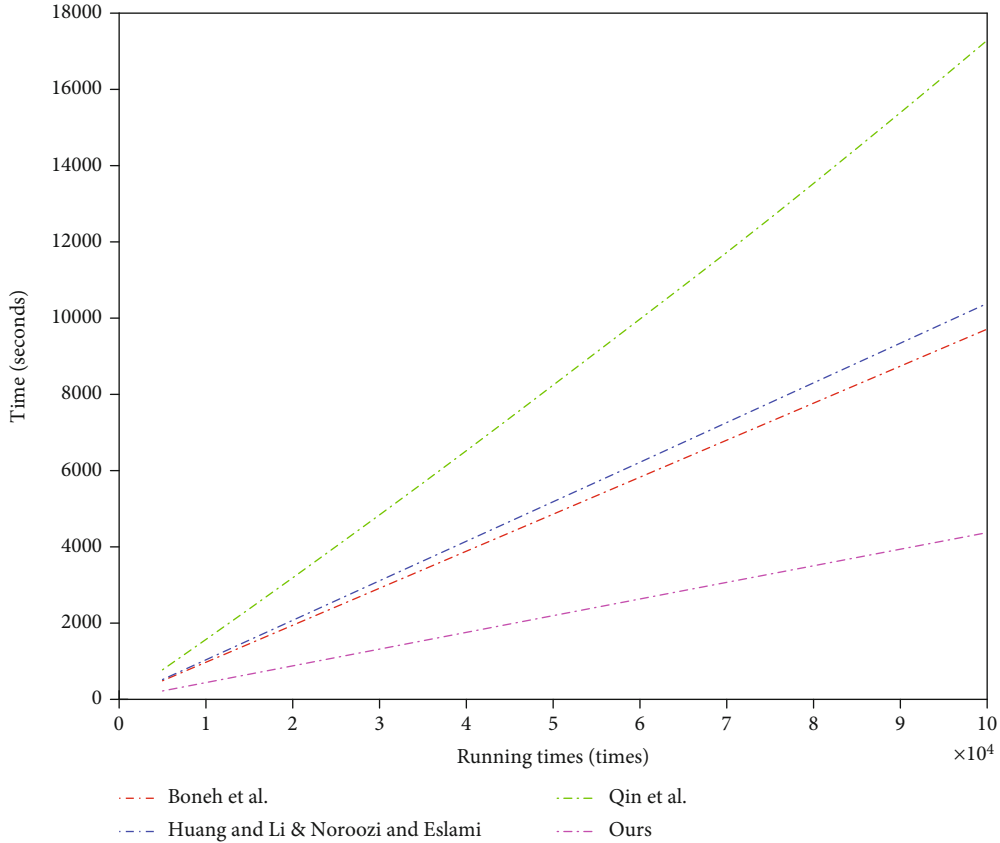


FIGURE 4: Running time of the encryption algorithm.

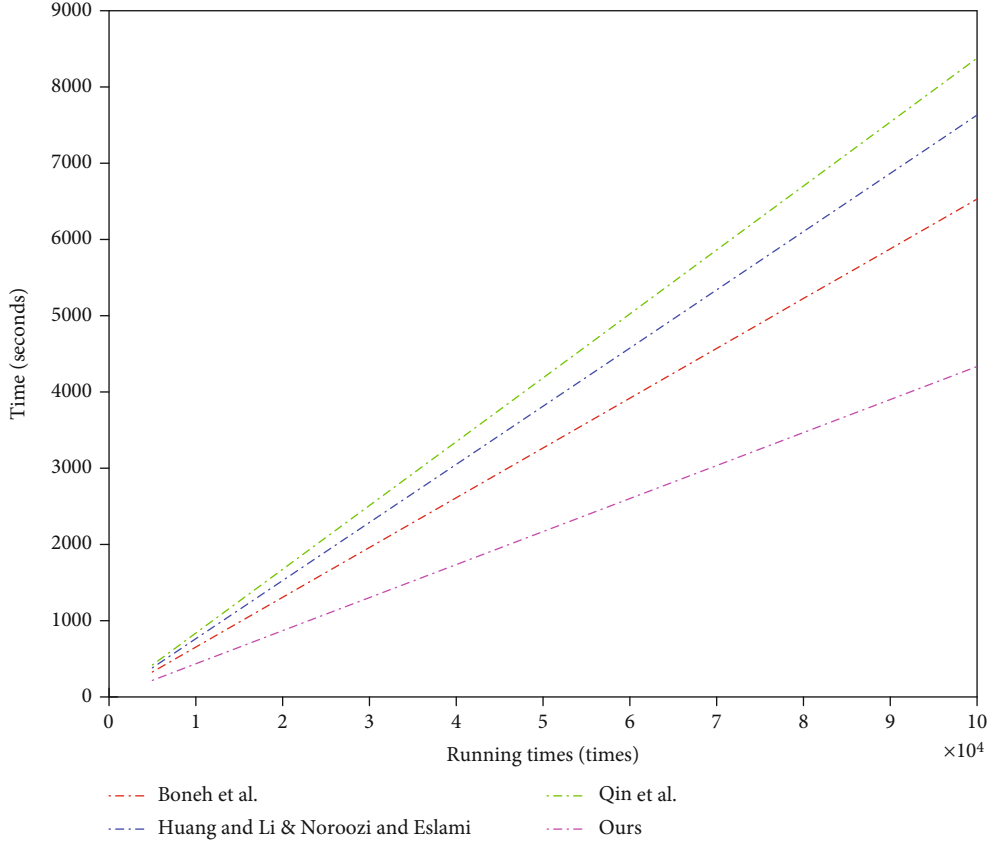


FIGURE 5: Running time of the trapdoor algorithm.

Taking Equations (6) to (10) together, it follows that

$$\text{Adv}_{\mathcal{A}}^{\text{TI}}(\lambda) \leq \epsilon_{\text{odh}} + \epsilon_E. \quad (11)$$

This completes the proof of Theorem 6.

The cipher text indistinguishability of our PAEKS scheme follows from the theorem below.

Theorem 7. *If the oracle Diffie-Hellman assumption holds and f, F are pseudorandom functions, then our PAEKS scheme achieves cipher text indistinguishability. Specifically, for any PPT adversary \mathcal{A} , we have*

$$\text{Adv}_{\mathcal{A}}^{\text{CI}}(\lambda) \leq \epsilon_{\text{odh}} + \epsilon_f + \epsilon_F, \quad (12)$$

where ϵ_{odh} , ϵ_f , and ϵ_F are the advantages to break the ODH assumption and the pseudorandomness of the PRFs f and F , respectively.

Proof. Similar to the proof of Theorem 6, we prove the above theorem also via a sequence of games. In each game, \mathcal{A} is a PPT adversary, aiming to break the cipher text indistinguishability of our PAEKS scheme. b is the challenge random bit, selected by the challenger, and b' is \mathcal{A} 's guess bit. We denote the event that $b' = b$ in each game as Suc_i .

Game 0. This is the original cipher text indistinguishability game as defined in Definition 5. So,

$$\text{Adv}_{\mathcal{A}}^{\text{CI}}(\lambda) = \left| \Pr [\text{Suc}_0] - \frac{1}{2} \right|. \quad (13)$$

Game 1. This game is the same as Game 0, except that the value $k' \| k''$ is chosen randomly from $\mathcal{K}' \times \mathcal{K}''$. Under the ODH assumption, these two games are computationally indistinguishable, i.e.,

$$|\Pr [\text{Suc}_1] - \Pr [\text{Suc}_0]| \leq \epsilon_{\text{odh}}. \quad (14)$$

The proof of the above equation is similar to that of Equation (7); we omit it here.

Game 2. This game is identical to Game 1, except the following modification to the challenge cipher text. Suppose that w_b^* is the challenge keyword and $X^* = E_{k'}(w_b^*)$ is the corresponding internal value of the cipher text. In this game, k^* is selected randomly from \mathcal{K} , instead of being computed via $k^* = f_{k''}(X^*)$. Note that, for normal keyword cipher text, k is still computed from $f_{k''}(X)$. Under the assumption that f is a pseudorandom function, these two games are computationally indistinguishable. Specially, we have

$$|\Pr [\text{Suc}_2] - \Pr [\text{Suc}_1]| \leq \epsilon_f. \quad (15)$$

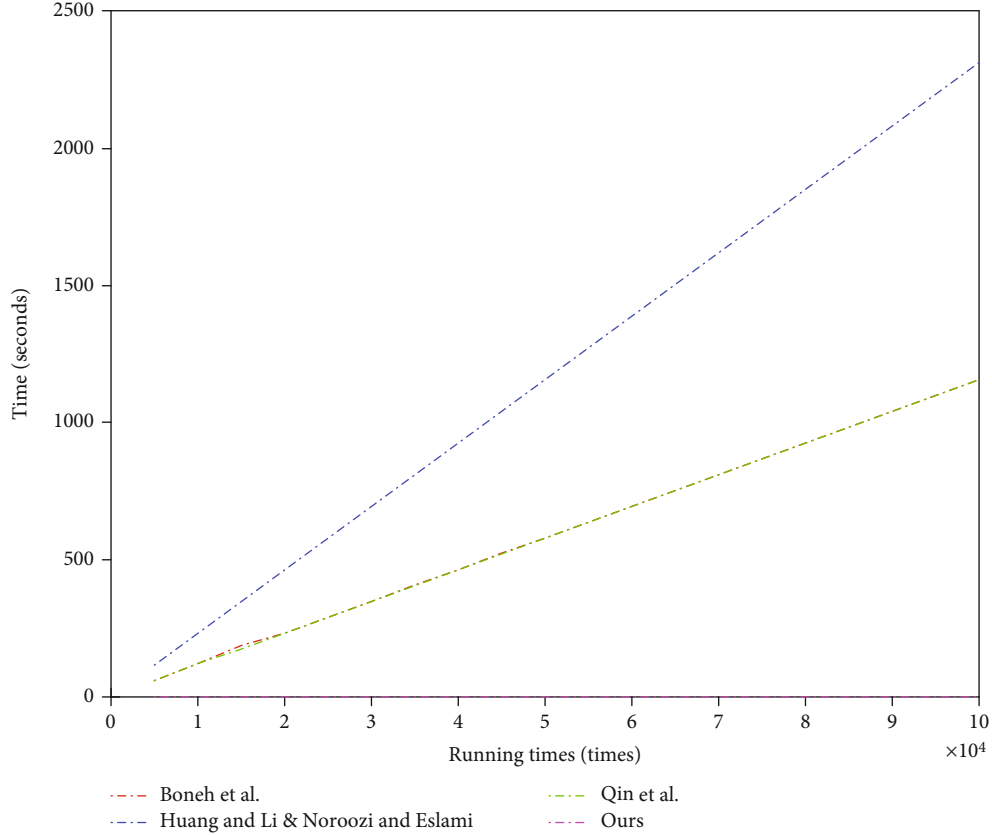


FIGURE 6: Running time of the test algorithm.

The proof of the above equation is similar to that of Equation (9); we omit it here.

Game 3. In this game, we replace the challenge value $U^* = S^* \parallel F_{k^*}(S^*)$ with a random string $U^* \leftarrow \{0, 1\}^n$. Recall that, in this game, k^* is sampled uniformly from \mathcal{K} . By the pseudorandomness of PRF F , $S^* \parallel F_{k^*}(S^*)$ is computationally indistinguishable from a random n -bit string. Similarly, we can prove that

$$|\Pr [\text{Suc}_3] - \Pr [\text{Suc}_2]| \leq \epsilon_F. \quad (16)$$

In Game 3, U^* is random and is independent of the challenge keywords. So, the adversary has no advantage in this game, i.e.,

$$\Pr [\text{Suc}_3] = \frac{1}{2}. \quad (17)$$

Taking Equations (13) to (17) together, we complete the proof of Theorem 7.

From Theorems 6 and 7, we conclude that our PAEKS scheme is semantically secure against inside keyword guessing attack assuming that the ODH problem is hard and E, f, F are PRFs.

4. Experiments and Efficiency Comparison

In this section, we analyze the efficiency of our PAEKS scheme and compare it with some other related schemes, including Boneh et al.'s PEKS scheme [4] and PAEKS schemes of [16, 18, 19]. Except our scheme, all the others are designed in bilinear groups. That is, besides group G , there are another group G_T and a bilinear map e defined from $G \times G$ to G_T .

Table 1 demonstrates the theoretical result of efficiency comparison in terms of keyword encryption, trapdoor generation, testing, and two security properties. In the table, we use symbols “ E ” and “ P ” to denote the evaluation of a modular exponentiation and a bilinear pairing, respectively. “ H_1 ” denotes a special hash function that maps an arbitrary string to a group element, while “ H_2 ” denotes a traditional hash function, e.g., MD5. We denote the pseudorandom function as “ F .”

Figure 3 shows the length of each parameter in different PEKS/PAEKS schemes. With the exception of Boneh et al.'s scheme, the other three schemes involve the sender's public key and secret key in the keyword encryption algorithm and trapdoor generation algorithm, respectively. It can be seen from the figure that our scheme has shorter trapdoor and cipher text than other schemes. For the other parameters, our scheme still has comparable length with other schemes.

Among these operations, the computation of the pairing is usually the most time-consuming. According to the construction of H_1 in [21], its computation is usually inefficient

with the comparison of the traditional hash function. In a random oracle model, it is easy to construct a PRF from an efficient hash function. From these observations, we can see that our keyword testing algorithm should be much faster than that of the other three schemes. For encryption and the trapdoor generation, the advantage of our scheme is not obvious among them. In terms of security, Boneh et al.'s scheme cannot resist against IKGA. The scheme of [16] can prevent IKGA, but it is not secure in a multiuser setting. The scheme of [19] did not show its security in a multiuser setting.

To evaluate the efficiency of these schemes in practice, we use a laptop with 1.7 GHz Intel i3 CPU, 2 GB memory, and a Windows 7 operating system to implement them. We use the jPBC library and choose a type A pairing, which makes use of the curve $y^2 = x^3 + x$ over the field \mathbb{F}_q for prime $q \equiv 3 \pmod{4}$. We run each algorithm with different times and record their time in seconds. The results are shown in Figures 4, 5, and 6, respectively. As the computations of Noroozi and Eslami and Huang and Li, they possess the same experimental results. Experiment results show that our encryption algorithm and trapdoor generation algorithm are slightly faster than those of the other schemes. But our keyword testing algorithm is significantly faster than that of the other schemes.

5. Conclusion

In this paper, we proposed a new public-key authenticated encryption scheme with keyword search. Our scheme uses the idea of the Diffie-Hellman key exchange protocol to generate a shared secret key between the sender and the receiver. The shared key can be viewed as the secret key of a symmetric-key searchable encryption scheme to encrypt keywords by the sender or to generate search trapdoors by the receiver. Under the ODH assumption, our PAEKS scheme can achieve both trapdoor indistinguishability and cipher text indistinguishability, and hence, it can resist inside keyword guessing attacks. The scheme is also efficient. Specifically, its keyword searching algorithm is very fast in the sense that it requires only one computation of PRF, while the previous schemes require at least one expensive pairing operation.

Data Availability

The data used to support the findings of this study are embedded in the programming. They are available from the corresponding author upon request (email: qinbaodong@xupt.edu.cn).

Conflicts of Interest

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (grant numbers 61872292 and 61772418), the Key Research and Development Program of Shaanxi (grant number 2020ZDLGY08-04), and the Basic Research Program of Qinghai Province (grant number 2020-ZJ-701).

References

- [1] R. Zhang, R. Xue, and L. Liu, "Searchable encryption for healthcare clouds: a survey," *IEEE Transactions on Services Computing*, vol. 11, no. 6, pp. 978–996, 2018.
- [2] W. Sun, Z. Cai, Y. Li, F. Liu, S. Fang, and G. Wang, "Security and privacy in the medical internet of things: a review," *Security and Communication Networks*, vol. 2018, Article ID 5978636, 9 pages, 2018.
- [3] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proceeding 2000 IEEE Symposium on Security and Privacy. S&P 2000*, pp. 44–55, Berkeley, CA, USA, May 2000.
- [4] D. Boneh, G. di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Advances in Cryptology - EUROCRYPT 2004, International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004*, C. Cachin and J. Camenisch, Eds., vol. 3027 of Lecture Notes in Computer Science, pp. 506–522, Springer, 2004.
- [5] H. S. Rhee, W. Susilo, and H. J. Kim, "Secure searchable public key encryption scheme against keyword guessing attacks," *IEICE Electronics Express*, vol. 6, no. 5, pp. 237–243, 2009.
- [6] L. Fang, W. Susilo, C. Ge, and J. Wang, "Public key encryption with keyword search secure against keyword guessing attacks without random oracle," *Information Sciences*, vol. 238, pp. 221–241, 2013.
- [7] P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: a provably secure scheme under keyword guessing attack," *IEEE Transactions on Computers*, vol. 62, no. 11, pp. 2266–2277, 2013.
- [8] C.-H. Wang and T.-Y. Tu, "Keyword search encryption scheme resistant against keyword-guessing attack by the untrusted server," *Journal of Shanghai Jiaotong University (Science)*, vol. 19, no. 4, pp. 440–442, 2014.
- [9] Z. Y. Shao and B. Yang, "On security against the server in designated tester public key encryption with keyword search," *Information Processing Letters*, vol. 115, no. 12, pp. 957–961, 2015.
- [10] Y. Wu, X. Lu, J. Su, and P. Chen, "An efficient searchable encryption against keyword guessing attacks for sharable electronic medical records in cloud-based system," *Journal of Medical Systems*, vol. 40, no. 12, article 258, 2016.
- [11] M. Ma, D. He, N. Kumar, K. K. R. Choo, and J. Chen, "Certificateless searchable public key encryption scheme for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 759–767, 2018.
- [12] W. C. Yau, R. C. W. Phan, S. H. Heng, and B. M. Goi, "Keyword guessing attacks on secure searchable public key encryption schemes with a designated tester," *International Journal of Computer Mathematics*, vol. 90, no. 12, pp. 2581–2587, 2013.
- [13] C. Li, C. Lee, C. Weng, T. Wu, and C. Chen, "Cryptanalysis of an efficient searchable encryption against keyword guessing attacks for shareable electronic medical records in cloud-based system," in *Information Science and Applications 2017-ICISA 2017, Macau, China, 20-23 March 2017*, K. Kim and N. Joukov, Eds., vol. 424 of Lecture Notes in Electrical Engineering, pp. 282–289, Springer, 2017.
- [14] Y. Lu, G. Wang, and J. Li, "Keyword guessing attacks on a public key encryption with keyword search scheme without random oracle and its improvement," *Information Sciences*, vol. 479, pp. 270–276, 2019.

- [15] T. Y. Wu, C. M. Chen, K. H. Wang, and J. M. T. Wu, "Security analysis and enhancement of a certificateless searchable public key encryption scheme for IIoT environments," *IEEE Access*, vol. 7, pp. 49232–49239, 2019.
- [16] Q. Huang and H. Li, "An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks," *Information Sciences*, vol. 403-404, pp. 1–14, 2017.
- [17] T.-Y. Wu, C.-M. Chen, K.-H. Wang, J. M.-T. Wu, and J.-S. Pan, "Security analysis of a public key authenticated encryption with keyword search scheme," in *Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceeding of the Fourteenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, November, 26-28, 2018, Sendai, Japan, Volume 1*, pp. 178–183, Springer, 2019.
- [18] M. Noroozi and Z. Eslami, "Public key authenticated encryption with keyword search: revisited," *IET Information Security*, vol. 13, no. 4, pp. 336–342, 2019.
- [19] B. Qin, Y. Chen, Q. Huang, X. Liu, and D. Zheng, "Public-key authenticated encryption with keyword search revisited: security model and constructions," *Information Sciences*, vol. 516, pp. 515–528, 2020.
- [20] M. Abdalla, M. Bellare, and P. Rogaway, "The oracle Diffie-Hellman assumptions and an analysis of DHIES," *Topics in Cryptology — CT-RSA 2001: The Cryptographers' Track at RSA Conference 2001 San Francisco, CA, USA, April 8–12, 2001 Proceedings*, D. Naccache, Ed., , pp. 143–158, Springer, 2001.
- [21] D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing," *SIAM Journal on Computing*, vol. 32, no. 3, pp. 586–615, 2003.

Research Article

Reliability Evaluation of Generalized Exchanged X -Cubes Based on the Condition of g -Good-Neighbor

Xiaoyan Li,^{1,2} Hongbin Zhuang,^{1,2} Shuming Zhou,^{3,4} Hongju Cheng,^{1,5} Cheng-Kuan Lin,¹ and Wenzhong Guo¹

¹College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China

²Fujian Provincial Key Laboratory of Information Security of Network Systems, Fuzhou University, Fuzhou 350108, China

³College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China

⁴Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350007, China

⁵Key Laboratory of Spatial Data Mining and Information Sharing, Ministry of Education, Fuzhou 350116, China

Correspondence should be addressed to Wenzhong Guo; guowenzhong@fzu.edu.cn

Received 19 February 2020; Accepted 4 May 2020; Published 26 May 2020

Academic Editor: Huaqun Wang

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

In the cloud computing environment with massive information services and decision-making resources, the accuracy and reliability of information are more important than previous single closed systems. Therefore, ensuring the reliability of information and the stable operation of the system are the core problems in the research fields such as the Internet Plus and the Internet of Things. The connectivity and diagnosability are two important measures for the fault tolerance of multiprocessor systems. The g -good-neighbor conditional connectivity (R^g -connectivity) is the minimum number of nodes that make the graph disconnected, and each node has at least g neighbors in every remaining component. The g -good-neighbor conditional diagnosability (g -GNCD) is the maximum number of faulty processors that has been correctly identified in a system, and any fault-free processor has no less than g fault-free neighbors. Exchanged X -cubes are a class of irregular networks, obtained by deleting links from hypercubes and some variant networks of hypercubes (X -cubes). They not only combine the advantages of X -cubes but also reduce the interconnection complexity. Exchanged X -cubes classify its nodes into two different classes clusters with a unique connecting rule. In this paper, we propose the generalized exchanged X -cubes framework so that architecture can be constructed by different connecting rules. Furthermore, we study the R^g -connectivity and g -GNCD of generalized exchanged X -cubes under the PMC and MM* models. As applications, the R^g -connectivity and g -GNCD of generalized exchanged hypercubes, dual-cube-like networks, generalized exchanged crossed cubes, and locally generalized exchanged twisted cubes are determined, respectively.

1. Introduction

With the expansion of network scale and the improvement of complexity, the reliability and stability of the system become more and more important. How to ensure the correct and efficient operation of the system is an important research topic for wireless sensor networks and distributed systems. The distributed system disperses the computing tasks which are originally collected on one computer to polymorphic computer for parallel processing. It has many advantages, such as resource sharing, openness, concurrency, scalability, and fault tolerance. In the operation of the multiprocessor system, processor failure is inevitable. It may slow down the

communication of information or even lead to paralyze of the system, thereby affecting the normal operation of the multiprocessor system and bringing huge losses. For example, on Google and Amazon systems, the failure of processors (servers) for several hours can bring millions of dollars of losses. Therefore, fault tolerance is very important for the construction and maintenance of systems [1, 2]. In fact, a multiprocessor system can be usually enlightened as a simple connected graph, where each processor represents a node of the graph, and each link between two processors represents an edge between two nodes in the graph. The graph is called the interconnection network of this multiprocessor system. Thus, some parameters of a graph as an interconnection

network can be used to measure the reliability of a multi-processor system. In the following, we do not distinguish among multiprocessor systems, interconnection networks, and graphs.

An important evaluating parameter for the fault tolerance of a system (modeled by graph G), the *connectivity*, is denoted by $\kappa(G)$, which is the minimum number of nodes that make the graph disconnected. So far, connectivities for many famous networks have been proven. Nevertheless, there is a shortcoming for using traditional connectivity as a parameter of fault tolerance, which is considered a highly unlikely phenomenon in reality that all nodes adjacent to a node have failed simultaneously. Therefore, Esfahanian and Hakimi [3] proposed a new measure to overcome this shortcoming, the *restricted connectivity*, which limits that all adjacent nodes of any node cannot fail at the same time. Later, a generalized restricted connectivity concept, the *g -restricted connectivity (R^g -connectivity)*, was proposed by Latifi et al. [4], which defines that each node of any remaining component after deleting all faulty nodes has degree at least g . In recent years, they have attracted much interest of theoretical computer scientists and mathematicians. Xu et al. [5] determined the R^g -connectivity of hierarchical cubic networks and complete cubic networks. Ning [6] studied the R^g -connectivity of exchanged crossed cubes. Yuan et al. [7] explored the R^g -connectivity of κ -ary n -cube networks. Lin et al. [8] obtained the R^g -connectivity of (n, k) -arrangement graphs.

Identifying all faulty processors in a multiprocessor system (in brief, system) is called *system-level diagnosis*. A system is t -diagnosable when all faulty processors can be detected, provided that the number of faulty processors in it does not exceed t . The maximum number of faulty processors that the system can precisely point out is as known as the *diagnosability* of the system. In *system-level diagnosis*, there are several well-known models.

The *PMC model* is the first model, proposed by Preparata et al. [9], which is a test-based model, assumes that the adjacent processors can perform tests on each other. For any adjacent processors in a system, the ordered pair $\langle x, y \rangle$ is called a test that x diagnoses its neighbor y , where x is a tester and y is a testee. In case x diagnoses y to be faulty (resp., fault-free), the outcome of the test $\langle x, y \rangle$ is 1 (resp., 0). Moreover, the outcome is reliable in the present of the tester x is fault-free. Another model, the *MM model*, was proposed by Maeng and Malek [10], which is a comparison-based model. In MM model, a comparator processor z sends the same test to its two neighbors x, y (i.e., comparison nodes) and then compares their responses. Let a labeled edge $(x, y)_z$ be a comparison performed that two processors x and y are compared by a processor z , where x is adjacent to z and y is also adjacent to z . If the comparator processor z is fault-free, and the responses of x and y are identical, then both comparison processors x and y are fault-free; on the other hand, if the responses of x and y are different, then at least one of x, y is faulty. Furthermore, if both comparison processors x and y are faulty, the responses of x and y are distinct. In addition, the comparison $(x, y)_z$ is unreliable in the present if the comparator node z is faulty. The *MM* model* (proposed by Sengupta and Dahbura) [11] is a special MM model, which is

assumed that each processor must compare each pair of its adjacent processors.

Since there is no restrictive condition on the distribution pattern of faulty processors, the classical diagnosability of a system is quite small. In order to increase the diagnosability, Lai et al. [12] proposed a more realistic parameter of diagnosability, *conditional diagnosability*, which limited that all the neighbors of any processor cannot be faulty at the same time in a system. Recently, Peng et al. [13] proposed the notion of *g -good-neighbor conditional diagnosability (g -GNCD)*, which is the maximum number of faulty processors that can be identified under the condition that every fault-free processor has no less than g fault-free neighbors. Peng et al. [13] (resp., Wang et al. [14]) established the g -GNCD of hypercubes under the PMC model (resp., MM* model). Li et al. [15] introduced the diagnosability and 1-good-neighbor conditional diagnosability of hypercubes with missing links and broken-down nodes under the PMC model. Yuan et al. [7] studied the g -good-neighbor conditional diagnosabilities of k -ary n -cube networks under the PMC model and the MM* model. Xu et al. [5] established the g -good-neighbor conditional diagnosabilities of complete cubic networks under the PMC model and the MM* model. Lin et al. [16] evaluated the g -good-neighbor conditional diagnosabilities of (n, k) -arrangement graphs under the PMC model and the MM* model. Guo et al. [17] studied the g -good-neighbor conditional diagnosability of the crossed cubes under the PMC model and the MM* model. Li et al. [18] introduced this concept into a family of data center networks—DCell—and determined the g -good-neighbor conditional diagnosabilities of DCell under the PMC model and the MM* model.

The R^g -connectivity (or g -GNCD) of different networks are usually determined independently. It is a very worthwhile topic to explore a unified method to get them in different networks. A family of exchanged networks (i.e., exchanged X -cubes) have some common properties, so that their R^g -connectivity (or g -GNCD) can be studied by a uniform method. The family of exchanged X -cubes not only combine the advantages of hypercubes and some variant networks of hypercubes (X -cubes) but also reduce the interconnection complexity. Exchanged X -cubes classify its nodes into two different classes clusters with a unique connecting rule. In this paper, we propose the generalized exchanged X -cubes framework so that architecture can be constructed by different connecting rules. There are some of the better properties in generalized exchanged X -cubes, such as smaller diameter, fewer edges, lower cost factor, and low latency. Based on the fine properties, the network's hardware and communication costs are reduced, and a greater balance between performance and cost can be achieved. Due to the excellent properties of the generalized exchanged X -cubes, they can be used as the logical topologies in the peer-to-peer environment [19].

In recent years, the research on the relationship between the R^g -connectivity and the g -GNCD of regular networks under certain conditions has been widely developed [20–24], while this paper will study the R^g -connectivity and the g -GNCD of a class of irregular networks (i.e., generalized

exchanged X -cubes). We first establish the R^g -connectivity of generalized exchanged X -cubes. Next, we evaluate the g -GNCD of generalized exchanged X -cubes. As applications, we obtain the R^g -connectivity and g -GNCD of generalized exchanged hypercubes, dual-cube-like networks, generalized exchanged crossed cubes, and locally generalized exchanged twisted cubes.

The remainder of this paper is organized as follows. Section 2 provides the terms and notations used throughout the paper. Section 3 evaluates the R^g -connectivity of generalized exchanged X -cubes. Section 4 establishes the g -GNCD of generalized exchanged X -cubes. Section 5 gives some applications based on the results in Section 3 and Section 4. In Section 6, we illustrate the advantages of R^g -connectivity and g -GNCD compared to traditional connectivity and traditional diagnosability, respectively. Finally, we finish the whole paper by concluding in Section 7.

2. Preliminaries

2.1. Terminology and Notations. In this paper, a multiprocessor system is usually represented by a simple undirected graph (in brief, a graph). For terminology and notations not defined in this paper, we follow the reference [25]. We use $G = (V(G), E(G))$ to represent a graph, where $V(G)$ representing a nonempty and finite *node set* and $E(G) = \{(u, v) \mid (u, v) \text{ is an unordered pair of } V(G)\}$ representing an *edge set*. Two nodes u and v are *adjacent*, denoted by $(u, v) \in E(G)$. The set of *neighbors* of node u in G is denoted by $N_G(u) = \{v \in V(G) \mid (u, v) \in E(G)\}$. If $R \subseteq V(G)$, let $G[R]$ denote the subgraph of G induced by the node subset R in G . And we denote $G - R$ as $G[V(G) \setminus R]$. We set $N_G(R) = \{v \in V(G) \setminus R \mid (u, v) \in E(G) \text{ and } u \in R\} = \bigcup_{u \in R} N_G(u) \setminus R$ and $N_G[R] = N_G(R) \cup R$. Two binary strings $u = u_1u_0$ and $v = v_1v_0$ are pair related, denoted by $u \sim v$, if and only if $(u, v) \in \{(00, 00), (01, 11), (10, 10), (11, 01)\}$. The case that u and v are not pair related is denoted by $u \not\sim v$ [26].

The *degree* of u in G is denoted by $\deg_G(u) = |N_G(u)|$. Let $\delta(G) = \min \{\deg_G(u) \mid u \in V(G)\}$, $\Delta(G) = \max \{\deg_G(u) \mid u \in V(G)\}$. K_n is defined as a complete graph with n nodes. A *path* P is a sequence of distinct nodes with any two consecutive nodes in P that are adjacent. We use $G_1 \cong G_2$ to represent the graph G_1 is isomorphic to the graph G_2 . A *component* is defined as a maximally connected subgraph of a graph.

Definition 1 (see [27]). Let $R \subseteq V(G)$. R is called a *node-cut* if $G - R$ is disconnected. If there exists a node-cut R with $|R| = k$, then R is called a k -*node-cut*. The connectivity $\kappa(G)$ of G is defined as the minimum k such that G has a k -node-cut.

Definition 2 (see [4]). Let g be a positive integer and $R \subseteq V(G)$. If $G - R$ is disconnected and each remaining component has minimum degree at least g , then R is called an R^g -*cut*.

Definition 3 (see [4]). The R^g -connectivity of G , denoted by $\kappa^g(G)$, is the minimum cardinality over all R^g -cuts of G .

2.2. The g -Good-Neighbor Conditional Diagnosability. Under the PMC model and MM* model, we call the notation Ω as the *syndrome* of the system, which is defined as the set of all test (comparison) results in a system G , where test results are based on the PMC model and comparison results are based on the MM* model. Define a faulty set F , where $\forall i \in F$, i is a faulty processor. Let $\Omega(F)$ be the set of test (comparison) results which could be produced if F is the faulty node set. We use \widehat{F}_1 and \widehat{F}_2 to represent two distinct faulty sets of $V(G)$. In case $\Omega(\widehat{F}_1) \cap \Omega(\widehat{F}_2) = \emptyset$, we call these two distinct faulty sets \widehat{F}_1 and \widehat{F}_2 *distinguishable*, and $(\widehat{F}_1, \widehat{F}_2)$ a *distinguishable pair*; otherwise, \widehat{F}_1 and \widehat{F}_2 are *indistinguishable*, and $(\widehat{F}_1, \widehat{F}_2)$ is an *indistinguishable pair*. Let $\widehat{F}_1 \Delta \widehat{F}_2$ be the symmetric difference $(\widehat{F}_1 - \widehat{F}_2) \cup (\widehat{F}_2 - \widehat{F}_1)$ between \widehat{F}_1 and \widehat{F}_2 . In [28], under the PMC model, the sufficient and necessary condition for two different subsets \widehat{F}_1 and \widehat{F}_2 is a distinguishable pair proposed by Dahbura and Masson. Moreover, under the MM* model, the sufficient and necessary condition for two different subsets \widehat{F}_1 and \widehat{F}_2 is a distinguishable pair proposed by Sengupta and Dahbura [11].

Lemma 4 (see [28]). Let $G = (V(G), E(G))$ be a multiprocessor system. For any two distinct sets $\widehat{F}_1, \widehat{F}_2 \subseteq V(G)$, \widehat{F}_1 and \widehat{F}_2 are distinguishable under the PMC model if and only if there exists at least one test from $V(G) - (\widehat{F}_1 \cup \widehat{F}_2)$ to $\widehat{F}_1 \Delta \widehat{F}_2$ (see Figure 1(a)).

Lemma 5 (see [11]). Let $G = (V(G), E(G))$ be a multiprocessor system. For any two distinct sets $\widehat{F}_1, \widehat{F}_2 \subseteq V(G)$, \widehat{F}_1 and \widehat{F}_2 are distinguishable under the MM* model if and only if there is a node $w \in V(G) - \widehat{F}_1 \cup \widehat{F}_2$ such that one of the following conditions holds (see Figure 1(b)):

- (1) $|N_G(w) - (\widehat{F}_1 \cup \widehat{F}_2)| \geq 1$ and $|N_G(w) \cap (\widehat{F}_1 \Delta \widehat{F}_2)| \geq 1$,
- (2) $|N_G(w) \cap (\widehat{F}_1 - \widehat{F}_2)| \geq 2$,
- (3) $|N^{-G}(w) \cap ((F^{-2})^\wedge - (F^{-1})^\wedge)| \geq 2$.

The concept of g -GNCD of a system was proposed in the literature [13].

Definition 6 (see [13]).

- (1) Let $\widehat{F} \subseteq V(G)$ and \widehat{F} be a fault-set. If any node of $V(G) - \widehat{F}$ has at least g neighbors in $G - \widehat{F}$, then \widehat{F} is called a g -good-neighbor conditional fault-set.
- (2) A system G is g -good-neighbor conditional t -diagnosable if each distinct pair of g -good-neighbor conditional faulty (g -GNCF) sets \widehat{F}_1 and \widehat{F}_2 of $V(G)$ with $|\widehat{F}_1| \leq t$ and $|\widehat{F}_2| \leq t$ are distinguishable.

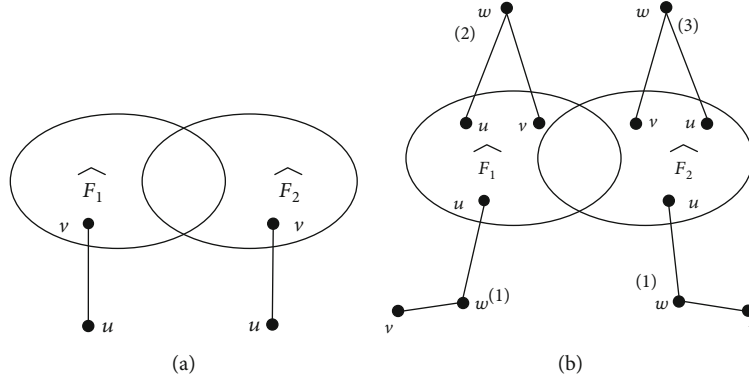


FIGURE 1: (a) An illustration for Lemma 4. (b) An illustration for Lemma 5.

- (3) The g -GNCD, denoted by $t_g(G)$, is defined as the maximum value of t such that G is g -good-neighbor conditionally t -diagnosable. Let $t_g^P(G)$ and $t_g^M(G)$ be the g -GNCD of G under the PMC model and MM^* model, respectively.

2.3. *Generalized Exchanged X-Cubes.* In this subsection, we give the definition of the family of generalized exchanged networks, denoted by *generalized exchanged X-cubes*, which have some common properties, so that their R^g -connectivity (or g -GNCD) can be studied by a uniform method. Since generalized exchanged X-cubes are derived by BC networks (bijective connection networks), we first review the definition of the BC network.

Definition 7 (see [29]). The one-dimensional BC network X_1 contains only two nodes which forms an edge. We use \mathbb{L}_1 to represent the family of the one-dimensional BC network with $\mathbb{L}_1 = \{X_1\}$. A graph G belongs to the family of n -dimensional BC networks \mathbb{L}_n if and only if there exists $V_0, V_1 \subset V(G)$ such that the following two conditions hold:

$$\begin{aligned} V(G) &= V_0 \cup V_1, V_0 \neq \emptyset, V_1 \neq \emptyset, V_0 \cap V_1 \\ &= \emptyset, \text{ and } G[V_0], G[V_1] \in \mathbb{L}_{n-1} \end{aligned} \quad (1)$$

$E(V_0, V_1)$ is a perfect matching M between V_0 and V_1 in G

For any $X_n \in \mathbb{L}_n$, by Definition 7, there exist V_0, V_1, M satisfying the conditions. We use X_{n-1}^0, X_{n-1}^1 to denote the induced subgraph $G[V_0], G[V_1]$, respectively. Clearly, they are both $(n-1)$ -dimensional BC networks, and $E(X_{n-1}^0), E(X_{n-1}^1), M$ is a decomposition of $E(X_n)$. We define the decomposition as $X_n = G(X_{n-1}^0, X_{n-1}^1; M)$.

BC networks are a class of networks containing a number of famous networks such as hypercubes [13], the Möbius cubes [30], crossed cubes [31], and locally twisted cubes [32] as members. An n -dimensional BC network X_n is n -regular and consisting of 2^n nodes. Figure 2 shows two three-dimensional BC networks.

Lemma 8 (see [33]). For $0 \leq g \leq n$ and $Y \subset V(X_n)$, if $\delta(X_n[Y]) \geq g$, then $|Y| \geq 2^g$.

Lemma 9 (see [34]).

- (1) For $Y \subset V(X_n)$ and $0 \leq g \leq n$, if $\delta(X_n[Y]) = g$, then $|N_{X_n}(Y)| \geq (n-g)2^g$.
- (2) For $0 \leq g \leq n-2$, $\kappa^g(X_n) = (n-g)2^g$.

Lemma 10 (see [35]). For $n \geq 2$, there are at most two common neighbors between any two nodes of X_n .

Next, we introduce the definition of generalized exchanged X-cubes.

Definition 11. The (s, t) -dimensional generalized exchanged X-cubes is defined as a graph $GEX(s, t) = (V(GEX(s, t)), E(GEX(s, t)))$, for $s \geq 1$ and $t \geq 1$. $GEX(s, t)$ consists of two disjoint subgraphs \widetilde{L}' and \widetilde{R}' . And \widetilde{L}' consists of 2^t subgraphs, denoted by \widetilde{L}_i' for $i = 1, 2, \dots, 2^t$. Similarly, \widetilde{R}' consists of 2^s subgraphs, denoted by \widetilde{R}_j' for $j = 1, 2, \dots, 2^s$. Moreover, $GE X(s, t)$ satisfies the following conditions (see Figure 3):

- (a) For any integers $1 \leq i \leq 2^t$ and $1 \leq j \leq 2^s$, $\widetilde{L}_i' \cong X_s$ and $\widetilde{R}_j' \cong X_t$. Further, $|V(\widetilde{L}_i')| = 2^s$ and $|V(\widetilde{R}_j')| = 2^t$
- (b) Each node in $V(\widetilde{L}')$ has a sole neighbor in $V(\widetilde{R}')$ and vice versa. In addition, for distinct nodes in each \widetilde{L}_i' , their neighbors of \widetilde{R}' lie in different \widetilde{R}_j'
- (c) For any two different subgraphs \widetilde{L}_i' and \widetilde{L}_h' with $i \neq h$, there exists no edge between them. Similar for \widetilde{R}_j' and \widetilde{R}_k' with $j \neq k$.

By Definition 11, we can deduce that $|V(GEX(s, t))| = 2^{s+t+1}$. Let each of \widetilde{L}_i' and \widetilde{R}_j' be a cluster of $GEX(s, t)$. Obviously, $GEX(s, t)$ consists of $2^t + 2^s$ clusters. If we contract each cluster as a node, then $GEX(s, t)$ is contracted into a complete bipartite graph $K_{2^t, 2^s}$. The edges that connect

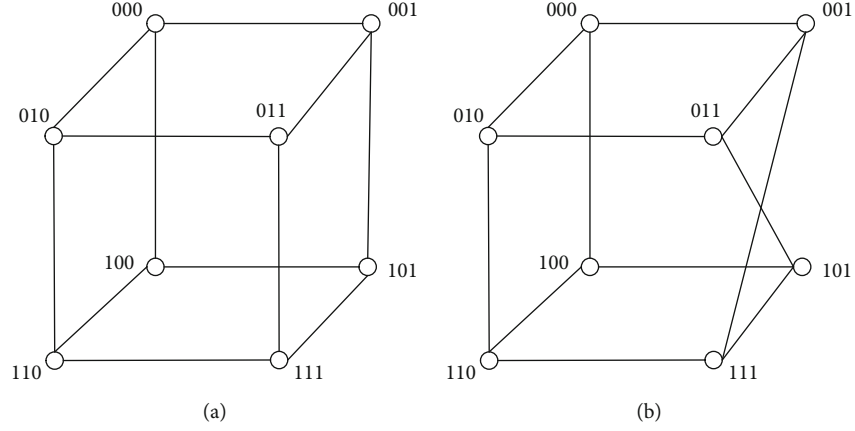
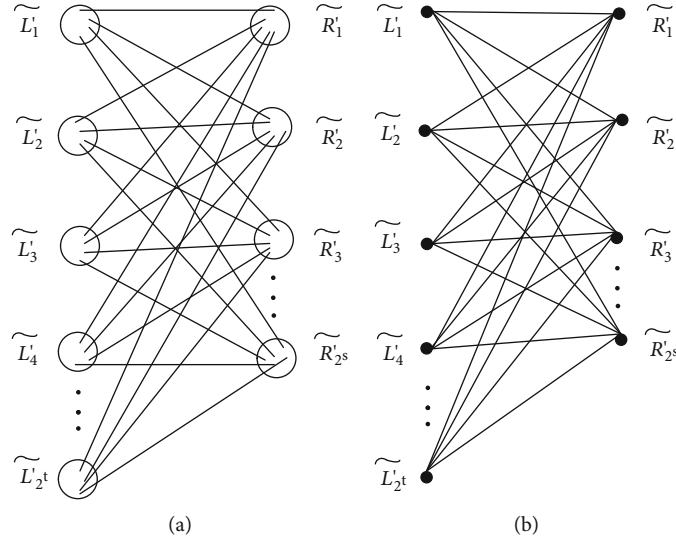


FIGURE 2: Two three-dimensional BC networks.


 FIGURE 3: (a) The partition of $GEX(s, t)$. (b) The contraction of $GEX(s, t)$.

different clusters are called cross edges. In the following discussion, we consider $s \leq t$, and thus, $\delta(GEX(s, t)) = s + 1$, $\Delta(GEX(s, t)) = t + 1$.

3. The R^g -Connectivity of $GEX(s, t)$

In this section, we establish the R^g -connectivity of $GEX(s, t)$ with $1 \leq g \leq s - 2$. In what follows, we exploit some useful lemmas for our further investigation.

Lemma 12. *For any integers $s \geq 3$ and $1 \leq g \leq s$, let H be a subgraph of X_s with $\delta(H) \geq g$, and let T be a subgraph of X_s such that $T \cong X_g$. Then $|N_{X_s}[H]| \geq |N_{X_s}[T]| = (s - g + 1)2^g$.*

Proof. We conduct induction on s .

If $s = 3$, by fixing g , the lemma holds obviously. Suppose that the lemma holds for $s = \tau - 1$, let H be a subgraph of $X_{\tau-1}$ with $\delta(H) \geq g$ and T_1 be a subgraph of $X_{\tau-1}$ such that $T_1 \cong X_g$, then $|N_{X_{\tau-1}}[H]| \geq |N_{X_{\tau-1}}[T_1]| = (\tau - g)2^g$ for $1 \leq g \leq \tau - 1$ and $\tau \geq 4$. In the following, we will prove that

the lemma holds for $s = \tau$. Since X_τ can be merged through a perfect matching by two $X_{\tau-1}$, namely $X_{\tau-1}^0$ and $X_{\tau-1}^1$, we discuss the two cases below.

Case 1. $H \cap X_{\tau-1}^0 \neq \emptyset$ and $H \cap X_{\tau-1}^1 \neq \emptyset$.

Let $H_1 = H - X_{\tau-1}^1$ and $H_2 = H - X_{\tau-1}^0$. Then $\delta(H_1) \geq g - 1$ and $\delta(H_2) \geq g - 1$. Let T_2 and T_3 be two subgraphs of $X_{\tau-1}^0$ and $X_{\tau-1}^1$ with $T_2 \cong X_{g-1}$ and $T_3 \cong X_{g-1}$, respectively. Thus, by the induction hypothesis, we have

$$\begin{aligned} |N_{X_\tau}[H]| &= |N_{X_{\tau-1}}[H_1]| + |N_{X_{\tau-1}}[H_2]| \\ &\geq |N_{X_{\tau-1}}[T_2]| + |N_{X_{\tau-1}}[T_3]| \\ &\geq 2[\tau - g + 1]2^{g-1} = (\tau - g + 1)2^g. \end{aligned} \quad (2)$$

Then, for $s = \tau$, the lemma holds.

Case 2. $H \subseteq X_{\tau-1}^0$ or $H \subseteq X_{\tau-1}^1$.

Without loss of generality, we suppose that $H \subseteq X_{\tau-1}^0$. By Lemma 8, we have $|V(H)| \geq 2^g$. Further, by the induction hypothesis, $|N_{X_{\tau-1}}[H]| \geq |N_{X_{\tau-1}}[T_1]| = (\tau - g)2^g$. Then

$$\begin{aligned} |N_{X_\tau}[H]| &= |N_{X_{\tau-1}}[H]| + |V(H)| \\ &\geq |N_{X_{\tau-1}}[T_1]| + |V(H)| \\ &\geq (\tau - g)2^g + 2^g = (\tau - g + 1)2^g. \end{aligned} \quad (3)$$

Hence, the lemma holds.

Lemma 13. For any integers $s \geq 3$ and $1 \leq g \leq s - 2$, $\kappa^g(GEX(s, t)) \leq (s - g + 1)2^g$.

Proof. By Definition 11, $GEX(s, t)$ can be decomposed into two disjoint subgraphs \widetilde{L}' and \widetilde{R}' , where \widetilde{L}' can be partitioned into 2^t subgraphs (clusters) and \widetilde{R}' can be partitioned into 2^s subgraphs (clusters). Without loss of generality, let $A \subseteq V(\widetilde{L}_1')$ such that $GEX(s, t)[A] \cong X_g$. Clearly, $|A| = 2^g$. By Definition 7 and Lemma 9, $|N_{\widetilde{L}_1'}(A)| = (s - g)2^g$. Further, by Definition 11, each node in $V(\widetilde{L}_1')$ has a sole neighbor in $V(\widetilde{R}')$. And, for distinct nodes in each \widetilde{L}_i' , their neighbors of \widetilde{R}' lie in different \widetilde{R}_j' . In addition, for any two different subgraphs \widetilde{L}_i' and \widetilde{L}_h' with $i \neq h$, the edge between them is nonexistent. Thus, each node in \widetilde{L}_1' has exactly one neighbor in $GEX(s, t) - \widetilde{L}_1'$. Then $|N_{GEX(s, t) - \widetilde{L}_1'}(A)| = |A| = 2^g$. Thus, we have

$$\begin{aligned} |N_{GEX(s, t)}(A)| &= |N_{\widetilde{L}_1'}(A)| + |N_{GEX(s, t) - \widetilde{L}_1'}(A)| \\ &= (s - g)2^g + 2^g = (s - g + 1)2^g \end{aligned} \quad (4)$$

Since $|N_{GEX(s, t)}[A]| = (s - g + 1)2^g + 2^g = (s - g + 2)2^g$ and $|V(GEX(s, t))| = 2^{s+t+1} > (s - g + 2)2^g$, $GEX(s, t) - N_{GEX(s, t)}(A)$ is disconnected. Then $N_{GEX(s, t)}(A)$ is a node-cut of $GEX(s, t)$.

In what follows, $N_{GEX(s, t)}(A)$ as an R^g -cut of $GEX(s, t)$ will be proved. That is, $\delta(GEX(s, t) - N_{GEX(s, t)}(A)) \geq g$.

Since $GEX(s, t)[A] \cong X_g$, $\delta(GEX(s, t)[A]) = g$. By Lemma 9, $N_{\widetilde{L}_1'}(A)$ is an R^g -cut of \widetilde{L}_1' , where $1 \leq g \leq s - 2$. As a result, $\delta(\widetilde{L}_1' - N_{\widetilde{L}_1'}(A)) \geq g$ with $1 \leq g \leq s - 2$. Moreover, by Definition 11, each node in \widetilde{L}_1' has exactly one neighbor in $GEX(s, t) - \widetilde{L}_1'$, and for distinct nodes in \widetilde{L}_1' , their neighbors in \widetilde{R}' lie in different \widetilde{R}_j' . Since $\delta(GEX(s, t)) = s + 1$, $\delta(GEX(s, t) - \widetilde{L}_1' - N_{GEX(s, t) - \widetilde{L}_1'}(A)) \geq s + 1 - 1 > g$ for any node $w \in GEX(s, t) - \widetilde{L}_1' - N_{GEX(s, t) - \widetilde{L}_1'}(A)$.

Summary of the above discussion, we have $\delta(GEX(s, t) - N_{GEX(s, t)}(A)) \geq g$. Then $N_{GEX(s, t)}(A)$ is a g -good-neighbor cut of $GEX(s, t)$. Hence, $\kappa^g(GEX(s, t)) \leq (s - g + 1)2^g$ with $1 \leq g \leq s - 2$ and $s \geq 3$, the lemma holds.

Lemma 16. For any integers $s \geq 3$ and $1 \leq g \leq s - 2$, $\kappa^g(GEX(s, t)) \geq (s - g + 1)2^g$.

Proof. We assume U as a minimum R^g -cut of $GEX(s, t)$. Let $U \cap V(\widetilde{L}_i') = U_{L_i'}$ and $U \cap V(\widetilde{R}_j') = U_{R_j'}$, where $1 \leq i \leq 2^t$, $1 \leq j \leq 2^s$. Then we will show that $\kappa^g(GEX(s, t)) = |U| \geq (s - g + 1)2^g$ with $1 \leq g \leq s - 2$ and $s \geq 3$. We consider three cases as follows.

Case 1. $\widetilde{L}_i' - U_{L_i'}$ and $\widetilde{R}_j' - U_{R_j'}$ are connected for each i, j , where $1 \leq i \leq 2^t$, $1 \leq j \leq 2^s$.

We prove this case by contradiction. Suppose that $|U| \leq (s - g + 1)2^g - 1$. In the following, we will prove that U is not an R^g -cut of $GEX(s, t)$.

Since U is a minimum R^g -cut of $GEX(s, t)$, $GEX(s, t) - U$ is disconnected. In addition, there must exist a component C with C traverses r clusters, where $1 \leq r \leq 2^{s-1}$. Let $U_i = C_i \cap U$, where C_i be one of these r clusters with $1 \leq i \leq r$. As a result, $C = \bigcup_{i=1}^r (C_i - U_i)$. By Definition 11, for any node in $V(\widetilde{L}')$, it has a sole neighbor in $V(\widetilde{R}')$. And, for distinct nodes in each \widetilde{L}_i' , their neighbors of \widetilde{R}' lie in different \widetilde{R}_j' . In addition, for any two different subgraphs \widetilde{L}_i' and \widetilde{L}_h' with $i \neq h$, there exists no edge between them. Then there exist at most $r - 1$ cross edges between $C_i - U_i$ and C_{I-p} , where $I = \{1, 2, \dots, r\}$. Moreover, there are at least $2^s - |U_i| - (r - 1)$ cross edges between $C_i - U_i$ and C_j with $J = \{r + 1, r + 2, \dots, 2^s + 2^t\}$. Clearly, $C_j = \bigcup_{j=r+1}^{2^s+2^t} C_j$ and $U_j = \bigcup_{j=r+1}^{2^s+2^t} U_j$. Since there is no edge between $C_i - U_i$ and $C_j - U_j$, $|U_j| \geq \sum_{i=1}^r [2^s - |U_i| - (r - 1)]$. Then, we have

$$\begin{aligned} |U| &= |U_1| + |U_2| + \dots + |U_r| + |U_J| \\ &\geq |U_1| + |U_2| + \dots + |U_r| + \sum_{i=1}^r [2^s - |U_i| - (r - 1)] \\ &= r[2^s - r + 1]. \end{aligned} \quad (5)$$

Let $f(r) = r(2^s - r + 1)$ with $1 \leq r \leq 2^{s-1}$. We obtain $\partial f(r)/\partial r = 2^s - 2r + 1 > 0$. Thus, $f(r)$ is an increasing function. Therefore, $f(r) \geq f(1) = 2^s$ and $|U| \geq f(r) \geq 2^s$. In addition, let $f(g) = 2^s - [(s - g + 1)2^g - 1]$ with $1 \leq g \leq s - 2$. We obtain that $\partial f(g)/\partial g = 2^g[g \ln 2 + 1 - (s + 1) \ln 2] < 0$. Thus, $f(g)$ is a decreasing function. Therefore, $f(g) \geq f(s - 2) = 2^{s-2} + 1 > 0$. Then $|U| \geq 2^s > (s - g + 1)2^g - 1$, which results in a contradiction with $|U| \leq (s - g + 1)2^g - 1$.

Case 2. Only one of $\widetilde{L}_i' - U_{L_i'}$ and $\widetilde{R}_j' - U_{R_j'}$ is disconnected, where $1 \leq i \leq 2^t$, $1 \leq j \leq 2^s$.

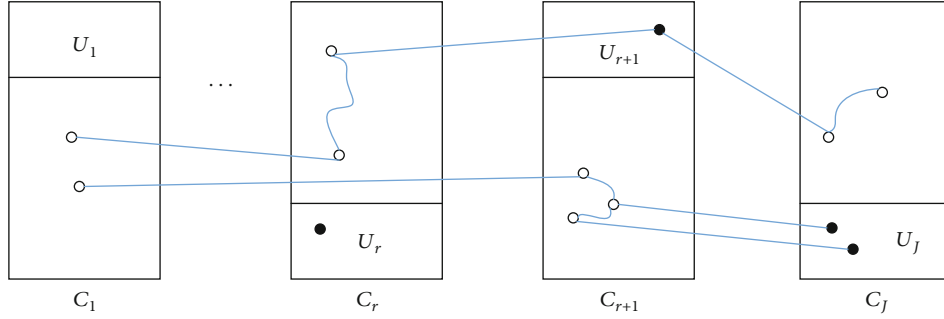


FIGURE 4: An illustration for Case 2 of Lemma 16.

Without loss of generality, assume that $\widetilde{L}_1' - U_{L_1}'$ is disconnected. Since U is an R^g -cut of $GEX(s, t)$, $\delta(\widetilde{L}_1' - U_{L_1}') \geq g - 1$. Then U_{L_1}' is a $(g - 1)$ -good-neighbor cut of $GEX(s, t)$. By Lemma 9, $|U_{L_1}'| \geq (s - g + 1)2^{g-1}$. By contradiction, suppose that $|U| \leq (s - g + 1)2^g - 1$ with $1 \leq g \leq s - 2$ and $s \geq 3$. Let $U_{L_1}' = U_1$ and $U_I = \bigcup_{i=2}^{2^s+2^t} U_i$. Then $|U_I| = |U - U_1| \leq (s - g + 1)2^{g-1} - 1$.

Assume that $GEX(s, t) - \widetilde{L}_1' - U_I$ is disconnected. Then there must exist a component C such that C traverses r clusters, where $1 \leq r \leq \lfloor 2^{s-1} - 1 \rfloor$. Let C_i be one of these r clusters for $2 \leq i \leq r + 1$ and $U_i = C_i \cap U$. As a result, $C = \bigcup_{i=2}^{r+1} (C_i - U_i)$. By Definition 11, for any node in $V(\widetilde{L}_1')$, it has a sole neighbor in $V(\widetilde{R}')$. And for distinct nodes in each \widetilde{L}_i' , their neighbors of \widetilde{R}' lie in different \widetilde{R}_j' . In addition, for any two different subgraphs \widetilde{L}_i' and \widetilde{L}_h' with $i \neq h$, the edge between them is nonexistent. Then there exist at most r cross edges between $C_i - U_i$ and $C_{I'} - U_{I'}$, where $I' = \{1, 2, \dots, r, r + 1\}$. Moreover, there are at least $2^s - |U_i| - r$ cross edges between $C_i - U_i$ and C_J with $J = \{r + 2, r + 3, \dots, 2^s + 2^t\}$. Clearly, $C_J = \bigcup_{j=r+2}^{2^s+2^t} C_j$ and $U_J = \bigcup_{j=r+2}^{2^s+2^t} U_j$. Since there is no edge between $C_i - U_i$ and $C_J - U_J$, $|U_J| \geq \sum_{i=2}^{r+1} [2^s - |U_i| - r]$. Figure 4 shows an illustration for this case. Then, we have

$$\begin{aligned} |U_I| &= |U_2| + \dots + |U_{r+1}| + |U_J| \\ &\geq |U_2| + \dots + |U_{r+1}| + \sum_{i=2}^{r+1} [2^s - |U_i| - r] = r(2^s - r). \end{aligned} \quad (6)$$

Let $f(r) = r(2^s - r)$ with $1 \leq r \leq \lfloor 2^{s-1} - 1 \rfloor$. We obtain that $\partial f(r)/\partial r = 2^s - 2r > 0$. Thus, $f(r)$ is an increasing function. Therefore, $f(r) \geq f(1) = 2^s - 1$. And $|U_I| \geq f(r) \geq 2^s - 1$. In addition, let $f(g) = 2^s - 1 - [(s - g + 1)2^{g-1} - 1]$ with $1 \leq g \leq s - 2$ and $s \geq 3$. We get that $\partial f(g)/\partial g = 2^{g-1}[g \ln 2 + 1 - (s + 1) \ln 2] < 0$. Thus, $f(g)$ is a decreasing function. Therefore, $f(g) \geq f(s - 2) > 0$. Then $|U_I| \geq 2^s - 1 > (s - g + 1)2^{g-1} - 1$, which results in a contradiction with $|U_I| \leq (s - g + 1)2^{g-1} - 1$.

Thus, $GEX(s, t) - \widetilde{L}_1' - U_I$ is connected. Since $\widetilde{L}_1' - U_1$ is disconnected, there must exist a component H in $\widetilde{L}_1' - U_1$ such that there is no edge between H and $GEX(s, t) - \widetilde{L}_1' - U_I$. Then $|U_I| \geq |V(H)|$. Since $N_{L_1}'(H) \subseteq U_1$, $|U| = |U_1| + |U_I| \geq |N_{L_1}'(H)| + |V(H)| = |N_{L_1}'[H]|$. Moreover, since U is an R^g -cut of $GEX(s, t)$, $\delta(H) \geq g$. By Lemma 8 and $\widetilde{L}_1' \cong X_s$, we have $|U| = |N_{L_1}'[H]| \geq (s - g + 1)2^g$, which results in a contradiction with $|U| \leq (s - g + 1)2^g - 1$.

Case 3. For any integers $1 \leq i \leq 2^t$, $1 \leq j \leq 2^s$, there are at least two of $\widetilde{L}_i' - U_{L_i}'$ and $\widetilde{R}_j' - U_{R_j}'$ that are disconnected.

Without loss of generality, suppose that $\widetilde{L}_1' - U_{L_1}'$ and $\widetilde{L}_2' - U_{L_2}'$ are disconnected. Since U is an R^g -cut of $GEX(s, t)$ and by Definition 11, we have $\delta(\widetilde{L}_1' - U_{L_1}') \geq g - 1$ and $\delta(\widetilde{L}_2' - U_{L_2}') \geq g - 1$. Then U_{L_1}' and U_{L_2}' are two $(g - 1)$ -good-neighbor cuts of $GEX(s, t)$. By Lemma 9, $|U_{L_1}'| \geq (s - g + 1)2^{g-1}$ and $|U_{L_2}'| \geq (s - g + 1)2^{g-1}$. Then

$$\begin{aligned} |U| &\geq |U_{L_1}'| + |U_{L_2}'| \geq (s - g + 1)2^{g-1} + (s - g + 1)2^{g-1} \\ &= (s - g + 1)2^g. \end{aligned} \quad (7)$$

Thus, $\kappa^g(GEX(s, t)) = |U| \geq (s - g + 1)2^g$.

Hence, the lemma holds.

Combining Lemma 13 and Lemma 16, the following theorem holds.

Theorem 14. For any integers $s \geq 3$ and $1 \leq g \leq s - 2$, $\kappa^g(GEX(s, t)) = (s - g + 1)2^g$.

4. The g -Good-Neighbor Conditional Diagnosability of $GEX(s, t)$

In this section, we will determine the g -GNCD of $GEX(s, t)$ under the PMC model and MM* model, respectively, where $1 \leq g \leq s - 2$.

Theorem 15. For any integers $s \geq 3$ and $1 \leq g \leq s-2$, $t_g^P(GEX(s, t)) = (s-g+2)2^g - 1$.

Proof. First, we show that $t_g^P(GEX(s, t)) \leq (s-g+2)2^g - 1$ with $1 \leq g \leq s-2$ and $s \geq 3$. Let $A \subseteq V(C_1)$ with $C_1 \cong X_s$ such that $GEX(s, t)[A] \cong X_g$. Clearly, $|A| = 2^g$. Suppose that $\widehat{F}_1 = N_{GEX(s, t)}(A)$ and $\widehat{F}_2 = N_{GEX(s, t)}[A]$. By Lemma 13, we have $|\widehat{F}_1| = |N_{GEX(s, t)}(A)| = (s-g+1)2^g$, and $|\widehat{F}_2| = |N_{GEX(s, t)}[A]| = (s-g+2)2^g$, where $\delta(GEX(s, t) - \widehat{F}_2) \geq g$. Since $\delta(GEX(s, t) - \widehat{F}_1) \geq g$ and $\delta(GEX(s, t) - \widehat{F}_2) \geq g$, \widehat{F}_1 and \widehat{F}_2 are two g -GNCF sets of $GEX(s, t)$ with $|\widehat{F}_1| \leq (s-g+2)2^g$ and $|\widehat{F}_2| \leq (s-g+2)2^g$. On the other hand, since $V(A) = \widehat{F}_1 \Delta \widehat{F}_2$ and $N_{GEX(s, t)}(A) = \widehat{F}_1$, there is no edge between $\widehat{F}_1 \Delta \widehat{F}_2$ and $GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$. By Lemma 4, \widehat{F}_1 and \widehat{F}_2 are indistinguishable under the PMC model. By Definition 6 (2), $GEX(s, t)$ is not g -good-neighbor conditional $(s-g+2)2^g$ -diagnosable under the PMC model. That is, $t_g^P(GEX(s, t)) \leq (s-g+2)2^g - 1$ for $1 \leq g \leq s-2$.

Next, we prove that $t_g^P(GEX(s, t)) \geq (s-g+2)2^g - 1$ with $1 \leq g \leq s-2$ and $s \geq 3$. We suppose, to the contrary, that $t_g^P(GEX(s, t)) \leq (s-g+2)2^g - 2$ for $1 \leq g \leq s-2$. And assume that there are two indistinguishable g -GNCF sets \widehat{F}_1 and \widehat{F}_2 with $|\widehat{F}_1| \leq (s-g+2)2^g - 1$ and $|\widehat{F}_2| \leq (s-g+2)2^g - 1$. In what follows, we consider two cases.

Case 1. $V(GEX(s, t)) = \widehat{F}_1 \cup \widehat{F}_2$.

Since $s \leq t$, by Definition 11, we have $|V(GEX(s, t))| = 2^{s+t+1} \geq 2^{2s+1}$.

Since $|\widehat{F}_1 \cup \widehat{F}_2| \leq |\widehat{F}_1| + |\widehat{F}_2| \leq 2[(s-g+2)2^g - 1]$, we have

$$|V(GEX(s, t))| - |\widehat{F}_1 \cup \widehat{F}_2| \geq 2^{2s+1} - 2[(s-g+2)2^g - 1]. \quad (8)$$

Let $f(g) = 2^{2s+1} - (s-g+2)2^{g+1} + 2$ with $1 \leq g \leq s-2$ and $s \geq 3$. We obtain that $\partial f(g)/\partial g = [(g-s-2) \ln 2 + 1] 2^{g+1} < 0$. Thus, $f(g)$ is a decreasing function. Therefore, for $1 \leq g \leq s-2$ and $s \geq 3$,

$$f(g) \geq f(s-2) = 2^{2s+1} - 2^{s+1} + 2 > 0, \quad (9)$$

which induces a contradiction since $V(GEX(s, t)) = \widehat{F}_1 \cup \widehat{F}_2$.

Case 2. $V(GEX(s, t)) \neq \widehat{F}_1 \cup \widehat{F}_2$.

Since $\widehat{F}_1 \neq \widehat{F}_2$, we may assume that $\widehat{F}_2 - \widehat{F}_1 \neq \emptyset$. There exists no edge between $V(GEX(s, t)) - \widehat{F}_1 \cup \widehat{F}_2$ and $\widehat{F}_1 \Delta \widehat{F}_2$ because \widehat{F}_1 and \widehat{F}_2 are indistinguishable. Moreover, since \widehat{F}_1 is a g -good-neighbor conditional faulty set, it is easy to verify that $\delta(GEX(s, t)[\widehat{F}_2 - \widehat{F}_1]) \geq g$. By Lemma 8, $|\widehat{F}_2 - \widehat{F}_1| \geq 2^g$. On the other hand, since both \widehat{F}_1 and \widehat{F}_2 are g -

GNCF sets, $\widehat{F}_1 \cap \widehat{F}_2$ is also a g -good-neighbor conditional faulty set. Moreover, there is no edge between $V(GEX(s, t)) - \widehat{F}_1 \cup \widehat{F}_2$ and $\widehat{F}_1 \Delta \widehat{F}_2$; thus, $GEX(s, t) - \widehat{F}_1 \cap \widehat{F}_2$ is disconnected. Then $\widehat{F}_1 \cap \widehat{F}_2$ is an R^g -cut of $GEX(s, t)$. By Theorem 14, $|\widehat{F}_1 \cap \widehat{F}_2| \geq (s-g+1)2^g$ with $s \geq 3$ and $1 \leq g \leq s-2$. Hence,

$$\begin{aligned} |\widehat{F}_2| &= |\widehat{F}_2 - \widehat{F}_1| + |\widehat{F}_1 \cap \widehat{F}_2| \geq 2^g + (s-g+1)2^g \\ &= (s-g+2)2^g, \end{aligned} \quad (10)$$

which results in a contradiction since $|\widehat{F}_2| \leq (s-g+2)2^g - 1$.

To sum up, we can conclude that $t_g^P(GEX(s, t)) \geq (s-g+2)2^g - 1$ for any integers $1 \leq g \leq s-2$ and $s \geq 3$.

Hence, the theorem holds.

Theorem 24. For any integers $1 \leq g \leq s-2$ and $s \geq 4$, $t_g^M(GEX(s, t)) = (s-g+2)2^g - 1$.

Proof. The proof of $t_g^M(GEX(s, t)) \leq (s-g+2)2^g - 1$ with $1 \leq g \leq s-2$ and $s \geq 4$ is similar to Theorem 15, so it is omitted.

Next, we prove that $t_g^M(GEX(s, t)) \geq (s-g+2)2^g - 1$ with $s \geq 4$ and $1 \leq g \leq s-2$. We suppose, to the contrary, that $t_g^M(GEX(s, t)) \leq (s-g+2)2^g - 2$ with $s \geq 4$ and $1 \leq g \leq s-2$. Moreover, we assume that there are two indistinguishable g -GNCF sets \widehat{F}_1 and \widehat{F}_2 , where $|\widehat{F}_1| \leq (s-g+2)2^g - 1$ and $|\widehat{F}_2| \leq (s-g+2)2^g - 1$.

Since $s \leq t$, by Definition 11, we have $|V(GEX(s, t))| = 2^{s+t+1} \geq 2^{2s+1}$.

Furthermore, it is easy to get that $|\widehat{F}_1 \cup \widehat{F}_2| \leq |\widehat{F}_1| + |\widehat{F}_2| \leq 2[(s-g+2)2^g - 1]$. Then, we have

$$|V(GEX(s, t))| - |\widehat{F}_1 \cup \widehat{F}_2| \geq 2^{2s+1} - 2[(s-g+2)2^g - 1]. \quad (11)$$

Let $f(g) = 2^{2s+1} - (s-g+2)2^{g+1} + 2$ with $s \geq 4$ and $1 \leq g \leq s-2$. We obtain that $\partial f(g)/\partial g = [g \ln 2 + 1 - (s+2) \ln 2] 2^{g+1} < 0$. Thus, $f(g)$ is a decreasing function. Therefore, for $s \geq 4$ and $1 \leq g \leq s-2$,

$$f(g) \geq f(s-2) = 2^{2s+1} - 2^{s+1} + 2 > 0, \quad (12)$$

which results in a contradiction since $|\widehat{F}_2| \leq (s-g+2)2^g - 1$.

Thus, $V(GEX(s, t)) \neq \widehat{F}_1 \cup \widehat{F}_2$. In addition, an important claim is given as follows.

Claim 25. $GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$ has no isolated node.

By contradiction, suppose that $GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$ has at least one isolated node. Then, we prove that the two cases both contradict the supposition.

Case 1. $g = 1$.

Since $\widehat{F}_1 \neq \widehat{F}_2$, without loss of generality, we suppose that $\widehat{F}_2 - \widehat{F}_1 \neq \emptyset$. When $\widehat{F}_1 \subset \widehat{F}_2$, since \widehat{F}_2 is a 1-GNCF set, $GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$ has no isolated node. Now, we consider $\widehat{F}_1 \not\subset \widehat{F}_2$. The given W is the set of all isolated nodes and $B = GEX(s, t)[V(GEX(s, t)) - (\widehat{F}_1 \cup \widehat{F}_2) - W]$. Since \widehat{F}_1 is a 1-GNCF set, $|N_{GEX(s, t) - \widehat{F}_1}^{\widehat{F}_1}(w)| \geq 1$ for any $w \in W$.

Since \widehat{F}_1 and \widehat{F}_2 are indistinguishable, there exists at most one node $u \in \widehat{F}_2 - \widehat{F}_1$ with u is adjacent to w by Lemma 5. Thereby, there exists only one node $u \in \widehat{F}_2 - \widehat{F}_1$ with u adjacent to w . It is easy to see that there is only one node $v \in \widehat{F}_1 - \widehat{F}_2$ with v adjacent to w . Since $\Delta(GEX(s, t)) = t + 1$, there are at most $t - 1$ neighbors of w in $\widehat{F}_1 \cap \widehat{F}_2$ with any isolated node $w \in W$. Since $|\widehat{F}_2| \leq (s - g + 2)2^g - 1$ and $g = 1$, $|\widehat{F}_2| \leq 2s + 1$. Hence,

$$\begin{aligned} & \sum_{w \in W} \left| N_{GEX(s, t)}^{\widehat{F}_1 \cap \widehat{F}_2}(w) \right| \\ & \leq |W|(t - 1) \leq \sum_{x \in \widehat{F}_1 \cap \widehat{F}_2} \deg_{GEX(s, t)}(x) \\ & \leq |\widehat{F}_1 \cap \widehat{F}_2|(t + 1) \\ & \leq \left(|\widehat{F}_2| - 1 \right)(t + 1) \leq 2s(t + 1). \end{aligned} \quad (13)$$

It follows that $|W| \leq 2s(t + 1)/t - 1 \leq 4s$. Thus,

$$\begin{aligned} |\widehat{F}_1 \cup \widehat{F}_2| + |W| &= |\widehat{F}_1| + |\widehat{F}_2| - |\widehat{F}_1 \cap \widehat{F}_2| + |W| \\ &\leq 2(2s + 1) - (s - 1) + |W| \leq 7s + 3. \end{aligned} \quad (14)$$

Let $f(s) = 2^{2s} - 7s - 3$. We can deduce that $\partial f(s)/\partial s > 0$. Then $f(s)$ is an increasing function. Therefore, $f(s) \geq f(4) > 0$, a contradiction. Thus, $V(B) \neq \emptyset$.

Since the fault-pair $(\widehat{F}_1, \widehat{F}_2)$ does not satisfy Lemma 5 and any node in $V(B)$ is not isolated, there exists no edge between $V(B)$ and $\widehat{F}_1 \Delta \widehat{F}_2$. Moreover, $\widehat{F}_1 \cap \widehat{F}_2$ is also a 1-GNCF set. Thus, $\widehat{F}_1 \cap \widehat{F}_2$ is an R^1 -cut of $GEX(s, t)$. By Theorem 14, $|\widehat{F}_1 \cap \widehat{F}_2| \geq 2s$. Since $|\widehat{F}_1| \leq 2s + 1$, $|\widehat{F}_2| \leq 2s + 1$, and $\widehat{F}_1 \neq \widehat{F}_2$, $|\widehat{F}_1 - \widehat{F}_2| = |\widehat{F}_2 - \widehat{F}_1| = 1$. Let $\widehat{F}_1 - \widehat{F}_2 = \{v\}$ and $\widehat{F}_2 - \widehat{F}_1 = \{u\}$. Then $|N_{GEX(s, t)[\widehat{F}_1 - \widehat{F}_2]}^{\widehat{F}_1}(w)| = |N_{GEX(s, t)[\widehat{F}_2 - \widehat{F}_1]}^{\widehat{F}_2}(w)| = 1$. Hence, we have $(v, w) \in E(GEX(s, t))$ and $(u, w) \in E(GEX(s, t))$ for any isolated node $w \in W$. By Lemma 10, there are at most two common neighbors between any two nodes in $V(X_s)$. In addition, by Definition 11, any two cross edges have no common end node. Then we deduce that any two nodes in $V(GEX(s, t))$ have at most two common neighbors. Thus, $|W| \leq 2$. Since there is no common node between any two cross edges and X_s is triangle-free, $GEX(s, t)$ is triangle-free. Thereby,

$$\begin{aligned} \left| \widehat{F}_1 \cap \widehat{F}_2 \right| &\geq \left| N_{GEX(s, t)}(v) - W \right| + \left| N_{GEX(s, t)}(w) - \{u, v\} \right| \\ &\quad + \left| N_{GEX(s, t)}(v) - W \right| - (2 - |W|) \\ &\geq 2(s + 1 - |W|) + s - 1 - 2 + |W| \\ &= 3s - 1 - |W| \geq 3s - 3. \end{aligned} \quad (15)$$

Therefore, for $s \geq 4$, we have

$$\left| \widehat{F}_2 \right| = \left| \widehat{F}_2 - \widehat{F}_1 \right| + \left| \widehat{F}_1 \cap \widehat{F}_2 \right| \geq 1 + 3s - 3 = 3s - 2 > 2s + 1, \quad (16)$$

which results in a contradiction since $|\widehat{F}_2| \leq 2s + 1$.

Case 2. $g \geq 2$.

Without loss of generality, we suppose that $\widehat{F}_2 - \widehat{F}_1 \neq \emptyset$. Since \widehat{F}_1 is a g -GNCF set of $GEX(s, t)$, $|N_{GEX(s, t) - \widehat{F}_1}^{\widehat{F}_1}(x)| \geq g$ with any node $x \in V(GEX(s, t) - \widehat{F}_1)$. For any $w \in V(GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2)$, there exists at most one neighbor in $\widehat{F}_2 - \widehat{F}_1$ because the fault-pair $(\widehat{F}_1, \widehat{F}_2)$ is indistinguishable by Lemma 5. Therefore, $|N_{GEX(s, t) - (\widehat{F}_1 \cup \widehat{F}_2)}^{\widehat{F}_1}(w)| \geq g - 1 \geq 1$.

Since $w \in V(GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2)$ is arbitrary, every node of $GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$ is not an isolated one.

To sum up, Claim 25 holds.

Since there exists no isolated node in $GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$ by Claim 25 we have, for any $w \in GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$, there exists some node $v \in GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$ such that $(w, v) \in E(GEX(s, t))$. If $(u, w) \in E(GEX(s, t))$ for any $u \in \widehat{F}_1 \Delta \widehat{F}_2$, $(u, v)_w$ satisfies condition in Lemma 5. Therefore, the g -GNCF sets \widehat{F}_1 and \widehat{F}_2 are distinguishable, which results in a contradiction. By the arbitrariness of $w \in GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2$, there exists no edge between $V(GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2)$ and $\widehat{F}_1 \Delta \widehat{F}_2$.

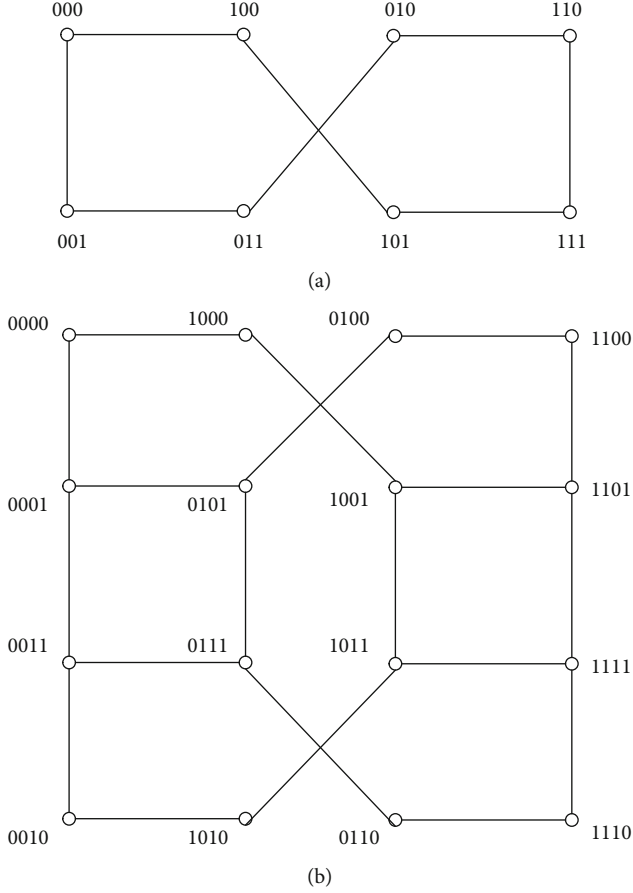
Since \widehat{F}_1 is a g -GNCF set and $\widehat{F}_2 - \widehat{F}_1 \neq \emptyset$, $\delta(GEX(s, t)[\widehat{F}_2 - \widehat{F}_1]) \geq g$. Thus, $|\widehat{F}_2 - \widehat{F}_1| \geq g + 1$. Since \widehat{F}_1 and \widehat{F}_2 are both g -GNCF sets and there exists no edge between $V(GEX(s, t) - \widehat{F}_1 \cup \widehat{F}_2)$ and $\widehat{F}_1 \Delta \widehat{F}_2$, $\widehat{F}_1 \cap \widehat{F}_2$ is also an R^g -cut of $GEX(s, t)$. By Theorem 14, $|\widehat{F}_1 \cap \widehat{F}_2| \geq (s - g + 1)2^g$. Then $|\widehat{F}_2| = |\widehat{F}_2 - \widehat{F}_1| + |\widehat{F}_1 \cap \widehat{F}_2| \geq (s - g + 1)2^g + 2^g = (s - g + 2)2^g$, which contradicts with $|\widehat{F}_2| \leq (s - g + 2)2^g - 1$.

Thus, $t_g^M(GEX(s, t)) \geq (s - g + 2)2^g - 1$ for any integers $1 \leq g \leq s - 2$ and $s \geq 4$.

Hence, the proof of theorem is completed.

5. Applications to a Family of Famous Networks

In Section 2, the definition of the generalized exchanged X -cube $GEX(s, t)$ has been given. Furthermore, we determine the R^g -connectivity and g -GNCD of $GEX(s, t)$ in Section 3 and Section 4, respectively. Applying the theorems of

FIGURE 5: (a) $GEH(1, 1)$, (b) $GEH(1, 2)$.

Section 3 and Section 4, we can directly establish the R^g -connectivity and g -GNCD of some generalized exchanged X -cubes, including generalized exchanged hypercubes, dual-cube-like networks, generalized exchanged crossed cubes, and locally generalized exchanged twisted cubes. In this section, we will give the applications to these networks.

5.1. The Generalized Exchanged Hypercube. In 2005, Loh et al. [36] proposed the exchanged hypercube, which obtained by removing edges from a hypercube H_{s+t+1} . We denote $I_r = \{1, 2, \dots, r\}$, where r is a given position integer. For each $n \in I_r$, the sequence $x_r, x_{r-1} \dots x_1$ is a binary string of length r if $x_n \in \{0, 1\}$. The definition of exchanged hypercubes is presented as follows.

Definition 16 (see [36]). Let $s, t \geq 1$, the exchanged hypercube $EH(s, t)$ consists of the node set $V(EH(s, t))$ and the edge set $E(EH(s, t))$, two nodes $u = u_{s+t} \dots u_{t+1} u_t \dots u_1 u_0$ and $v = v_{s+t} \dots v_{t+1} v_t \dots v_1 v_0$ are linked by an edge, called r -dimensional edge, if and only if the following conditions are satisfied:

- (i) u and v differ exactly in one bit on the r -th bit or on the last bit
- (ii) if $r \in I_t$, then $u_0 = v_0 = 1$

- (iii) $r \in I_{s+t} - I_t$, then $u_0 = v_0 = 0$.

The generalized exchange hypercube was proposed by Cheng et al. [37]. Let $s, t \geq 1$, the generalized exchanged hypercube $GEH(s, t, f)$ consists of two classes of hypercubes: one class contains $2^t H_s$'s, referred to as the Class-0 clusters; and the other contains $2^s H_t$'s, referred to as the Class-1 clusters. Class-0 and Class-1 clusters will be referred to as clusters of opposite class of each other, same class otherwise. The function f is a bijection between nodes of Class-0 clusters and those of Class-1 clusters; for two nodes u, v in the same cluster, $f(u)$ and $f(v)$ are in two different clusters, and the edge $(u, f(u))$ is a cross edge. The bijection f ensures the existence of a perfect matching between nodes of Class-0 clusters and those in the Class-1 clusters but ignores the specifics of the perfect matching. Hence, we present the following proposition.

Proposition 17. $GEH(s, t)$ can be decomposed into two subgraphs \tilde{L}' and \tilde{R}' . Further, \tilde{L}' can be partitioned into 2^t subgraphs, denoted by \tilde{L}_i' for $i = 1, 2, \dots, 2^t$. Similarly, \tilde{R}' can be partitioned into 2^s subgraphs, denoted by \tilde{R}_j' for $j = 1, 2, \dots, 2^s$. And $GEH(s, t)$ satisfies the following conditions (Figure 5 shows the $GEH(1, 1)$ and $GEH(1, 2)$):

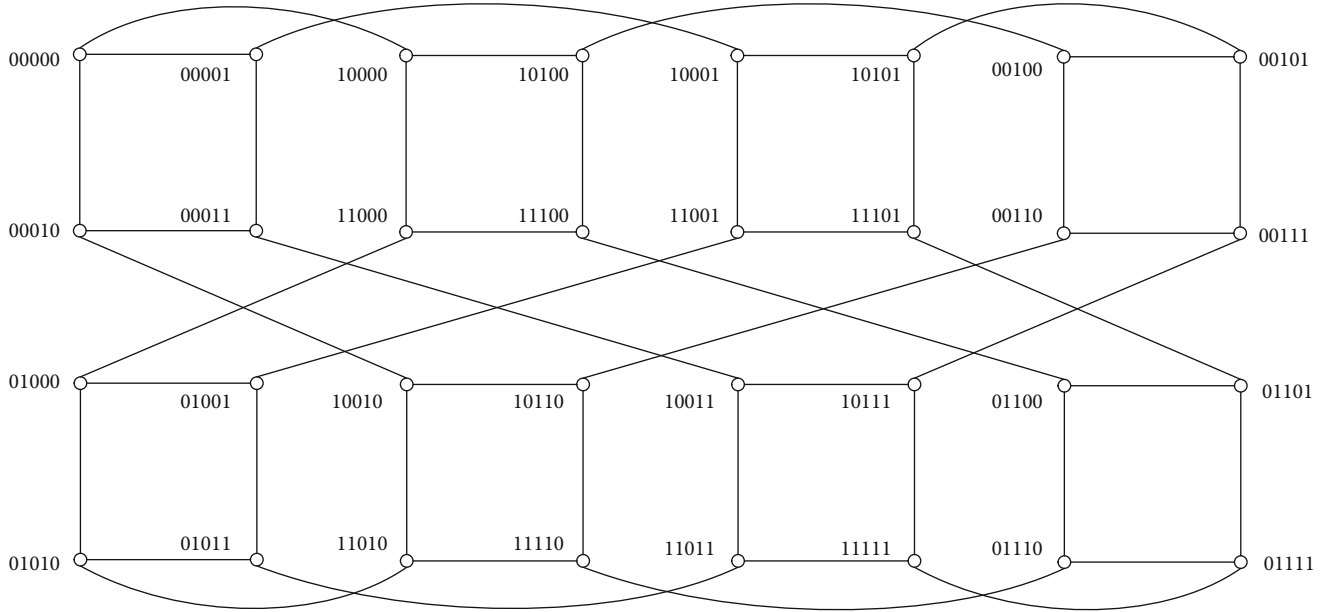
- (a) For any i, j , $\tilde{L}_i' \cong H_s$ and $\tilde{R}_j' \cong H_t$. Further, $|V(\tilde{L}_i')| = 2^s$ and $|V(\tilde{R}_j')| = 2^t$
- (b) Each node in $V(\tilde{L}')$ has a sole neighbor in $V(\tilde{R}')$ and vice versa. In addition, for distinct nodes in each \tilde{L}_i' , their neighbors of \tilde{R}' lie in different \tilde{R}_j'
- (c) For any two different subgraphs \tilde{L}_i' and \tilde{L}_h' with $i \neq h$, there exists no edge between them. Similar for \tilde{R}_j' and \tilde{R}_k' with $j \neq k$.

The dual-cube is a special case of the exchanged hypercube when $s = t$, proposed by Li and Peng [38]. That is, $EH(n, n) \cong D_n$. The dual-cube-like network DC_n [39], which is a generalization of dual-cubes, is isomorphic to $EH(n-1, n-1)$, a special case of $GEH(n-1, n-1)$ (see DC_3 in Figure 6).

By Proposition 17, the generalized exchanged hypercube $GEH(s, t)$ is the member of generalized exchanged X -cubes, where the X -cube is a hypercube. Then, the following theorems hold obviously.

Theorem 18.

- (1) For any integers $1 \leq g \leq s-2$ and $3 \leq s \leq t$, $\kappa^g(GE H(s, t)) = (s-g+1)2^g$
- (2) For any integers $1 \leq g \leq n-3$ and $n \geq 4$, $\kappa^g(DC_n) = (n-g)2^g$.

FIGURE 6: The dual-cube-like network DC_3 .**Theorem 19.**

- (1) For any integers $1 \leq g \leq s-2$ and $3 \leq s \leq t$, $t_g^P(\text{GEH}(s, t)) = (s-g+2)2^g - 1$
- (2) For any integers $1 \leq g \leq n-3$ and $n \geq 4$, $t_g^P(\text{DC}_n) = (n-g+1)2^g - 1$.

Theorem 20.

- (1) For any integers $1 \leq g \leq s-2$ and $4 \leq s \leq t$, $t_g^M(\text{GEH}(s, t)) = (s-g+2)2^g - 1$
- (2) For any integers $1 \leq g \leq n-3$ and $n \geq 5$, $t_g^M(\text{DC}_n) = (n-g+1)2^g - 1$.

5.2. *The Generalized Exchanged Crossed Cube.* Li et al. [26] give the definition of $ECQ(s, t)$, which is obtained by removing edges from a crossed cube CQ_{s+t+1} . In what follows, we review the definition of exchanged crossed cubes.

Definition 21 (see [26]). The (s, t) -dimensional exchanged crossed cube is defined as a graph $ECQ(s, t) = G(V(ECQ(s, t)), E(ECQ(s, t)))$ for $s, t \geq 1$. The node set $V(ECQ(s, t)) = \{a_{s-1} \cdots a_0 b_{t-1} \cdots b_0 c | a_j, b_j, c \in \{0, 1\}, \text{ where } 0 \leq i \leq t-1 \text{ and } 0 \leq j \leq s-1\}$. The edge set $E(ECQ(s, t))$ consisting of three types of disjoint sets E_1, E_2 , and E_3 is shown as follows.

$E_1 : |u[0] \neq v[0], u \oplus v = 1$, where \oplus is the exclusive-OR operator.

$E_2 : |u[s+t:t+1] = v[s+t:t+1], u[0] = v[0] = 1, u[t:1]$ is denoted by $b = b_{t-1} \cdots b_0$ and $v[t:1]$ is denoted by $b' = b'_{t-1} \cdots b'_0$. And u and v are adjacent by the following rule: for any integer $t \geq 1$, if and only if there is an $l(1 \leq l \leq t)$ with

$b_{t-1} \cdots b_l = b'_{t-1} \cdots b'_l; b_{l-1} \neq b'_{l-1}, b_{l-2} = b'_{l-2}$ if l is even; $b_{2i+1}b_{2i} \sim b'_{2i+1}b'_{2i}$, where $0 \leq i < \lfloor (l-1)/2 \rfloor$
 $E_3 : |u[t:1] = v[t:1], u[0] = v[0] = 0, u[s+t:t+1]$ is denoted by $a = a_{s-1} \cdots a_0$ and $y[s+t:t+1]$ is denoted by $a' = a'_{t-1} \cdots a'_0$. And u and v are adjacent by the following rule: for any integer $s \geq 1$, if and only if there is an $l(1 \leq l \leq s)$ with $a_{s-1} \cdots a_l = a'_{s-1} \cdots a'_l; a_{l-1} \neq a'_{l-1}, a_{l-2} = a'_{l-2}$ if l is even; $a_{2i+1}a_{2i} \sim a'_{2i+1}a'_{2i}$, where $0 \leq i < \lfloor (l-1)/2 \rfloor$.
 $x[u:v]$ is the bit pattern of x from dimension u to dimension v .

Let $s, t \geq 1$, the generalized crossed cube $GECQ(s, t, f)$ comprises two classes of crossed cubes, referred to as the Class-0 clusters and the Class-1 clusters, respectively. The Class-0 clusters contain 2^t CQ_s 's and the Class-1 clusters contain 2^s CQ_t 's. They will be referred to as clusters of opposite class of each other, same class otherwise. The function f is a bijection between nodes of Class-0 clusters and those of Class-1 clusters such that, for u, v , two nodes of the same cluster, $f(u)$ and $f(v)$, are in two different clusters, and the edge $(u, f(u))$ is a cross edge. The bijection f ensures the existence of a perfect matching between two nodes in different clusters, but there is no requirement for the specifics of the perfect matching. Therefore, we have the following proposition.

Proposition 22. $GECQ(s, t)$ can be decomposed into two disjoint subgraphs \tilde{L}' and \tilde{R}' . And \tilde{L}' and \tilde{R}' are the subgraphs induced by $V(\tilde{L}')$ and $V(\tilde{R}')$, respectively, where

$$V(\tilde{L}') = \{a_{s-1}a_{s-2} \cdots a_0 b_{t-1} \cdots b_0 0 | a_j, b_i \in \{0, 1\}, \quad (17)$$

with $0 \leq j \leq s-1$ and $0 \leq i \leq t-1$.

$$V(\widetilde{R}') = \{a_{s-1}a_{s-2} \cdots a_0 b_{t-1} \cdots b_0 I | a_j, b_i \in \{0, 1\}, \quad (18)$$

with $0 \leq j \leq s-1$ and $0 \leq i \leq t-1$.

By Definition 21, \widetilde{L}' can be partitioned into 2^t subgraphs, denoted by \widetilde{L}_i' such that for $v_1, v_2 \in \widetilde{L}_i'$, $v_1[t:1] = v_2[t:1]$, where $i = 1, 2, \dots, 2^t$. Similarly, \widetilde{R}' can be partitioned into 2^s subgraphs, denoted by \widetilde{R}_j' such that $w_1, w_2 \in \widetilde{R}_j'$, $w_1[t+1:s+t] = w_2[t+1:s+t]$, for $j = 1, 2, \dots, 2^s$. And $GECQ(s, t)$ satisfies the following conditions (see $GECQ(1, 3)$ in Figure 7):

- (1) For any i, j , $\widetilde{L}_i' \cong CQ_s$ and $\widetilde{R}_j' \cong CQ_t$. Further, $|V(\widetilde{L}_i')| = 2^s$ and $|V(\widetilde{R}_j')| = 2^t$
- (2) Each node in $V(\widetilde{L}')$ has a sole neighbor in $V(\widetilde{R}')$ and vice versa. In addition, for distinct nodes in each \widetilde{L}_i' , their neighbors of \widetilde{R}' lie in different \widetilde{R}_j'
- (3) For any two different subgraphs \widetilde{L}_i' and \widetilde{L}_h' with $i \neq h$, there exists no edge between them. Similar for \widetilde{R}_j' and \widetilde{R}_k' with $j \neq k$.

By Proposition 22, the exchanged crossed cube $GECQ(s, t)$ is an exchanged X -cube, where the X -cube is a crossed cube. Then, the following theorems hold obviously.

Theorem 23. For any integers $1 \leq g \leq s-2$ and $3 \leq s \leq t$, $\kappa^g(GECQ(s, t)) = (s-g+1)2^g$.

Theorem 24.

- (1) For any integers $1 \leq g \leq s-2$ and $3 \leq s \leq t$, $t_g^P(GECQ(s, t)) = (s-g+2)2^g - 1$
- (2) For any integers $1 \leq g \leq s-2$ and $4 \leq s \leq t$, $t_g^M(GECQ(s, t)) = (s-g+2)2^g - 1$.

5.3. The Locally Generalized Exchanged Twisted Cube. The locally exchanged twisted cube proposed by Chang et al. [29], obtained by removing edges from a locally twisted cube LTQ_{s+t+1} . The definition of locally exchanged twisted cube is introduced as follows.

Definition 25 (see [29]). The (s, t) -dimensional locally exchanged twisted cube is defined as a graph $LETQ(s, t) = G(V(LETQ(s, t)), E(LETQ(s, t)))$ for $s, t \geq 1$. The node set $V(LETQ(s, t)) = \{x = x_{t+s} \cdots x_{t+1} x_t \cdots x_1 x_0 : x_i \in \{0, 1\} \text{ with } 0 \leq i \leq t+s\}$. $E(LETQ(s, t))$ is the edge set consisting of the following three types of disjoint sets E_1, E_2 , and E_3 .

$$\begin{aligned} E_1 &= \{(x, y) \in V \times V : x \oplus y = 2^0\} \\ E_2 &= \{(x, y) \in V \times V : x_0 = y_0 = 1, x_1 = y_1 = 0 \text{ and } x \oplus y = 2^k \text{ for } k \in [3, t]\} \cup \{(x, y) \in V \times V : x_0 = y_0 = x_1 = y_1 = 1 \text{ and } x \oplus y = 2^k + 2^{k-1} \text{ for } k \in [3, t]\} \\ &\quad \cup \{(x, y) \in V \times V : x_0 = y_0 = 1 \text{ and } x \oplus y \in \{2^1, 2^2\}\} \\ E_3 &= \{(x, y) \in V \times V : x_0 = y_0 = x_{t+1} = y_{t+1} = 0 \text{ and } x \oplus y = 2^k \text{ for } k \in [t+3, t+s]\} \\ &\quad \cup \{(x, y) \in V \times V : x_0 = y_0 = 0, x_{t+1} = y_{t+1} = 1 \text{ and } x \oplus y = 2^k + 2^{k-1} \text{ for } k \in [t+3, t+s]\} \\ &\quad \cup \{(x, y) \in V \times V : x_0 = y_0 = 0 \text{ and } x \oplus y \in \{2^{t+1}, 2^{t+2}\}\} \end{aligned} \quad (19)$$

Let $s, t \geq 1$; there are two classes of locally twisted cubes in the locally generalized exchanged twisted cube $LGETQ(s, t, f)$: one class, referred to as the Class-0 clusters, contains 2^t LTQ_s 's; and the other, referred to as the Class-1 clusters, contains 2^s LTQ_t 's. They will be referred to as clusters of opposite class of each other, same class otherwise. There exists a bijection function f between nodes of Class-0 clusters and those of Class-1 clusters. For two nodes u, v in the same cluster, $f(u)$ and $f(v)$ belong to two different ones, and the edge $(u, f(u))$ is a cross edge. The bijection f ensures the existence of a perfect matching between nodes of Class-0 clusters and those in the Class-1 clusters, but the specifics of the perfect matching can be ignored. Further, we obtain the proposition as follows.

Proposition 26. $LGETQ(s, t)$ can be decomposed into two disjoint subgraphs \widetilde{L}' and \widetilde{R}' . \widetilde{L}' can be partitioned into 2^t subgraphs, denoted by \widetilde{L}_i' for $i = 1, 2, \dots, 2^t$. Similarly, \widetilde{R}' can be partitioned into 2^s subgraphs, denoted by \widetilde{R}_j' for $j = 1, 2, \dots, 2^s$. And $LGETQ(s, t)$ satisfies the following conditions (see $LGETQ(1, 3)$ in Figure 8):

- (a) For any i, j , $\widetilde{L}_i' \cong LTQ_s$ and $\widetilde{R}_j' \cong LTQ_t$. Further, $|V(\widetilde{L}_i')| = 2^s$ and $|V(\widetilde{R}_j')| = 2^t$
- (b) Each node in $V(\widetilde{L}')$ has a sole neighbor in $V(\widetilde{R}')$ and vice versa. In addition, for distinct nodes in each \widetilde{L}_i' , their neighbors of \widetilde{R}' lie in different \widetilde{R}_j'
- (c) For any two different subgraphs \widetilde{L}_i' and \widetilde{L}_h' with $i \neq h$, there exists no edge connects them. Similar for \widetilde{R}_j' and \widetilde{R}_k' with $j \neq k$.

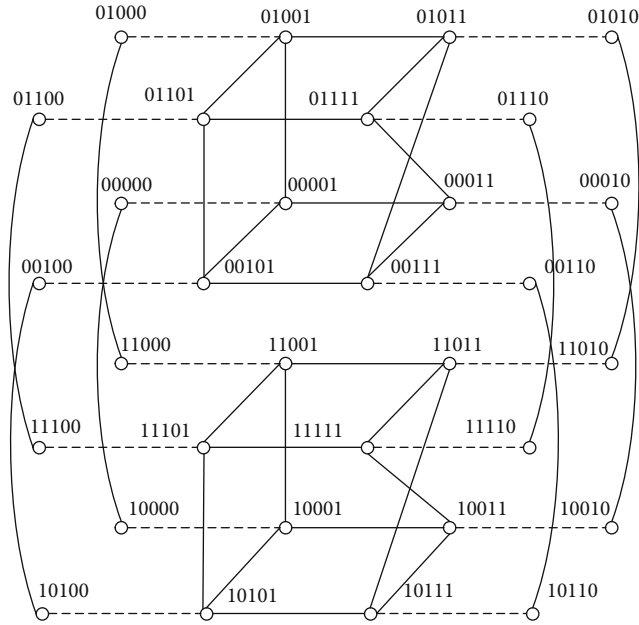


FIGURE 7: An exchanged crossed cube $GECQ(1, 3)$.

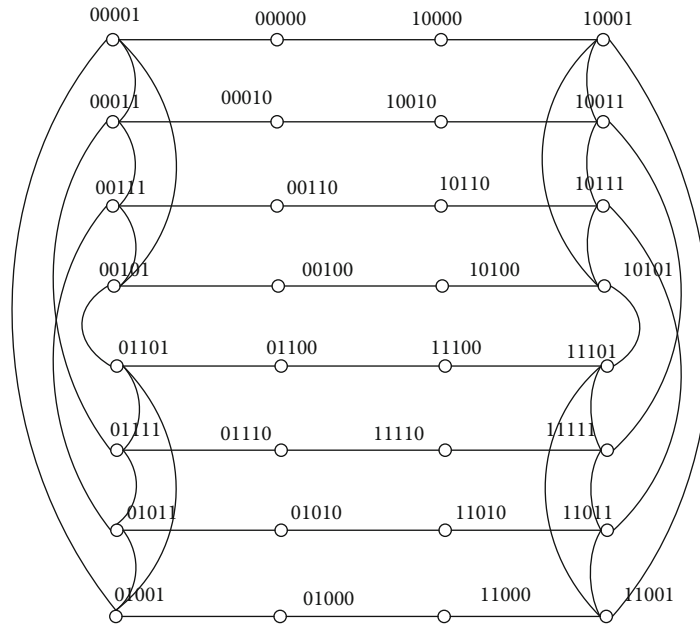


FIGURE 8: A locally exchanged twisted cube $LGETQ(1, 3)$.

By Proposition 26, the locally exchanged twisted cube $LGETQ(s, t)$ is a member of generalized exchanged X -cubes, where the X -cube is a locally twisted cube. Then, we have the following theorems.

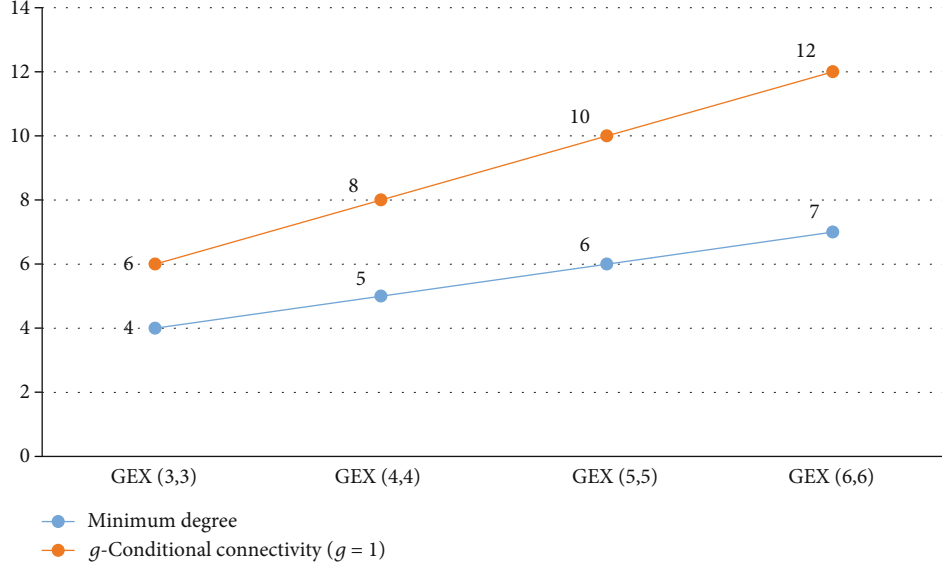
Theorem 27.

- (1) For any integers $1 \leq g \leq s - 2$ and $3 \leq s \leq t$, $\kappa^g(LGETQ(s, t)) = (s - g + 1)2^g$
- (2) For any integers $1 \leq g \leq s - 2$ and $3 \leq s \leq t$, $t_g^P(LGETQ(s, t)) = (s - g + 2)2^g - 1$

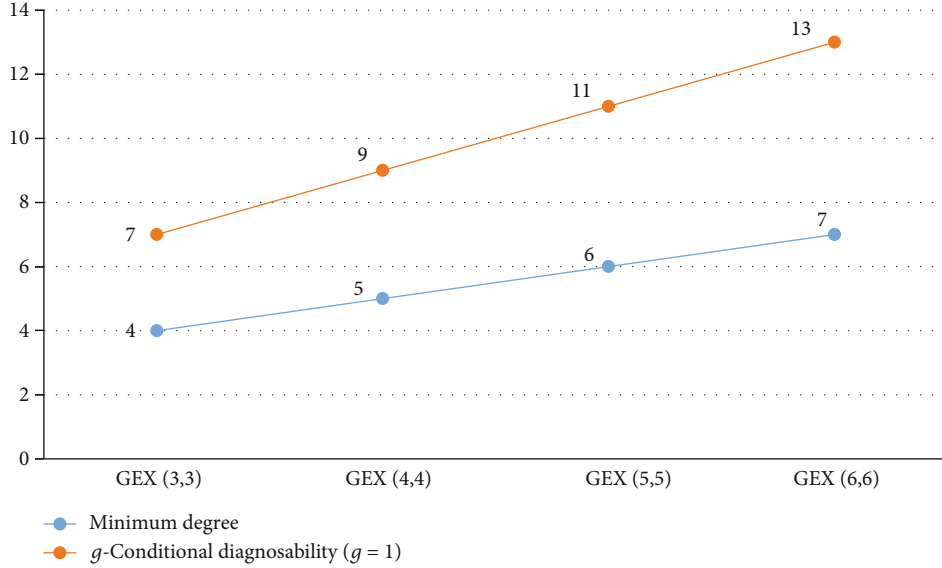
- (3) For any integers $1 \leq g \leq s - 2$ and $4 \leq s \leq t$, $t_g^M(LGETQ(s, t)) = (s - g + 2)2^g - 1$.

6. Compare Results

In this section, we will illustrate the advantages of R^g -connectivity and g -GNCD compared to traditional connectivity and traditional diagnosability, respectively. Let us review their definition. The *connectivity*, which is less than the minimum degree of graph, is the minimum number of nodes that make the graph disconnected. The maximum number of faulty processors that the system can precisely point out is



(a)



(b)

FIGURE 9: (a) The minimum degree and R^g -connectivity of $GEX(s, t)$. (b) The minimum degree and g -GNCD of $GEX(s, t)$.

as known as the *diagnosability* of the system, which is equal to the minimum degree of graph in most cases. The g -good-neighbor conditional connectivity (R^g -connectivity) is the minimum number of nodes that make the graph disconnected, and each node has at least g neighbors in every remaining component. The g -good-neighbor conditional diagnosability (g -GNCD) is the maximum number of faulty processors that can be identified under the condition that every fault-free processor has no less than g fault-free neighbors. We have determined that the R^g -connectivity of $GEX(s, t)$ is $(s - g + 1)2^g$ and the g -GNCD of $GEX(s, t)$ is $(s - g + 2)2^g - 1$. Figure 9 shows that R^g -connectivity and g -GNCD are both about 2^g times the minimum degree of graph. Therefore, we can speculate that R^g -connectivity is about 2^g times traditional connectivity and g -GNCD is about 2^g times traditional diagnosability, which means that R^g -

connectivity and g -GNCD can better evaluate the fault tolerance of network.

7. Conclusion

The R^g -connectivity and g -GNCD are two significant metrics for reliability of multiprocessor systems. Exchanged X -cubes are a class of irregular networks, obtained by deleting links from hypercubes and some variant networks of hypercubes (X -cubes). They not only combine the advantages of X -cubes but also reduce the interconnection complexity. Exchanged X -cubes classify its nodes into two different classes clusters with a unique connecting rule. In this paper, we propose the generalized exchanged X -cubes framework so that architecture can be constructed by different connecting rules. We first give the definition of a family of

generalized exchanged X -cubes, including generalized exchanged hypercubes, dual-cube-like networks, generalized exchanged crossed cubes, and locally exchanged twisted cubes as members. Then we determine the R^g -connectivity and g -GNCD of generalized exchanged X -cubes. Finally, the R^g -connectivity and g -GNCD of generalized exchanged hypercubes, dual-cube-like networks, generalized exchanged crossed cubes, and locally exchanged twisted cubes are established directly. As a future research, we attempt to evaluate the R^g -connectivity and g -GNCD of other generalized exchanged X -cubes using methods extended from the proposed method in this paper.

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 have declared that no conflict of interest exists.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61872257 and No. 61977016) and the Joint Found of the National Natural Science Foundation of China (No. U1905211).

References

- [1] Y. Hong, "Realizability of fault-tolerant graphs," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 39, no. 2, pp. 619–631, 2016.
- [2] W. Guo, J. Li, G. Chen, Y. Niu, and C. Chen, "A PSO-optimized real-time fault-tolerant task allocation algorithm in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 12, pp. 3236–3249, 2015.
- [3] A.-H. Esfahanian and S. L. Hakimi, "On computing a conditional edge-connectivity of a graph," *Information Processing Letters*, vol. 27, no. 4, pp. 195–199, 1988.
- [4] S. Latifi, M. Hegde, and M. Naraghi-Pour, "Conditional connectivity measures for large multiprocessor systems," *IEEE Transactions on Computers*, vol. 43, no. 2, pp. 218–222, 1994.
- [5] X. Xu, S. Zhou, and J. Li, "Reliability of complete cubic networks under the condition of g -good-neighbor," *The Computer Journal*, vol. 26, pp. 1165–1177, 2016.
- [6] W. Ning, "Erratum and corrections to "The h -connectivity of exchanged crossed cube" [Theoret. Comput. Sci. 696 (2017) 65–68]," *Theoretical Computer Science*, vol. 705, pp. 118–121, 2018.
- [7] J. Yuan, A. Liu, X. Ma, X. Liu, X. Qin, and J. Zhang, "The g -good-neighbor conditional diagnosability of k -ary n -cubes under the PMC model and MM^* model," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 1165–1177, 2015.
- [8] L. Lin and S. Zhou, "Conditional connectivity for (n, k) -arrangement graphs," *Journal of Mathematical Study*, vol. 45, no. 4, pp. 350–364, 2012.
- [9] F. P. Preparata, G. Metze, and R. T. Chien, "On the connection assignment problem of diagnosable systems," *IEEE Transactions on Computers*, vol. EC-16, no. 6, pp. 848–854, 1967.
- [10] J. Maeng and M. Malek, "A comparison connection assignment for self-diagnosis of multiprocessors systems," in *Proceedings 11th International Symposium. Fault-Tolerant Computing*, pp. 173–175, Portland, Maine, June 1981.
- [11] A. Sengupta and A. Dahbura, "On self-diagnosable multiprocessor systems: diagnosis by the comparison approach," *IEEE Transactions on Computers*, vol. 41, no. 11, pp. 1386–1396, 1992.
- [12] P.-L. Lai, J. J. M. Tan, C.-P. Chang, and L.-H. Hsu, "Conditional diagnosability measures for large multiprocessor systems," *IEEE Transactions on Computers*, vol. 54, no. 2, pp. 165–175, 2005.
- [13] S.-L. Peng, C.-K. Lin, J. J. M. Tan, and L.-H. Hsu, "The g -good-neighbor conditional diagnosability of hypercube under PMC model," *Applied Mathematics and Computation*, vol. 218, no. 21, pp. 10406–10412, 2012.
- [14] S. Wang and W. Han, "The g -good-neighbor conditional diagnosability of n -dimensional hypercubes under the MM^* model," *Information Processing Letters*, vol. 116, no. 9, pp. 574–577, 2016.
- [15] X. Li, Y.-H. Teng, T.-L. Kung, Q. Chen, and C.-K. Lin, "The diagnosability and 1-good-neighbor conditional diagnosability of hypercubes with missing links and broken-down nodes," *Information Processing Letters*, vol. 146, pp. 20–26, 2019.
- [16] L. Lin, L. Xu, D. Wang, and S. Zhou, "The g -good-neighbor conditional diagnosability of arrangement graphs," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 3, pp. 542–548, 2018.
- [17] J. Guo, D. Li, and M. Lu, "The g -good-neighbor conditional diagnosability of the crossed cubes under the PMC and MM^* model," *Theoretical Computer Science*, vol. 755, pp. 81–88, 2019.
- [18] X. Li, J. Fan, C.-K. Lin, and X. Jia, "Diagnosability evaluation of the data center network DCell," *The Computer Journal*, vol. 61, no. 1, pp. 129–143, 2018.
- [19] J. Xu, A. Kumar, and X. Yu, "On the fundamental tradeoffs between routing table size and network diameter in Peer-to-Peer networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 1, pp. 151–163, 2004.
- [20] Y. Wei and M. Xu, "The 1, 2-good-neighbor conditional diagnosabilities of regular graphs," *Applied Mathematics and Computation*, vol. 334, pp. 295–310, 2018.
- [21] X. Hu, W. Yang, Y. Tian, and J. Meng, "Equal relation between g -good-neighbor diagnosability under the PMC model and g -good-neighbor diagnosability under the MM^* model of a graph," *Discrete Applied Mathematics*, vol. 262, pp. 96–103, 2019.
- [22] L. Lin, S.-Y. Hsieh, R. Chen, L. Xu, and C.-W. Lee, "The relationship between g -restricted connectivity and g -good-neighbor fault diagnosability of general regular networks," *IEEE Transactions on Reliability*, vol. 67, no. 1, pp. 285–296, 2018.
- [23] D. Cheng, "A relationship between g -good-neighbor conditional diagnosability and g -good-neighbor connectivity in regular graphs," *International Journal of Computer Mathematics*, vol. 3, no. 1, pp. 47–52, 2018.
- [24] E. Cheng, K. Qiu, and Z. Shen, "A general approach to deriving the g -good-neighbor conditional diagnosability of

- interconnection networks,” *Theoretical Computer Science*, vol. 757, pp. 56–67, 2019.
- [25] L.-H. Hsu and C.-K. Lin, *Graph Theory and Interconnection Networks*, CRC Press, Boca Raton, 2008.
- [26] K. Li, Y. Mu, K. Li, and G. Min, “Exchanged crossed cube: a novel interconnection network for parallel computation,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 11, pp. 2211–2219, 2013.
- [27] M. Fiedler, “Algebraic connectivity of graphs,” *Czechoslovak Mathematical Journal*, vol. 23, no. 2, pp. 298–305, 1973.
- [28] Dahbura and Masson, “An $O(n^{2.5})$ fault identification algorithm for diagnosable systems,” *IEEE Transactions on Computers*, vol. C-33, no. 6, pp. 486–492, 1984.
- [29] J.-M. Chang, X.-R. Chen, J.-S. Yang, and R.-Y. Wu, “Locally exchanged twisted cubes: connectivity and super connectivity,” *Information Processing Letters*, vol. 116, no. 7, pp. 460–466, 2016.
- [30] J. Fan, “Diagnosability of the Mobius cubes,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 9, no. 9, pp. 923–928, 1998.
- [31] D. Wang, “On embedding hamiltonian cycles in crossed cubes,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 3, pp. 334–346, 2008.
- [32] X. Yang, D. J. Evans, and G. M. Megson, “The locally twisted cubes,” *International Journal of Computer Mathematics*, vol. 82, no. 4, pp. 401–413, 2005.
- [33] X.-J. Li and J.-M. Xu, “Edge-fault tolerance of hypercube-like networks,” *Information Processing Letters*, vol. 113, no. 19–21, pp. 760–763, 2013.
- [34] L. Ye and J. Liang, “On conditional h -vertex connectivity of some networks,” *Chinese Journal of Electronics*, vol. 25, no. 3, pp. 556–560, 2016.
- [35] Q. Zhu, “On conditional diagnosability and reliability of the BC networks,” *The Journal of Supercomputing*, vol. 45, no. 2, pp. 173–184, 2008.
- [36] P. K. K. Loh, W. J. Hsu, and Y. Pan, “The exchanged hypercube,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 16, no. 9, pp. 866–874, 2005.
- [37] E. Cheng, K. Qiu, and Z. Shen, “Structural properties of generalized exchanged hypercubes,” in *Emergent Computation, Emergence, Complexity and Computation*, A. Adamatzky, Ed., Springer, Cham, 2017.
- [38] Y. Li and S. Peng, “Dual-cubes: a new interconnection network for high-performance computer clusters,” in *Proceedings of the 2000 international computer symposium, workshop on computer architecture*, pp. 51–57, ChiaYi, Taiwan, December 2000.
- [39] A. Angjeli, E. Cheng, and L. Lipták, “Linearly many faults in dual-cube-like networks,” *Theoretical Computer Science*, vol. 472, pp. 1–8, 2013.
- [40] S. Wang and M. Wang, “The g -good-neighbor and g -extra diagnosability of networks,” *Theoretical Computer Science*, vol. 773, pp. 107–114, 2019.