

Security and Privacy Challenges for Intelligent Internet of Things Devices

Lead Guest Editor: AnMin Fu

Guest Editors: Shui Yu, Jinguang Han, and Jingyu Feng





Security and Privacy Challenges for Intelligent Internet of Things Devices

Security and Communication Networks

Security and Privacy Challenges for Intelligent Internet of Things Devices

Lead Guest Editor: AnMin Fu

Guest Editors: Shui Yu, Jinguang Han, and Jingyu
Feng






Copyright © 2021 Hindawi Limited. All rights reserved.

This is a special issue published in "Security and Communication Networks." 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

Roberto Di Pietro, Saudi Arabia

Associate Editors

Jiankun Hu , Australia
Emanuele Maiorana , Italy
David Megias , Spain
Zheng Yan , China

Academic Editors

Saed Saleh Al Rabae , United Arab Emirates
Shadab Alam, Saudi Arabia
Goutham Reddy Alavalapati , USA
Jehad Ali , Republic of Korea
Jehad Ali, Saint Vincent and the Grenadines
Benjamin Aziz , United Kingdom
Taimur Bakhshi , United Kingdom
Spiridon Bakiras , Qatar
Musa Balta, Turkey
Jin Wook Byun , Republic of Korea
Bruno Carpentieri , Italy
Luigi Catuogno , Italy
Ricardo Chaves , Portugal
Chien-Ming Chen , China
Tom Chen , United Kingdom
Stelvio Cimato , Italy
Vincenzo Conti , Italy
Luigi Coppolino , Italy
Salvatore D'Antonio , Italy
Juhriyansyah Dalle, Indonesia
Alfredo De Santis, Italy
Angel M. Del Rey , Spain
Roberto Di Pietro , France
Wenxiu Ding , China
Nicola Dragoni , Denmark
Wei Feng , China
Carmen Fernandez-Gago, Spain
AnMin Fu , China
Clemente Galdi , Italy
Dimitrios Geneiatakis , Italy
Muhammad A. Gondal , Oman
Francesco Gringoli , Italy
Biao Han , China
Jinguang Han , China
Khizar Hayat, Oman
Azeem Irshad, Pakistan

M.A. Jabbar , India
Minho Jo , Republic of Korea
Arijit Karati , Taiwan
ASM Kayes , Australia
Farrukh Aslam Khan , Saudi Arabia
Fazlullah Khan , Pakistan
Kiseon Kim , Republic of Korea
Mehmet Zeki Konyar, Turkey
Sanjeev Kumar, USA
Hyun Kwon, Republic of Korea
Maryline Laurent , France
Jegatha Deborah Lazarus , India
Huaizhi Li , USA
Jiguo Li , China
Xueqin Liang, Finland
Zhe Liu, Canada
Guangchi Liu , USA
Flavio Lombardi , Italy
Yang Lu, China
Vincente Martin, Spain
Weizhi Meng , Denmark
Andrea Michienzi , Italy
Laura Mongioi , Italy
Raul Monroy , Mexico
Naghme Moradpoor , United Kingdom
Leonardo Mostarda , Italy
Mohamed Nassar , Lebanon
Qiang Ni, United Kingdom
Mahmood Niazi , Saudi Arabia
Vincent O. Nyangaresi, Kenya
Lu Ou , China
Hyun-A Park, Republic of Korea
A. Peinado , Spain
Gerardo Pelosi , Italy
Gregorio Martinez Perez , Spain
Pedro Peris-Lopez , Spain
Carla Ràfols, Germany
Francesco Regazzoni, Switzerland
Abdalhossein Rezai , Iran
Helena Rifà-Pous , Spain
Arun Kumar Sangaiah, India
Nadeem Sarwar, Pakistan
Neetesh Saxena, United Kingdom
Savio Sciancalepore , The Netherlands


De Rosal Ignatius Moses Setiadi ,
Indonesia
Wenbo Shi, China
Ghanshyam Singh , South Africa
Vasco Soares, Portugal
Salvatore Sorce , Italy
Abdulhamit Subasi, Saudi Arabia
Zhiyuan Tan , United Kingdom
Keke Tang , China
Je Sen Teh , Australia
Bohui Wang, China
Guojun Wang, China
Jinwei Wang , China
Qichun Wang , China
Hu Xiong , China
Chang Xu , China
Xuehu Yan , China
Anjia Yang , China
Jiachen Yang , China
Yu Yao , China
Yinghui Ye, China
Kuo-Hui Yeh , Taiwan
Yong Yu , China
Xiaohui Yuan , USA
Sherali Zeadally, USA
Leo Y. Zhang, Australia
Tao Zhang, China
Youwen Zhu , China
Zhengyu Zhu , China

Contents

Secure Multi-Keyword Search and Access Control over Electronic Health Records in Wireless Body Area Networks

Yong Ding , Hui Xu, Yujue Wang, Fang Yuan, and Hai Liang 
Research Article (11 pages), Article ID 9520941, Volume 2021 (2021)






Privacy-Preserving Redactable Blockchain for Internet of Things

Yanli Ren , Xianji Cai, and Mingqi Hu
Research Article (12 pages), Article ID 4485311, Volume 2021 (2021)


A Fine-Grained IoT Data Access Control Scheme Combining Attribute-Based Encryption and Blockchain

Xiaofeng Lu , Songbing Fu , Cheng Jiang , and Pietro Lio 
Research Article (13 pages), Article ID 5308206, Volume 2021 (2021)


Attribution Classification Method of APT Malware in IoT Using Machine Learning Techniques

Shudong Li , Qianqing Zhang , Xiaobo Wu , Weihong Han , and Zhihong Tian 
Research Article (12 pages), Article ID 9396141, Volume 2021 (2021)

V-LDAA: A New Lattice-Based Direct Anonymous Attestation Scheme for VANETs System

Liquan Chen , Tianyang Tu, Kunliang Yu, Mengnan Zhao, and Yingchao Wang
Research Article (13 pages), Article ID 4660875, Volume 2021 (2021)





A LoRa-Based Lightweight Secure Access Enhancement System

Yu Jiang , Hua Fu, Aiqun Hu, and Wen Sun
Research Article (16 pages), Article ID 3530509, Volume 2021 (2021)


ICSTrace: A Malicious IP Traceback Model for Attacking Data of the Industrial Control System

Feng Xiao , Enhong Chen, Qiang Xu, and Xianguo Zhang
Research Article (14 pages), Article ID 7525092, Volume 2021 (2021)


GNFCVulFinder: NDEF Vulnerability Discovering for NFC-Enabled Smart Mobile Devices Based on Fuzzing

Zhiqiang Wang , Yuheng Lin , Zihan Zhuo , Jieming Gu , and Tao Yang
Research Article (14 pages), Article ID 9946022, Volume 2021 (2021)

Modeling and Control of Malware Propagation in Wireless IoT Networks




Qing Yan, Lipeng Song , Chenlu Zhang, Jing Li, and Shanshan Feng
Research Article (13 pages), Article ID 4133474, Volume 2021 (2021)

An Adaptive Industrial Control Equipment Safety Fault Diagnosis Method in Industrial Internet of Things

Hanrui Zhang , Qianmu Li , Shunmei Meng, Zhuoran Xu, and Chaoxian Lv
Research Article (8 pages), Article ID 5562275, Volume 2021 (2021)



Revisiting a Multifactor Authentication Scheme in Industrial IoT

Ding Wang , Shuhong Hong , and Qingxuan Wang 

Research Article (7 pages), Article ID 9995832, Volume 2021 (2021)

Research Article

Secure Multi-Keyword Search and Access Control over Electronic Health Records in Wireless Body Area Networks

Yong Ding ^{1,2} Hui Xu,³ Yujue Wang,⁴ Fang Yuan,⁵ and Hai Liang ¹

¹Guangxi Key Laboratory of Cryptography and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

²Cyberspace Security Research Center, Peng Cheng Laboratory, Shenzhen 518055, China

³School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin 541004, China

⁴Hangzhou Innovation Institute, Beihang University, Hangzhou 310052, China

⁵Foreign Affairs and Communications Administration, Beijing 100045, China

Correspondence should be addressed to Hai Liang; lianghai@guet.edu.cn

Received 3 June 2021; Revised 21 July 2021; Accepted 19 August 2021; Published 29 September 2021

Academic Editor: Shui Yu

Copyright © 2021 Yong Ding 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 body area network (WBAN) consists of a number of sensors that are worn on patients to collect dynamic e-health records (EHRs) and mobile devices that aggregate EHRs. These EHRs are encrypted at mobile devices and then uploaded to the public cloud for storage and user access. To share encrypted EHRs with users effectively, help users retrieve EHRs accurately, and ensure EHRs confidentiality, a secure multi-keyword search and access control (SMKS-AC) scheme is proposed, which implements encrypted EHRs access control under the ciphertext-policy attribute-based encryption (CP-ABE). SMKS-AC provides multi-keyword search for accurate EHRs retrieval, supports the validation of decrypted EHRs, and traces and revokes the identity of users who leak private key. Security analysis shows that SMKS-AC is secure against chosen keyword and chosen plaintext attacks. Through theoretical analysis and experimental verification, the proposed SMKS-AC scheme requires less storage resources and computational costs on mobile devices than existing schemes.

1. Introduction

Wireless body area network (WBAN) consists of sensors and mobile devices. Sensors are used to monitor and collect patients's medical/health data. These data are aggregated on mobile devices, then transmitted to public cloud for storage, and shared with data users. However, due to the limited storage space, computing resources, and energy on mobile devices, it is important to keep computing lightweight on mobile devices. Besides, data security and privacy [1] are another important concern in WBAN, including the security of data transmission and storage [2] and access control at the user side [3].

In WBAN, electronic health records (EHRs) are outsourced to public cloud which cannot be directly controlled by the data owner. To ensure the privacy and security of data storage, data owners usually store EHRs in an encrypted

form in the cloud. However, the encrypted data should be shared with data users according to some access control policy. Attribute-based encryption (ABE) is an effective mechanism for fine-grained access control of data. In addition, when data users want to access the encrypted EHRs stored in the public cloud, they can retrieve the required data according to some keywords. Then, the data can be obtained through decryption. Although the single keyword search method can retrieve encrypted data, the search results may contain a large number of irrelevant data, which affects the retrieval accuracy. Therefore, an effective multi-keyword search on encrypted data and the validation of the correctness of decrypted data meet more practical needs.

The EHRs are not only highly private but also commercially valuable. These may promote authorized users to sell their private key for huge profits, so it is necessary to track the identity of the private key owners. Moreover, when

the private key of a data user is sold or compromised, the user's access authority shall be revoked. The decrypted data are able to be forged during transmission, and data users cannot recognize the forged data. Hence, it is imperative to verify the correctness of the data.

1.1. Our Contributions. Inspired by the LiST scheme [4], based on the schemes of LSABE [5] and Chen et al. [6], we propose a secure multi-keyword search and access control scheme (SMKS-AC) that supports multi-keyword search of encrypted EHRs and verification of decrypted EHRs in WBAN. It enables data users to search for EHRs more accurately and verify the correctness of decrypted EHRs after decryption so as to ensure the reliability and security of EHRs sharing between data owners and data users. In addition, CP-ABE is employed to achieve fine-grained access control over EHRs. If a user's private key is compromised, the system can also track and revoke the user's real identity. Our SMKS-AC scheme is suitable to the limited resources of mobile devices since only a few exponentiation operations are performed on mobile devices, while bilinear operations are transferred to the public cloud.

The rest of this paper is organized as follows. Section 2 reviews related works. Section 3 provides the preliminaries for the proposed scheme. Section 4 presents system model and security requirements. Section 5 presents a concrete SMKS-AC construction. Section 6 shows the security analysis, as well as the function and performance comparison with other schemes. Finally, Section 7 concludes this paper.

2. Related Works

To achieve fine-grained access control of outsourcing data, ABE provides a good method of data encryption and sharing. ABE is an extension of the identity-based encryption algorithm, which was first proposed by Amit and Waters [7]. It can be divided into two types, namely, key policy ABE (KP-ABE) and ciphertext-policy ABE (CP-ABE). The first KP-ABE, proposed by Goyal et al. [8], associates ciphertext with a set of attributes, while the private key is associated with an access policy that controls which ciphertext users can decrypt. The CP-ABE is first proposed by Bethencourt et al. [9], which can realize complex access control on encrypted data. The main idea is to embed the user's attribute set into the private key, and the ciphertext contains the access policy that determines which user can access the ciphertext. The key can be used to decrypt the ciphertext for the user who can access the data. The advantage of the CP-ABE scheme is that encrypted data can be kept secret even if the storage server is untrusted [10], such as the public cloud.

In order to save storage resources and computing costs at mobile devices, Go et al. [11] introduced an outsourcing scheme that reduces the storage space of mobile devices by outsourcing encrypted medical data to cloud servers. It is also the first outsourcing scheme that can resist the malicious cloud server attacks. Li et al. [12] proposed an attribute-based

encryption scheme for verifiable outsourcing encryption and decryption. Their scheme not only reduces the storage cost of mobile devices but also requires only a small amount of computing overhead to complete the decryption. However, Guo et al. [13] found that there was a security problem in Li et al.'s scheme [12] and put forward an improved scheme. Fan et al. [14] proposed a secure and effective outsourcing computing algorithm to solve the problem of a large amount of computing in existing schemes for data encrypting on mobile devices. This algorithm aims to outsource most of the computing involved in encryption and decryption to the cloud, thereby reducing the cost of computing on mobile devices. In order to reduce the overhead of data transmission and the storage space occupied by encrypted data, Zhang et al. [15] proposed an outsourced data access control scheme with constant size, which can keep the encrypted data and the user's private key constant size.

When some users sell private keys for profit, the outsourced data will face the risk of disclosure. Yu et al. [16] proposed a scheme that can effectively protect the outsourced data. If a malicious user shares the decryption key for profit, the user's identity is required to be publicly verified and the request cannot be denied. Therefore, it is necessary to trace the user's identity in ABE. Zhang et al. [17] proposed a scheme that can change user attributes and track the identity of traitors. The results show that the scheme provides feasibility and reliability for practical application. Sethi et al. [18] constructed a multiauthority CP-ABE scheme, which not only provides the function of tracking the identity of malicious users who leak decryption keys but also provides the ability to outsource decryption to reduce the computational burden on users and also supports access policy updates.

Although ABE scheme can preserve the security and privacy of outsourced encrypted medical data, it still faces another problem in retrieving the required medical data in a large number of ciphertexts. The encryption algorithm directly makes the outsourced EHRs unreadable. In order to solve the problem of encrypted data search, Song et al. [19] proposed a scheme of using keywords to search encrypted data stored on untrusted servers without revealing any text information. After that many schemes have been proposed to search for encrypted data. For example, Vaanchig et al. [20] proposed a temporary and fuzzy keyword search public key encryption scheme, which can resist keyword guessing attacks and limit data retrieval time, thus enhancing the security of keyword search. Zhou et al. [21] proposed a public key encryption scheme, in which CP-ABE technology is used for fine-grained access control and keyword search of encrypted data. Their scheme is indistinguishable against adaptive selection keyword attacks.

Besides, the ABE scheme for keyword search of outsourced encrypted data has also been specifically studied in [22–25]. However, these schemes only support single keyword search, which limits the flexibility and accuracy of data retrieval. Sun et al. [26] proposed a multi-keyword search scheme based on CP-ABE, which supports auditing on search results. This scheme reduces a large number of irrelevant search results from cloud servers by narrowing the

search scope. Moreover, Long et al. [27] proposed a lightweight multi-keyword search algorithm based on attribute encryption, which not only supports multi-keyword search but also reduces the computing cost of mobile devices.

3. Preliminaries

3.1. Linear Secret-Sharing Scheme (LSSS)

Definition 1 (see [28]). A secret-sharing scheme Π over a set of parties \mathcal{P} is called linear (over \mathbb{Z}_p) if

- (1) The shares for all parties form a vector over \mathbb{Z}_p .
- (2) There exists a matrix M with l rows and n columns called the share-generating matrix for Π . For all $i = 1, \dots, l$, the i -th row of M is labeled by a party $\rho(i)$, where ρ is a function from i to \mathcal{P} . Set the column vector $v = (s, r_2, \dots, r_n)$, where $s \in \mathbb{Z}_p$ is the secret to be shared and r_2, \dots, r_n are randomly chosen. Mv is the vector of l shares of the secret s according to Π . The share $(Mv)_i$ belongs to party $\rho(i)$.

It is shown in [28] that every LSSS defined above has the linear reconstruction property. Let Π be an LSSS for access structure \mathbb{A} , $S \in \mathbb{A}$ be any authorized set, and $I \subset \{1, 2, \dots, l\}$ be defined as $I = \{i: \rho(i) \in S\}$. Then, there exist constants $\{\omega_i \in \mathbb{Z}_p\}_{i \in I}$ such that if $\{\lambda_i\}_{i \in I}$ are valid shares of any secret s according to Π , then $\sum_{i \in I} \omega_i \lambda_i = s$. Moreover, these constants $\{\omega_i\}_{i \in I}$ can be found in time polynomial in the size of the share-generating matrix M .

3.2. Bilinear Groups. Let \mathbb{G}_1 and \mathbb{G}_2 be two multiplicative cyclic groups of prime order p and g be a generator of \mathbb{G}_1 . If $e: \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ satisfies the following properties:

- (1) Bilinearity: for any $u, v \in \mathbb{G}_1$ and $a, b \in \mathbb{Z}_p$, we have $e(u^a, v^b) = e(u, v)^{ab}$.
- (2) Nondegeneracy: there exist two elements $u, v \in \mathbb{G}_1$ such that $e(u, v) \neq 1$, where 1 is the identity element in \mathbb{G}_2 .
- (3) Computability: $e(u, v)$ can be efficiently computed, for any $u, v \in \mathbb{G}_1$ then e is a bilinear map.

3.3. DBDH Assumption. The security of the proposed SMKS-AC construction is based on the following decisional bilinear Diffie–Hellman (DBDH) assumption.

Assumption 1 (decisional bilinear Diffie–Hellman (DBDH) assumption). Let \mathbb{G}_1 be a multiplicative cyclic group of prime order p with generator g . Let $a, b, c \in \mathbb{Z}_p^*$ be randomly selected. If an adversary \mathcal{A} is given g, g^a, g^b, g^c , and $e(g, g)^{abc}$, it is difficult for \mathcal{A} to distinguish $e(g, g)^{abc}$ from a random element R in \mathbb{G}_2 . The adversary \mathcal{A} has advantage ϵ in solving the DBDH assumption if $\text{Adv}_{DBDH} = |\Pr[\mathcal{A}(g, g^a, g^b, g^c, e(g, g)^{abc}) = 0] - \Pr[\mathcal{A}(g, g^a, g^b, g^c, R) = 0]| \geq \epsilon$.

4. System Model and Security Requirements

In this section, we introduce the system model and security requirements of SMKS-AC.

4.1. System Model. As shown in Figure 1, the system of SMKS-AC mainly consists of the following four types of entities: data owner, medical staff who are regarded as data users, public cloud (PC), and key generation center (KGC).

- (1) *WBAN (Data Owner).* WBAN contains many sensors, which are embedded or worn on patients to collect medical data. The collected data are aggregated and transmitted to personal digital assistant (PDA) through wireless channel. Keywords are extracted from EHRs to describe health information, and an EHR can be described by multiple keywords. Then, these keywords form a keyword group, and the corresponding EHR is encrypted under a specific access policy. Finally, these encrypted EHRs are outsourced to the PC.
- (2) *Medical Staff (Data User).* Medical staff, as data users, have their own specific set of attributes. Data users are authorized to search encrypted EHRs according to their set of attributes. A data user can generate keyword trapdoor and send it to the PC to realize data retrieval. As long as the search keywords are included in the keyword group describing the corresponding EHR and the user's set of attributes satisfies the access policy, the searched encrypted EHR will be returned. Then, the user uses his/her private key to decrypt and verify the decrypted EHR to ensure the correctness of EHR.
- (3) *PC.* PC has almost unlimited storage space and computing power, which can be used to store encrypted EHRs and respond to users' data retrieval requests. In SMKS-AC, PC first verifies whether the keywords in the retrieval request are in the keyword group of the retrieved data. Then, the PC helps transform the retrieved ciphertext into a ciphertext that users can decrypt through lightweight computing.
- (4) *KGC.* KGC generates public parameters for the whole system and distributes private key to each data user. The user's attributes set is embedded in the private key to implement access control. When the user's private key is maliciously disclosed, KGC can trace the identity of the private key holder and add it to the revocation list.

4.2. Security Requirements. In WBAN, in order to ensure the availability, privacy, and security of EHRs, a secure data access control scheme supporting multi-keyword search needs to meet the following security requirements.

4.2.1. Confidentiality of EHRs. The EHRs should be encrypted before being outsourced to the PC for remote

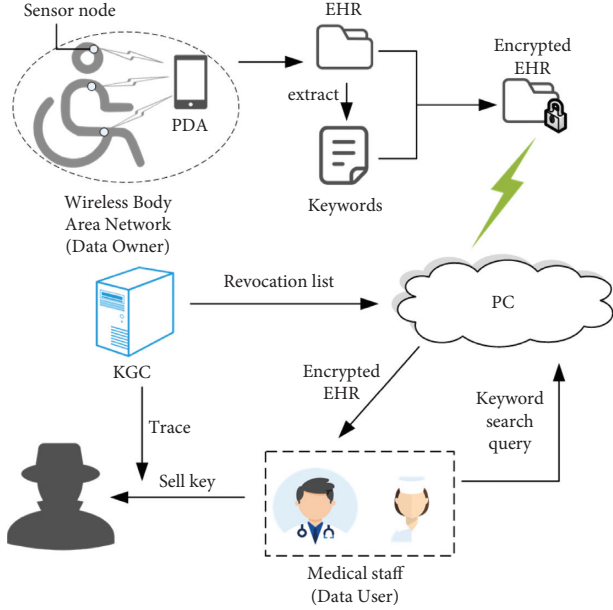


FIGURE 1: System model.

storage. Only when the user's set of attributes meets the access policy in the ciphertext, the data can be decrypted by the user.

4.2.2. Accuracy of Retrieved EHRs. When the users retrieve data, there will be a lot of redundant data in the single keyword search results. Thus, it is necessary to use multi-keyword search to improve the accuracy of retrieval results.

4.2.3. Verifiability of EHRs. Since the ciphertext is outsourced to PC, in order to prevent PC from tampering with the ciphertext, data users need to verify the accuracy of the retrieved data after decrypting with their own private key.

5. A SMKS-AC Construction

In this section, we present a SMKS-AC construction and depict the running procedure in Figure 2.

5.1. System Setup. Let $g \in \mathbb{G}_1$ be a generator of group \mathbb{G}_1 of prime order p . Let $e: \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ be a bilinear map, $H: \{0, 1\}^* \rightarrow \mathbb{Z}_p^*$ and $H_1: \{0, 1\}^* \rightarrow \kappa$ be two collision-resistant hash functions, and κ be the key space of the symmetric encryption algorithm (i.e., AES, DES, and so on). KGC selects random elements $\alpha, a, \lambda, \tau \in \mathbb{Z}_p^*$ and $k_1 \in \kappa$ and computes

$$\begin{aligned} f &= g^\tau, \\ Y &= e(g, g)^\alpha, \\ Y_0 &= e(g, f), \\ h &= g^\lambda. \end{aligned} \quad (1)$$

The public parameter and the master secret key of the system are $PP = (g, p, e, h, f, Y, Y_0, g^\alpha, f^\alpha, H, H_1)$ and $MSK = (\alpha, a, \lambda, \tau, k_1)$, respectively.

5.2. Key Generation. Let id be the user's identity and $S = \{\xi_1, \xi_2, \dots, \xi_k\} \subseteq \mathbb{Z}_p^*$ be the data user's attributes set, where k represents the number of attributes in the attribute set. KGC selects random elements $r, \varrho, u', u'' \in \mathbb{Z}_p^*$ and computes

$$\begin{aligned} \delta &= \text{Enc}_{k_1}(id), \\ D_1 &= g^{\alpha/(\lambda+\delta)}, \\ D_2 &= \delta, \\ D_{3,i} &= g^{r/(\xi_i+\tau)}, \\ D_4 &= g^{(\alpha-ar)/(\lambda+\delta)}, \\ D_5 &= \varrho, \\ \Psi_1 &= D_1^{u'}, \\ \Psi_2 &= (D_4^\varrho)^{u''}, \\ \Psi_{3,i} &= (D_{3,i}^\varrho)^{u''}. \end{aligned} \quad (2)$$

The public key $PK_{id,S}$ and secret key $SK_{id,S}$ of the data user are $PK_{id,S} = (\Psi_1, \Psi_2, \{\Psi_{3,i}\}_{i \in [k]})$ and $SK_{id,S} = (D_1, D_2, \{D_{3,i}\}_{i \in [k]}, D_4, D_5, u', u'')$, respectively.

5.3. Encryption. Data owner chooses keywords for an EHR m to form a keyword set $KW = \{kw_1, kw_2, \dots, kw_{l_1}\}$ and constructs the l_1 degree polynomial:

$$\varphi(x) = \eta_{l_1} x^{l_1} + \eta_{l_1-1} x^{l_1-1} + \dots + \eta_0 \text{ mod } p, \quad (3)$$

where $H(kw_1), H(kw_2), \dots, H(kw_{l_1})$ are the l_1 roots of $\varphi(x) = 1$. Let M be a $l \times n$ matrix and ρ be a function that associates M_i to attributes, where l and n represent the rows and columns of M , respectively, and M_i is the i -th row of M , $i = 1, 2, \dots, l$. Data owner randomly selects $Y \in \mathbb{G}_2^*$ and computes

$$\begin{aligned} \zeta &= H_1(Y), \\ C_m &= \text{Enc}_\zeta(m), \\ s &= H(m, Y). \end{aligned} \quad (4)$$

Then, the data owner chooses a random vector $\vec{v} = (s, y_2, y_3, \dots, y_n)^T \in \mathbb{Z}_p^n$, where $y_2, y_3, \dots, y_n \in \mathbb{Z}_p^*$ are random elements. For $i \in [l]$, the data owner computes

$$s_i = M_i \cdot \vec{v} \text{ mod } p, \quad (5)$$

selects random elements $s', s'' \in \mathbb{Z}_p^*$, and calculates

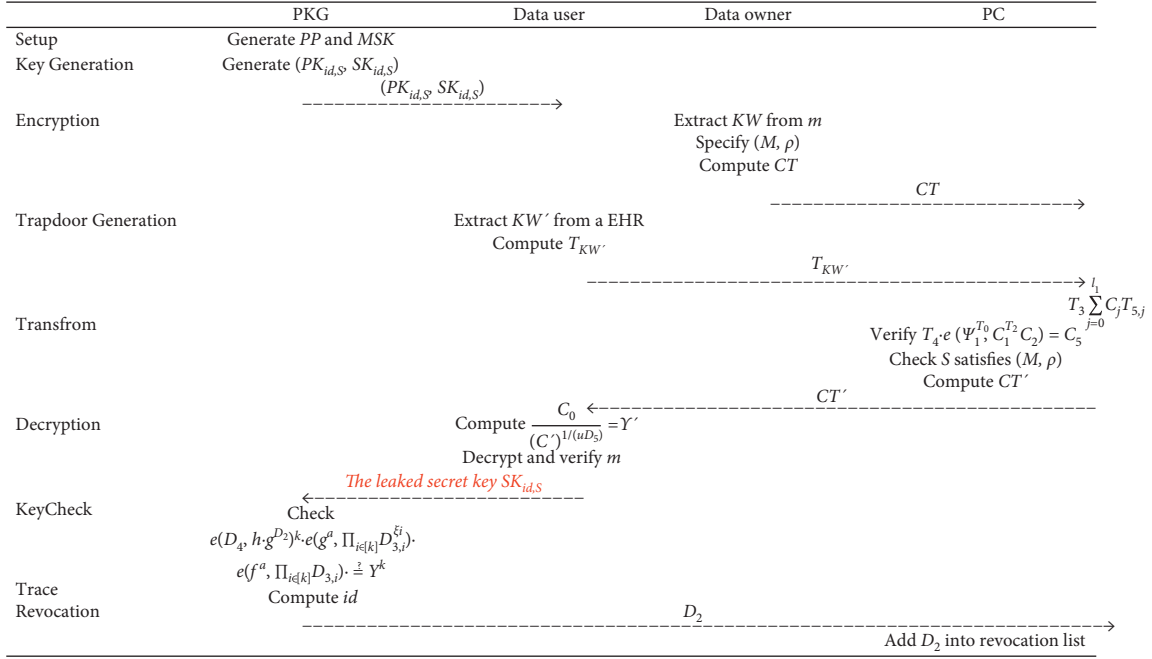


FIGURE 2: The flow of SMKS-AC construction.

$$\begin{aligned}
C_0 &= Y \cdot Y^s, & T_0 &= u \cdot (u')^{-1}, \\
C_1 &= g^s, & T_1 &= u \cdot (u'')^{-1}, \\
C_2 &= h^s, & T_2 &= D_2, \\
C_{3,i} &= \rho(i) \frac{s_i}{s}, & T_3 &= u \cdot u_0 \cdot l_2^{-1}, \\
C_{3,i}' &= \frac{s_i}{s'}, & T_4 &= Y_0^u, \\
C_4 &= f^{as''}, & T_{5,j} &= u_0^{-1} \sum_{i=1}^{l_2} H(kw_i)^j, \quad j \in \{0, 1, \dots, l_1\}. \\
C_5 &= Y_0^s Y^{ss'}, & & \\
C_6 &= g^{as'}, & & \\
C_j &= \frac{\eta_j}{s'}, \quad j \in \{0, 1, \dots, l_1\}. & &
\end{aligned} \tag{6}$$

Then, the ciphertext $CT = (C_0, C_1, C_2, \{C_{3,i}, C_{3,i}'\}_{i \in \{1, \dots, l_1\}}, C_4, C_5, C_6, \{C_j\}_{j \in \{0, 1, \dots, l_1\}}, C_m)$ and the access policy (M, ρ) are outsourced to PC.

5.4. Trapdoor Generation. If a data user wants to search for EHRs containing the keyword set $KW' = \{kw_1, kw_2, \dots, kw_{l_2}\}$, the user needs to construct keywords trapdoor $T_{KW'}$. The data user chooses random elements $u, u_0 \in \mathbb{Z}_p^*$ and computes

The keyword trapdoor $T_{KW'} = (T_0, T_1, T_2, T_3, T_4, \{T_{5,j}\}_{j \in \{0, 1, \dots, l_1\}})$ is sent to PC.

5.5. Transform. After receiving keywords trapdoor $T_{KW'}$ from a data user, the PC first verifies whether the following equation is true:

$$T_4 \cdot e(\Psi_1^{T_0}, C_1^{T_2} C_2) = C_5 \sum_{j=0}^{l_1} C_j T_{5,j}. \tag{8}$$

If so, the PC outputs 1 means $KW' \subset KW$. Otherwise, the PC outputs 0. And then, PC checks whether the attribute set S associated with $T_{KW'}$ meets the access policy (M, ρ) associated with ciphertext CT .

- (1) If the PC outputs 0 or the attribute set S associated with $T_{KW'}$ does not satisfy the access policy (M, ρ) , the Transform algorithm aborts.
- (2) If the PC outputs 1 and S associated with $T_{KW'}$ satisfies (M, ρ) , the Transform algorithm continues as follows. Let $I \subset [l]$ be $I = \{i: \rho(i) \in S\}$, and there exists a set of constants $\{\omega_i \in \mathbb{Z}_p\}_{i \in I}$ such that

$$\sum_{i \in I} \omega_i M_i = (1, 0, \dots, 0). \quad (9)$$

The PC computes

$$\begin{aligned} I' &= e \left(C_6, \prod_{i \in I} \Psi_{3,i}^{C_{3,i} \omega_i} \right) \cdot e \left(C_4, \prod_{i \in I} \Psi_{3,i}^{C_{3,i}' \omega_i} \right), \\ \Gamma &= (I')^{T_1}, \\ \Gamma' &= e(\Psi_2, C_1^{T_2} C_2)^{T_1}. \end{aligned} \quad (10)$$

Therefore, $C' = \Gamma \cdot \Gamma' = e(g, g)^{\alpha \text{qsu}}$.

Finally, the transformed ciphertext $CT' = (C_0, C', C_m)$ is sent back to the data user.

5.6. Decryption and Verification. After receiving the transformed ciphertext $CT' = (CT_1, CT_2, C_m)$, the data user calculates

$$\begin{aligned} Y' &= \frac{CT_1}{(CT_2)^{1/(uD_5)}}, \\ \zeta t &= H_1(Y'), \end{aligned} \quad (11)$$

$$\begin{aligned} mt &= \text{Dec}_{\zeta'}(C_m), \\ s_1 &= H(m', Y'). \end{aligned} \quad (12)$$

If

$$CT_1 = Y' \cdot Y^{s_1}, \quad (13)$$

$$CT_2 = Y^{s_1 u D_5}, \quad (14)$$

both hold, then m' is outputted.

5.7. User Tracing. If the private key of some authorized user is leaked, KGC is able to perform the following two algorithms to track its real identity.

5.7.1. Keycheck. Suppose $S = \{\xi_1, \xi_2, \dots, \xi_k\}$. KGC checks whether the private key $SK_{\text{id}, S}$ is in the form of $(D_1, D_2, \{D_{3,i}\}_{i \in [k]}, D_4, D_5, u', u'')$, where $D_1, D_{3,i}, D_4 \in \mathbb{G}_1$ and $D_2, D_5, u', u'' \in \mathbb{Z}_p^*$. Then, the KGC verifies the following equation:

$$e(D_4, h \cdot g^{D_2})^k \cdot e \left(g^a, \prod_{i \in [k]} D_{3,i}^{\xi_i} \right) \cdot e \left(f^a, \prod_{i \in [k]} D_{3,i} \right) = Y^k. \quad (15)$$

If it is satisfied, the algorithm outputs 1 or 0 otherwise.

5.7.2. Trace. If the KeyCheck algorithm outputs 0, it means that the private key $SK_{\text{id}, S}$ does not need to be traced; in this case, the Trace algorithm outputs \perp . Otherwise, KGC can use the master key MSK to recover the identity of the private key holder by calculating $\text{id} = \text{Dec}_{k_1}(D_2)$.

5.8. User Revocation. After the Trace algorithm is completed, the access right of the holder who leaked the private key $SK_{\text{id}, S}$ needs to be revoked. Therefore, KGC stores the component $D_2 = \delta$ of the key containing the user's identity information on the PC. When a user sends a data retrieval request and submits keywords trapdoor $T_{KW'}$, the PC checks whether the component $D_2 = \delta$ in $T_{KW'}$ is in the revocation list. If so, the user's data retrieval request is rejected.

6. Scheme Analysis

6.1. Correctness and Security Analysis. This section analyzes and proves the correctness and security of SMKS-AC construction.

Theorem 1. *The proposed SMKS-AC construction is correct.*

Proof. In order to prove the correctness of the SMKS-AC construction, we only need to show that equations (8), (11), and (15) hold.

First, PC can determine whether the keyword set KW' searched by the data user satisfies $KW' \subset KW$ by verifying equation (8), where KW is keyword set in CT . Since

$$\begin{aligned} T_4 \cdot e(\Psi_1^{T_0}, C_1^{T_2} C_2) &= Y_0^u \cdot e \left(D_1^{u' \cdot u \cdot (t)^{-1}}, g^{s\delta} g^{\lambda s} \right) \\ &= e(g, f)^u \cdot e \left(g^{au/(\lambda+\delta)}, g^{(\lambda+\delta)s} \right) \\ &= e(g, f)^u \cdot e(g, g)^{\alpha u s}, \\ C_5^{T_3} \sum_{j=0}^{l_1} C_j^{T_{5,j}} &= \left(Y_0^{s'} Y^{ss'} \right)^{u \cdot u_0 \cdot (l_2)^{-1}} \sum_{j=0}^{l_1} \left(\eta_j / s' \cdot (u_0)^{-1} \sum_{\ell=1}^{l_2} H(kw_i)^\ell \right)^j \\ &= (e(g, f) \cdot e(g, g)^{\alpha s})^{u \cdot (l_2)^{-1}} \sum_{j=0}^{l_1} \left(\eta_j \cdot \sum_{\ell=1}^{l_2} H(kw_i)^\ell \right)^j \\ &= (e(g, f) \cdot e(g, g)^{\alpha s})^u, \end{aligned} \quad (16)$$

we have $KW' \subset KW$.

Second, after receiving the PC's transformed ciphertext, the data user can recover the random element Y by calculating equation (11), so as to decrypt and verify the message. Since $C' = \Gamma \cdot \Gamma' = (I')^{T_1} \cdot \Gamma'$, we have

$$\frac{C_0}{(C')^{1/(uD_s)}} = \frac{Y \cdot e(g, g)^{\alpha s}}{(e(g, g)^{\alpha q s u})^{1/uq}} = Y, \quad (17)$$

where

$$\begin{aligned} I' &= e\left(C_6, \prod_{i \in I} \Psi_{3,i}^{C_{3,i} \omega_i}\right) \cdot e\left(C_4, \prod_{i \in I} \Psi_{3,i}^{C'_{3,i} \omega_i}\right) \\ &= e\left(g^a, g^{rqu'' \sum_{i \in I} \rho(i) s_i \omega_i / (\xi_i + \tau)}\right) \cdot e\left(g^{\tau a}, g^{rqu'' \sum_{i \in I} s_i \omega_i / (\xi_i + \tau)}\right) \\ &= e(g, g)^{arqu'' \sum_{i \in I} \rho(i) s_i \omega_i / (\xi_i + \tau) + \sum_{i \in I} \tau s_i \omega_i / (\xi_i + \tau)} \\ &= e(g, g)^{arqu'' s}, \\ \Gamma' &= e(\Psi_2, C_1^{T_2} C_2)^{T_1} \\ &= e\left(g^{(\alpha - ar)/(\lambda + \delta) qu''}, g^{s(\lambda + \delta)}\right)^{u(u'')^{-1}} \\ &= e(g, g)^{(\alpha - ar) q s u}, \\ C' &= \left(e(g, g)^{arqu'' s}\right)^{T_1} e(g, g)^{(\alpha - ar) q s u} \\ &= e(g, g)^{arqu'' s \cdot u(u'')^{-1}} e(g, g)^{(\alpha - ar) q s u} \\ &= e(g, g)^{\alpha q s u}. \end{aligned} \quad (18)$$

Third, if the data user sells the private key for profit, KGC can verify the authenticity of the sold key by verifying equation (15) to determine whether the identity of the user holding the private key is worth tracking due to

$$\begin{aligned} &e(D_4, h \cdot g^{D_2})^k \cdot e\left(g^a, \prod_{i \in [k]} D_{3,i}^{\xi_i}\right) \cdot e\left(f^a, \prod_{i \in [k]} D_{3,i}\right) \\ &= e\left(g^{(\alpha - ar)/(\lambda + \delta)}, g^\lambda g^\delta\right)^k \cdot e\left(g^a, g^{i \in [k]} \sum_{i \in [k]} r \xi_i / (\xi_i + \tau)\right) \cdot e\left(g^{\tau a}, g^{i \in [k]} \sum_{i \in [k]} r / (\xi_i + \tau)\right) \\ &= e(g, g)^{(\alpha - ar)k} \cdot e(g, g)^{ar \cdot \sum_{i \in [k]} \xi_i / (\xi_i + \tau)} \cdot e(g, g)^{ar \cdot \sum_{i \in [k]} r / (\xi_i + \tau)} \\ &= e(g, g)^{\alpha k} \end{aligned} \quad (19)$$

Therefore, the proposed SMKS-AC construction is correct. \square

Theorem 2. *The proposed SMKS-AC construction can ensure the EHRs confidentiality. Under the DBDH assumption, the SMKS-AC construction is indistinguishable against chosen keyword and chosen plaintext attack (IND-CKCPA).*

Proof. The form of ciphertext CT in the SMKS-AC construction is similar to that in [4]. Compared with the SMKS-AC construction, the data owners of [4] only extract one keyword in processing an EHR. In our construction, in order to improve the accuracy of the data retrieval by data users, the data owners of SMKS-AC construction are able to extract multiple keywords from an EHR. The specific proof of Theorem 2 is similar to Theorem 1 in [4]. Therefore, the SMKS-AC construction is IND-CKCPA secure. \square

Theorem 3. *The proposed SMKS-AC construction can resist collusion attacks against multiple users.*

Proof. In a multi-user system, collusion attack is an important attack type. Authorized users can collude with each other to generate a new key and gain extra privileges. Whereas in our solution, for each user, KGC selects a set of random numbers based on the user's attributes to generate the user's private key. Users who intend to collude with each other cannot combine their private keys to generate a new valid key. Since each user's private key is generated by different random numbers, they are not compatible with each other. Therefore, the proposed SMKS-AC construction is secure against collusion attacks. \square

6.2. Comparison. This section compares SMKS-AC construction with existing schemes in terms of function, storage, and computation overheads. The comparison is shown in Tables 1–3, respectively.

As shown in Table 1, in addition to the proposed SMKS-AC construction, both schemes [5, 29] provide multi-keyword search function. Data users are allowed to verify the message after decryption in [4] and SMKS-AC construction. Since the user's private key may be used maliciously, both [4] and the SMKS-AC construction provide user trace and user revocation functions, which are not considered in [5, 22, 29].

The storage and computing resources of mobile devices in WBAN are limited. In practical applications, storage and computation costs on mobile devices need to be considered. Let $|\text{PP}|$, $|\text{SK}|$, $|\text{CT}|$, and $|T_{KW'}|$ represent the sizes of the public parameter, private key of data user, ciphertext, and the keywords trapdoor, respectively. $|\mathbb{G}_1|$, $|\mathbb{G}_2|$, and $|\mathbb{Z}_p|$ denote the length of an element in groups \mathbb{G}_1 , \mathbb{G}_2 , and \mathbb{Z}_p , respectively. Let $|S|$ be the size of attribute set S , l be the number of rows in M , $|U|$ be the size of the universe attribute set U , and l_1 be the size of the keyword set KW . Besides, E_1 and E_2 represent exponentiation operations on groups \mathbb{G}_1 and \mathbb{G}_2 , respectively. P is the symbol of bilinear pairing operation.

Table 2 shows the storage cost comparison. It can be seen that only the size of public parameter in Wang et al.'s scheme [29] is related to U , which is the number of all the attributes in the whole system. As the number of system attributes increases, the size of public parameter also increases, which do not meet the actual needs. Although the length of public parameters of schemes [5, 22] is smaller than that of SMKS-AC construction, the sizes of user's private key, ciphertext, and the keywords trapdoor are, respectively, larger than

TABLE 1: Function comparison.

	Access control	Multi-keyword search	Verify decryption	User trace	User revocation
Yang et al.'s scheme [4]	✓	×	✓	✓	✓
Zhang et al.'s scheme [5]	✓	✓	×	×	×
Wang et al.'s scheme [22]	✓	×	×	×	×
Wang et al.'s scheme [29]	✓	✓	×	×	×
SMKS-AC scheme	✓	✓	✓	✓	✓

“✓” represents supported; “×” represents not supported.

TABLE 2: Storage cost comparison.

	PP	SK	CT	$ T_{KW} $
Yang et al.'s scheme [4]	$5 \mathbb{G}_1 + 2 \mathbb{G}_2 $	$(S + 1) \mathbb{G}_1 + 5 \mathbb{Z}_p $	$4 \mathbb{G}_1 + 2 \mathbb{G}_2 + 2l \mathbb{Z}_p $	$7 \mathbb{Z}_p $
Zhang et al.'s scheme [5]	$4 \mathbb{G}_1 + \mathbb{G}_2 $	$(S + 3) \mathbb{G}_1 + \mathbb{Z}_p $	$(l + 4) \mathbb{G}_1 + 2 \mathbb{G}_2 + (l_1 + 4) \mathbb{Z}_p $	$ \mathbb{G}_1 + \mathbb{G}_2 + (l_1 + 2) \mathbb{Z}_p $
Wang et al.'s scheme [22]	$5 \mathbb{G}_1 $	$(2 S) \mathbb{G}_1 + \mathbb{Z}_p $	$(l + 2) \mathbb{G}_1 + 3 \mathbb{G}_2 + \mathbb{Z}_p $	$2 \mathbb{G}_1 + \mathbb{Z}_p $
Wang et al.'s scheme [29]	$(4 + U) \mathbb{G}_1 + 3 \mathbb{G}_2 $	$3 \mathbb{G}_1 + \mathbb{Z}_p $	$(2l + l_1 + 4) \mathbb{G}_1 + \mathbb{G}_2 $	$(3 + l_1) \mathbb{G}_1 $
SMKS-AC scheme	$5 \mathbb{G}_1 + 2 \mathbb{G}_2 $	$(S + 2) \mathbb{G}_1 + 4 \mathbb{Z}_p $	$4 \mathbb{G}_1 + 2 \mathbb{G}_2 + (2l + l_1) \mathbb{Z}_p $	$ \mathbb{G}_2 + (l_1 + 4) \mathbb{Z}_p $

TABLE 3: Computation cost comparison.

	Key generation	Encryption	Trapdoor	Decryption
Yang et al.'s scheme [4]	$(S + 1)E_1$	$4E_1 + 3E_2$	0	E_2
Zhang et al.'s scheme [5]	$(S + 4)E_1$	$(2l + 4)E_1 + 3E_2 + 3P$	$E_1 + E_2 + P$	E_2
Wang et al.'s scheme [22]	$3E_1$	$5E_1 + 3E_2$	$2E_1$	$E_1 + P$
Wang et al.'s scheme [29]	$(S + 6)E_1$	$(4 + l + l_1)E_1 + E_2 + P$	$(3 + l_1)E_1$	$E_1 + E_2 + P$
SMKS-AC scheme	$(S + 2)E_1$	$4E_1 + 3E_2$	E_2	E_2

those of SMKS-AC construction, which will increase the storage burden and data transmission time of mobile devices. Although private key size of scheme [29] is smaller than that of SMKS-AC construction, the sizes of ciphertext and the keywords trapdoor in SMKS-AC construction are, respectively, smaller than those of [29]. Note that the length of the elements in group \mathbb{Z}_p is much smaller than that in group \mathbb{G} . In addition, the sizes of private key, ciphertext and keywords trapdoor in SMKS-AC scheme are, respectively, larger than those in scheme [4]. The main reason is that the SMKS-AC construction realizes multi-keyword search, which can ensure the accuracy of data retrieval, while scheme [4] can only support single keyword search.

Table 3 shows the computation cost comparison, where only the cost of exponentiation and bilinear pairing operations are considered. In the Key Generation algorithm, the KGC can use $|S| + 2$ exponentiation operations on group \mathbb{G}_1 to get the user's private key in SMKS-AC construction. Except for [4], other schemes require more computation than ours. The Encryption algorithm is executed on mobile devices with limited resources. In order to save storage space on mobile devices, the EHRs should be encrypted immediately and then transferred to the PC, which require high encryption efficiency. In our scheme and the scheme [4], only four exponentiation operations on group \mathbb{G}_1 and three exponentiation operations on group \mathbb{G}_2 are required to generate a ciphertext. However, the other schemes require other redundant operations overhead. Since the Trapdoor

Generation and Decryption algorithms are executed on the user's mobile device, a small amount of computations is required to get the keywords trapdoor and decryption data to meet the system needs. In addition, compared with the scheme [4] and our construction, the other schemes need to carry out bilinear pairing operations and exponentiation operations.

6.3. Experimental Analysis. In this section, we implement the code based on the Pairing-Based Cryptography Library (PBC-0.5.14, <https://crypto.stanford.edu/pbc/>). The experimental simulation is run on a virtual machine with 4-core 8 GB memory, 64 bit Linux Ubuntu 18.04.5 operating system, and Intel (R) Core (TM) i5-8265U CPU @ 1.60 GHz 1.80 GHz. The element of cyclic group is 512 bits, and the length of p is 160 bits.

Figure 3 compares the running time of each phase in SMKS-AC construction with that in other schemes. Since other schemes had no trace and revocation phases, we only compared the time of Key Generation algorithm, Encryption algorithm, Trapdoor Generation algorithm, Transform algorithm, and Decryption algorithm. Due to the limited resources of mobile devices, it is particularly necessary to consider the complexity of algorithms executed by data owners and users.

Figure 3(a) shows the time required for the Key Generation algorithm, which is executed by the KGC. It can be

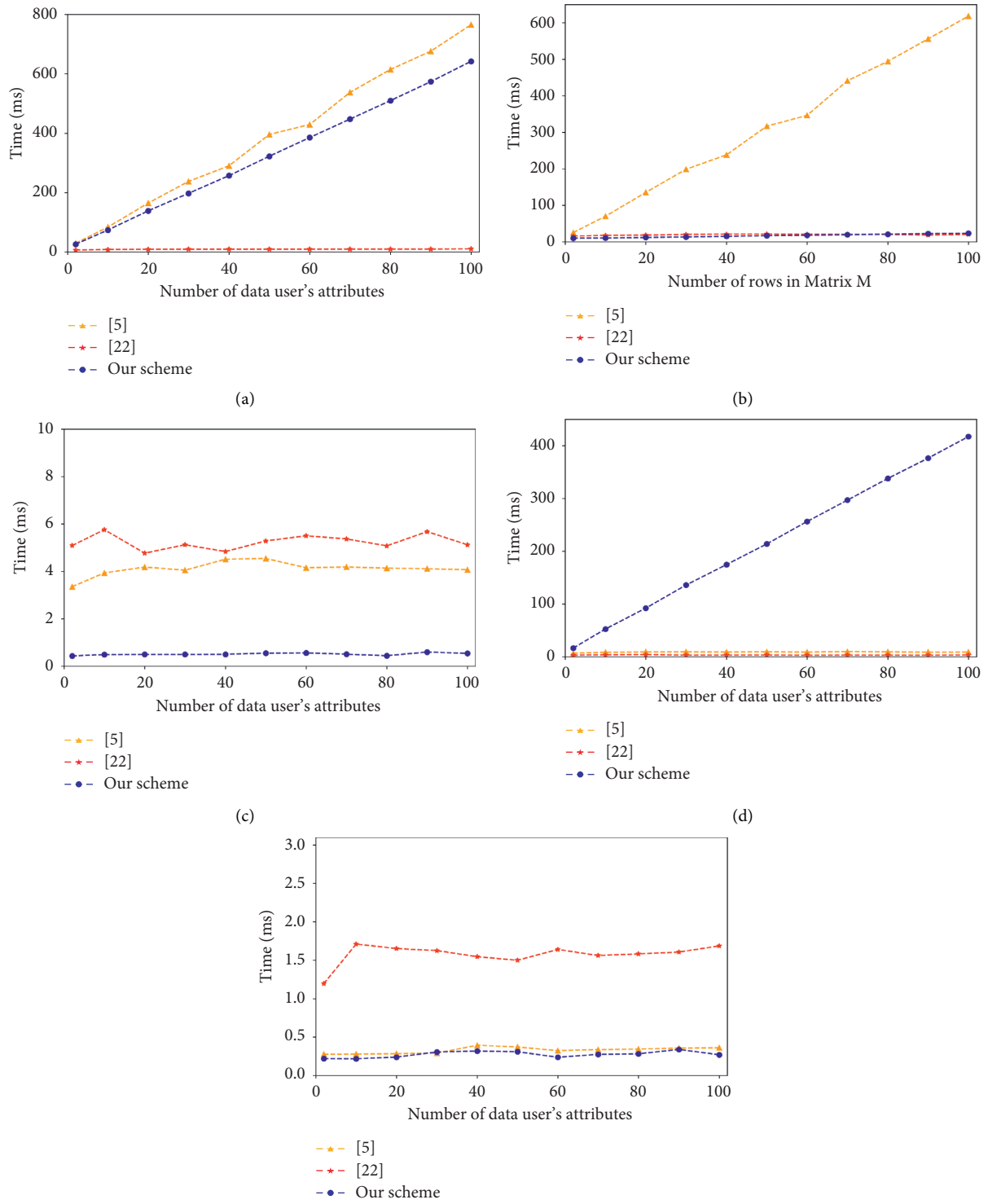


FIGURE 3: Computation overhead. (a) Key Generation time. (b) Encryption time. (c) Trapdoor Generation time. (d) Transform time. (e) Decryption time.

seen that the scheme in [22] takes the least time to generate the key. With the increase in the number of attributes, the time required remains unchanged. However, our SMKS-AC construction generates different keys for different attributes in the attribute set of data users; thus, the time user key generation will increase as the number of attributes in the attribute set increases.

Figure 3(b) indicates the time required for the data owner to execute the Encryption algorithm. The computational cost of SMKS-AC construction and the scheme in [22] roughly keeps constant, while the computational cost of [5] is increasing dynamically. The main reason is that the generation time of ciphertext in scheme [5] is related to the number of rows in matrix M . As the number of rows increases, the time to run the Encryption algorithm also increases.

It can be seen from Figure 3(c) that the time cost of the Trapdoor Generation algorithm has nothing to do with the number of attributes, and the generation time of trapdoor by SMKS-AC construction is shorter than that of the existing schemes. Note that the SMKS-AC construction and scheme [5] support multi-keyword search to improve the search accuracy, while scheme [22] only supports single keyword search.

As indicated in Figure 3(d), the SMKS-AC construction takes more time than the existing schemes. It is worth noting that the Transform algorithm is implemented by the PC, which has almost unlimited computing power and resources.

Figure 3(e) shows the decryption time of encrypted data by data users. The SMKS-AC construction takes almost the same decryption time as the scheme [5], and scheme [22] takes relatively much time. Also, the SMKS-AC construction satisfies the verification on the message after decryption.

7. Conclusion

In WBAN, in order to achieve secure sharing of outsourced EHRs with a large number of users, we proposed a SMKS-AC construction supporting secure multi-keyword search and access control. SMKS-AC provides fine-grained access control and verifiability of decrypted EHRs, multi-keyword search over encrypted EHRs, user's identity tracking, and revocation. Security analysis showed that the SMKS-AC construction can resist chosen keyword and chosen plaintext attacks and collusion attacks. Theoretical analysis and experiments demonstrate that our SMKS-AC construction is more effective and takes lower computational cost than existing related solutions.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This article was supported in part by the National Key R&D Program of China under project 2020YFB1006003, the

National Natural Science Foundation of China under projects 61772150, 61862012, and 61962012, the Guangdong Key R&D Program under project 2020B0101090002, the Guangxi Natural Science Foundation under grants 2018GXNSFDA281054, 2019GXNSFFA245015, and 2019GXNSFGA245004, the Peng Cheng Laboratory Project of Guangdong Province under grant PCL2018KP004, the Innovation Projects of GUET Graduate Education under grants 2021YCXS116 and 2021YCXS115, the Guangxi Young Teachers' Basic Ability Improvement Program under grant 2021KY0214, and the Open Program of Guangxi Key Laboratory of Cryptography and Information Security under grant GCIS201930.

References

- [1] A. Gazi Imtiyaz, S. Jimmye, and J. G. Kaiser, *Security and Privacy of E-Health Data*, pp. 199–214, Springer, Singapore, 2021.
- [2] C. Yang, X. Tao, F. Zhao, and Y. Wang, "Secure data transfer and deletion from counting bloom filter in cloud computing," *Chinese Journal of Electronics*, vol. 29, no. 2, pp. 79–86, 2020.
- [3] X. Zhou, J. Liu, Z. Zhang, and Q. Wu, "Secure outsourced medical data against unexpected leakage with flexible access control in a cloud storage system," *Security and Communication Networks*, vol. 2020, no. 7, pp. 1–20, 2020.
- [4] Y. Yang, X. Liu, R. H. Deng, and Y. Li, "Lightweight sharable and traceable secure mobile health system," *IEEE Transactions on Dependable and Secure Computing*, vol. 17, no. 1, pp. 78–91, 2020.
- [5] K. Zhang, J. Long, X. Wang, H.-N. Dai, K. Liang, and M. Imran, "Lightweight searchable encryption protocol for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4248–4259, 2021.
- [6] Y. Chen, Z. L. Jiang, S. M. Yiu, J. K. Au, and X. Wang, "Fully secure ciphertext-policy attribute based encryption with security mediator," *Information and Communications Security*, vol. 8958, pp. 274–289, 2015.
- [7] S. Amit and B. Waters, "Fuzzy identity-based encryption," *Lecture Notes in Computer Science*, vol. 3494, no. 1, p. 457, 2005.
- [8] V. Goyal, O. Pandey, S. Amit, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, pp. 89–98, CCS 2006, Alexandria, VA, USA, November 2006.
- [9] J. Bethencourt, S. Amit, and B. Waters, "Ciphertext-policy attribute-based encryption," *IEEE Symposium on Security and Privacy*, vol. 321, 2007.
- [10] Y. Wang, L. Wei, X. Tong, X. Zhao, and M. Li, "CP-ABE based access control for cloud storage," *Advances in Intelligent Systems and Computing*, vol. 455, pp. 463–472, 2017.
- [11] O. Go, R. Safavi-Naini, and L. F. Zhang, "Outsourcing scheme of ABE encryption secure against malicious adversary," *Computers & Security*, vol. 86, pp. 437–452, 2019.
- [12] Z. Li, W. Li, Z. Jin, H. Zhang, and Q. Wen, "An efficient ABE scheme with verifiable outsourced encryption and decryption," *IEEE Access*, vol. 7, pp. 29023–29037, 2019.
- [13] Z. Guo, B. Liu, K. Zhao, and C. Feng, "Comment on "an efficient abe scheme with verifiable outsourced encryption and decryption,"" *IEEE Access*, vol. 8, pp. 202483–202486, 2020.
- [14] K. Fan, T. Liu, K. Zhang, H. Li, and Y. Yang, "A secure and efficient outsourced computation on data sharing scheme for

- privacy computing,” *Journal of Parallel and Distributed Computing*, vol. 135, pp. 169–176, 2020.
- [15] Z. Zhang, W. Zhang, and Z. Qin, “Fully constant-size CP-ABE with privacy-preserving outsourced decryption for lightweight devices in cloud-assisted IoT,” *Security and Communication Networks*, vol. 2021, no. 6676862, pp. 1–16, 2021.
- [16] G. Yu, X. Ma, Z. Cao, G. Zeng, and W. Han, “Accountable CP-ABE with public verifiability: how to effectively protect the outsourced data in cloud,” *International Journal of Foundations of Computer Science*, vol. 28, no. 6, pp. 705–723, 2017.
- [17] R. Zhang, L. Hui, Sm Yiu, X. Yu, L. Zoe, and Jiang, “A traceable outsourcing CP-ABE scheme with attribute revocation,” in *Proceedings of the 2017 16th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/11th IEEE International Conference on Big Data Science and Engineering/14th IEEE International Conference on Embedded Software and Systems*, pp. 363–370, IEEE, Sydney, Australia, August 2017.
- [18] K. Sethi, A. Pradhan, and P. Bera, “Practical traceable multi-authority CP-ABE with outsourcing decryption and access policy updation,” *Journal of Information Security and Applications*, vol. 51, Article ID 102435, 2020.
- [19] D. X. Song, D. Wagner, and P. Adrian, “Practical techniques for searches on encrypted data,” in *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 44–55, Berkeley, California, USA, May 2002.
- [20] N. Vaanchig, Z. G. Qin, and Z. Qin, “Public key encryption with temporary and fuzzy keyword search,” *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 3914–3935, 2019.
- [21] Y. Zhou, S. Zheng, and L. Wang, “Privacy-preserving and efficient public key encryption with keyword search based on CP-ABE in cloud,” *Cryptography*, vol. 4, no. 28, p. 28, 2020.
- [22] H. Wang, J. Ning, X. Huang, G. Wei, G. Sen Poh, and X. Liu, “Secure fine-grained encrypted keyword search for e-healthcare cloud,” *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 3, pp. 1307–1319, 2021.
- [23] Y. Miao, R. Deng, X. Liu, K.-K. Raymond Choo, H. Wu, and H. Li, “Multi-authority attribute-based keyword search over encrypted cloud data,” *IEEE Transactions on Dependable and Secure Computing*, vol. 18, pp. 1667–1680, 2019.
- [24] H. Yin, Z. Qin, J. Zhang, H. Deng, F. Li, and K. Li, “A fine-grained authorized keyword secure search scheme with efficient search permission update in cloud computing,” *Journal of Parallel and Distributed Computing*, vol. 135, pp. 56–69, 2020.
- [25] J. Cui, H. Zhou, Y. Xu, and H. Zhong, “OOABKS: online/offline attribute-based encryption for keyword search in mobile cloud,” *Information Sciences*, vol. 489, pp. 63–77, 2019.
- [26] J. Sun, L. Ren, S. Wang, and X. Yao, “Multi-keyword searchable and data verifiable attribute-based encryption scheme for cloud storage,” *IEEE Access*, vol. 7, pp. 66655–66667, 2019.
- [27] J. Long, K. Zhang, X. Wang, and H.-N. Dai, “Lightweight distributed attribute based keyword search system for internet of things,” *Security, Privacy, and Anonymity in Computation, Communication, and Storage*, vol. 11637, pp. 253–264, 2019.
- [28] B. Amos, *Secure Schemes for Secret Sharing and Key Distribution*, PhD thesis, Israel Institute of Technology Sivan, Haifa, 1996.
- [29] S. Wang, S. Jia, and Y. Zhang, “Verifiable and multi-keyword searchable attribute-based encryption scheme for cloud storage,” *IEEE Access*, vol. 7, pp. 50136–50147, 2019.

Research Article

Privacy-Preserving Redactable Blockchain for Internet of Things

Yanli Ren , Xianji Cai, and Mingqi Hu

School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Yanli Ren; renyanli@shu.edu.cn

Received 8 June 2021; Revised 4 August 2021; Accepted 25 August 2021; Published 20 September 2021

Academic Editor: Jinguang Han

Copyright © 2021 Yanli Ren 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 traditional blockchain system, data is public and cannot be redacted. With the development of blockchain technology, the problem that the data cannot be altered will be more serious once it is written on the chain. Recently, some redactable blockchain schemes have been proposed. However, most of the schemes are based on the public blockchain, and the users' identities and transaction data may be disclosed. To solve the problem of privacy disclosure, we propose a privacy-preserving transaction-level redactable blockchain. In the proposed scheme, symmetric encryption and ring signature are used to protect transaction data and the users' identities, respectively. In order to prove the legality of data redaction, the transaction sender can reveal the invalid users' identities and transaction data in an anonymous environment. To construct a transaction-level redactable blockchain, the users only need to replace a single transaction to complete the data redaction instead of replacing the entire block. The experimental results show that the proposed scheme saves 20% of the redaction time compared to the previous privacy-preserving blockchains, so the redaction efficiency is higher.

1. Introduction

In January 2009, blockchain was invented as the underlying technology of Bitcoin [1], which merged the important achievements in some fields such as modern cryptography and distributed network. After the emergence of Bitcoin, the blockchain network has supported massive transfer transactions steadily in a purely distributed manner [2]. Especially, the combinations of blockchain technology and the Internet of things (IoT) begin to appear in large numbers, which proves that the blockchain is a good solution to the basic needs of distributed networks [3, 4]. For example, the blockchain is used to ensure the safety and anticounterfeiting traceability of the food supply chain [5].

With the wide application of blockchain technology in E-government, military, and medical fields [6–8], the problem that the data cannot be altered will be more serious. The continuous growth of data in the blockchain leads to a dramatical increase in storage space and data verification cost, which will result in the continuous decline of the number of the full-nodes and intensify the trend of centralization in blockchain. Moreover, the computational overhead for the nodes to verify historical data will also

increase, which is unfriendly to some IoT devices with limited computing resource [9]. Therefore, redacting or deleting the useless data will be an important means to improve the performance and scalability of the blockchain.

Recently, some redactable blockchains have been proposed [10, 11]. However, the users' identities and transactions are public. Even if the IoT devices simply apply pseudonyms when acting in the blockchain, they can only provide limited identity privacy. The adversary can still trace the transparent block data to find out the relations between the identities and transactions and analyze the real identities of the users [12]. Moreover, in the process of data redaction, the personal information of the redactors may be revealed. Due to the wide applications of IoT devices, the data owners do not want to leak the identity privacy or business secrets to their competitors. Therefore, it is extremely important to propose a privacy-preserving redactable blockchain to simultaneously protect the privacy of the users and rectify the incorrect data.

1.1. Privacy-Preserving Blockchain. An important feature of blockchain is data transparency, which means that the

identity and transaction data in public blockchain are transparent. However, for some companies or users, data is extremely important to maintain their strong competitiveness, and sharing data may bring some security challenges. Therefore, more and more privacy-preserving mechanisms on blockchain have been proposed by researchers. At present, the most popular solutions to protect the identity privacy are ring signature [13], mixing services [14], and noninteractive zero-knowledge proof [15], and the most common methods to protect the data privacy include zero-knowledge proof, homomorphic encryption, and commitment schemes [16].

1.2. Redactable Blockchain. In [10], the concept of redactable blockchain is first raised. The main idea of their protocol was to keep the block hash link compatible with the current state when a redaction happens. They use a special hash function called chameleon hash [17] to replace the traditional collision-resistant hash function, where collisions of the hash can be computed with trapdoors. Thus, a collision can be calculated to keep the hash linked as usual. In this work, the authors improve the chameleon hash protocol and solve the key exposure problem in the previous chameleon hash and also provide the formal security analysis of their new protocol. However, in this scheme, a trusted-party is needed to control the trapdoor of chameleon hash, which violates the idea of decentralization of the blockchain, and the redaction operation does not represent the opinions of the whole system. Moreover, this scheme can only realize data redaction at the block-level. Once a data is redacted, the whole block needs to be replaced.

Later, Deuber et al. proposed a history dependent redactable scheme [11], which is dependent on the history of the transaction data. They propose a redactable protocol in the permissionless setting, which can be integrated to Bitcoin easily. The scheme preserves the original transaction merkle root to keep the hash linked if a block is altered. In their scheme, complex cryptographic tools or trusted parties are not used, and the experiment proves that their scheme is an efficient and feasible protocol. However, in this scheme, consensus-based voting is used during the process of redaction, and this will lead to the privacy disclosure of the redactor. Moreover, this scheme also realizes data redaction at the block-level, but not on the transaction-level. Once a data is redacted, the whole block needs to be replaced.

Most of the redactable blockchains are realized on the block-level, which means that transactions can only be redacted by a block as a unit. Recently, a transaction-level redactable blockchain was proposed [18], where only the hash collision of the redacted transaction should be found. They propose a policy-based chameleon hash, which combines the ciphertext-policy attributed based encryption (CP-ABE) algorithm with a chameleon hash scheme. In their scheme, the attributes of a user should satisfy the access structure of the CP-ABE algorithm [18]. The chameleon hash collision can only be found when a user obtains the short-term and the long-term trapdoors at the same time.

However, all of the solutions to blockchain redaction ignore the issue of data privacy disclosure. Recently, a deletable blockchain was proposed [19], where the identities and transactions of the users are well protected when a block is deleted. In order to delete a block in an anonymous environment, the real identities or the transaction data are disclosed through traceable ring signature or Petersen commitment. Moreover, a linkable multisignature scheme was proposed to prevent the disclosure of the users' identities, which can trace an adversary if he generates a signature more than one time. However, the block can only be deleted by the sender of a transaction but cannot be redacted. Thus, it is urgent and meaningful to propose a redactable blockchain without disclosing the identities and transaction data of the users.

1.3. Our Contributions

- (1) We propose a privacy-preserving redactable blockchain protocol without disclosing the identities and transaction data of the users. A threshold ring signature and a symmetric encryption algorithm are separately used to protect the users' identities and transaction data.
- (2) The proposed blockchain is a fine-grained transaction-level redactable one. The block data can be redacted on the transaction level instead of block level, which means that other transactions in the block do not need to be changed when one transaction is redacted.
- (3) We formalize four security requirements for a privacy-preserving redactable blockchain including identity privacy, data privacy, chain quality, and common prefix and prove that our proposed protocol is provably secure based on symmetric encryption and threshold ring signature.
- (4) The efficiency of the proposed protocol is demonstrated by a series of experiments. The result shows that the time of redacting a transaction is about twice that of creating a block, and the proposed protocol can save about 20% of computational costs than the previous ones.

2. Preliminaries

Some definitions and algorithms utilized in the proposed protocol are introduced in this section.

2.1. An Introduction of Blockchain. Assume that B_{j-1}, B_j, B_{j+1} are three adjacent blocks in a blockchain. As shown in Figure 1, the j -th block is denoted as $B_j = \langle s_j, x_j, \text{ctr}_j \rangle$. The block can point to its previous block through the hash value $s_j \in \{0, 1\}^k$. All of the transactions packaged in a block are contained in $x_j \in \{0, 1\}^*$, and $\text{ctr} \in \mathbb{N}$ is the result of the PoW puzzle [20]. The users check whether the block B_j is valid or not by the following inequality:

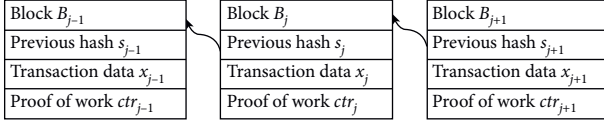


FIGURE 1: Block structure in a blockchain.

$$\text{validateBlockD}(B_j) = H_1(\text{ctr}_j, H_2(s_j, x_j)) < D, \quad (1)$$

where the parameter D is the target difficulty coefficient of the PoW consensus mechanism. $H_1, H_2: \{0, 1\}^* \rightarrow \{0, 1\}^k$ are two different cryptographic hash functions.

A blockchain is composed of a series of blocks, which have solved the PoW puzzle. The newly created block points to its previous block through the hash value. We denote the length of a chain $\text{len}(\mathcal{C})$, and the rightmost block of chain \mathcal{C} can be denoted as $\text{len}(\mathcal{C}) = \langle x', s', \text{ctr}' \rangle$. It should be noted that a blockchain can be an empty chain, which is denoted as $\mathcal{C} = \mathcal{E}$.

The chain \mathcal{C} will be updated as $\mathcal{C}' = \mathcal{C} \parallel B_{j+1}$ if a new block B_{j+1} is generated and validated by the users, and at present, the last block of the chain \mathcal{C}' is the block B_{j+1} . For any $q \geq 0$, \mathcal{C}^q is denoted as the chain after deleting the rightmost q blocks. Note that \mathcal{C}^q can be an empty chain if q is larger than n , where n is the length of a chain \mathcal{C} . If the prefix of chain \mathcal{C}_1 is \mathcal{C}_2 , we write $\mathcal{C}_1 \preceq \mathcal{C}_2$.

2.2. Redactable Protocol in the Public Blockchain. In this subsection, we introduce a redactable blockchain protocol [11], which will be used in our paper. The main contribution is that it preserves the old state $o_j = H_2(s_j, x_j)$ of a block, which is actually the merkle root of the old transactions. As shown in Figure 2, a block is linked to its predecessor by two links, an old hash link (solid arrow) and a new hash link (dashed arrow), and $s_{j+1} = H_1(\text{ctr}_j, H_2(s_j, x_j), o_j)$ holds [11]. When the block B_j is redacted, the new hash link between the block B_j and B_{j+1} is broken (marked by a red cross) because $s_{j+1} \neq H_1(\text{ctr}_j, H_2(s_j, x_j^*), o_j)$. However, B_{j+1} can still point to B_j successfully through the old hash link (solid arrow) since the merkle root of the old transaction (old state) is still preserved. The reason is shown as follows:

$$\begin{aligned} s_{j+1} &= H_1(\text{ctr}_j, o_j, o_j) \\ &= H_1(\text{ctr}_j, H_2(s_j, x_j), o_j). \end{aligned} \quad (2)$$

Please see reference [11] for the details.

2.3. Threshold Ring Signature. The following algorithms are commonly included in a (t, n) threshold ring signature (TRS) scheme, where t ($0 \leq t \leq n$) is the value of the threshold, n is the number of the ring members, and $t < n$ [21]:

TRS_Setup: on input security parameter λ , it outputs the public keys pk_1, \dots, pk_n and private keys sk_1, \dots, sk_n of n ring members

TRS_Sign: on input a message $m \in \{0, 1\}^*$, a ring containing n members with n public keys pk_1, \dots, pk_n ,

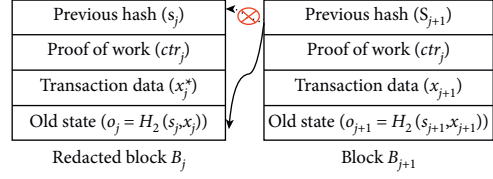


FIGURE 2: The redacted block can still be linked by the old state.

and private keys sk_1, \dots, sk_n of t members, it outputs a threshold ring signature σ on the message m

TRS_Verify: on input a message m , a signature σ and public keys pk_1, \dots, pk_n of n members in a ring, it outputs 0 or 1 to indicate accepting or rejecting the signature

3. Privacy-Preserving Redactable Blockchain

This section describes a privacy-preserving redactable blockchain protocol Γ' , which can redact the block data on the transaction-level without disclosing the transaction data and the identities of the users.

3.1. Syntax of Privacy-Preserving Redactable Blockchain. We construct a privacy-preserving redactable blockchain protocol Γ' , which contains the following four algorithms:

$\{0, 1\} \leftarrow \Gamma'.\text{validateReq}(\mathcal{C}, j, i, tx_i^*, m, K)$: on inputting a chain \mathcal{C} , an index j of the redacted block, an index i of the redacted transaction, an initial candidate transaction tx_i^* , the old transaction data m , and the old encryption key K , it returns 1 if the redaction request is valid; otherwise, it returns 0.

$\{0, 1\} \leftarrow \Gamma'.\text{validateCandTx}(\mathcal{C}, tx_i^{**}, j)$: on inputting a chain \mathcal{C} , a final candidate transaction tx_i^{**} and an index j , it returns 0 if tx_i^{**} is valid; otherwise, it returns 1.

$\{0, 1\} \leftarrow \Gamma'.\text{validateChain}(\mathcal{C})$: on inputting a chain \mathcal{C} , it returns 1 if the chain is valid; otherwise, it returns 0.

$\{0, 1\} \leftarrow \Gamma'.\text{validateBlock}(B)$: on inputting a block B , it returns 1 if the block is valid; otherwise, it returns 0.

3.2. Our Proposed Protocol. In the proposed protocol, a regular transaction is denoted as $tx = \langle E_K(m), \text{RSig} \rangle$, where m is the transaction data including the input and output of a transaction, $E(\cdot)$ is a symmetric encryption algorithm, $E_K(m)$ is the ciphertext of the transaction data using a key K and RSig is a ring signature to protect the identity privacy of the users. $H: \{0, 1\}^* \rightarrow \{0, 1\}^k$ is a cryptographic hash function. The block data is denoted as $x = \{tx_1, tx_2, \dots, tx_p\}$, where p is the number of the transactions packaged in each block.

3.2.1. Proposing a Redaction Request. Once a user joins in the blockchain system, he needs to generate a public key as his identity, and a private key, which will be used to generate a signature. If a miner wants to participate in a redaction

process, he can apply for the key generation algorithm of the threshold ring signature system to get another key pair when he joins the blockchain system. The key generation algorithm is executed by a trusted party, such as certification authority (CA). It should be noted that the private key generated by the user as the identity is different from that generated by the TRS system, where the former is used to generate the ring signature R_{Sign} to protect the identity of the users, and the latter is used to generate the threshold ring signature TRS to allow redacting a transaction in a block.

If a user wants to propose a redaction request for the transaction tx_i in block B_j , he discloses the transaction encryption key K and the encrypted transaction data m first. Then, he generates a new transaction mt , picks up a new encryption key Kt , and calculates $H(m, Kt, mt, K)$, $E_{Kt}(mt)$, ring signature R_{Sign} t and a ring signature S on $H(m, Kt, mt, K)$. Next, the user generates an initial candidate transaction $tx_i^* = \langle E_{Kt}(mt), R_{Sign}t, H(m, Kt, mt, K), S \rangle$. Finally, the user broadcasts the redaction request $r_{j,i} = \langle j, i, tx_i^*, m, K \rangle$ to the whole network. Note that two ring signatures are both generated by the data owner, who randomly chooses two rings including some users on the chain, and compute the signatures by the public keys of these users.

3.2.2. Joining a Redaction Operation. When a miner in the blockchain receives a redaction request, he invokes Algorithm 1 Γ' .validateReq to judge whether the redaction request is valid or not. If the redaction request is valid, he votes for the redaction by putting his vote result in the latest block he mines. If the number of miners who approve of the redaction request exceeds the threshold t of the threshold ring signature, these miners generate a threshold ring signature TRS jointly on the redaction message. Finally, a final candidate transaction tx_i^{**} is generated and broadcasted to the network, where $tx_i^{**} = tx_i^* \parallel TRS$.

3.2.3. Validating a Redaction Operation. Everyone can verify whether a final candidate transaction is valid or not by invoking Γ' .validateCandTx (Algorithm 2). If the final candidate transaction tx_i^{**} is valid, the user replaces the old transaction tx_i in block B_j with the final candidate transaction tx_i^{**} in their local chain. The process of redacting a transaction is shown in Figure 3.

4. Protocol Description

In this subsection, we introduce each algorithm in the proposed protocol in detail.

Algorithm 1 can validate whether a redaction request is valid or not. The miners generate a threshold ring signature if the validation of step 5 to step 9 all passed.

Algorithm 2 can validate whether a final candidate transaction is valid or not. This algorithm first checks whether the redacted block B_j^* is valid or not. In line 6, Algorithm 2 checks the validity of the threshold ring signature TRS. In line 11, it checks the old hash link from B_j^* to B_{j-1} and the hash link from B_{j+1} to B_j^* . The final candidate

transaction can be considered valid if the hash link of block B_j can point to the previous block correctly. The user can replace the old transaction with the final candidate transaction in his local chain.

Algorithm 3 describes how to validate a chain \mathcal{C} in our system. This algorithm checks the chain \mathcal{C} from the beginning to the end. Users only have to validate the head of the chain \mathcal{C} if the length of a chain is 1. Otherwise, blocks should be validated one by one. In line 8, the users first check the current link between the neighboring blocks; if the link is broken, then the users check the old link in line 10. The users accept the block if the old link can point to the previous block successfully, and the candidate block is valid.

Algorithm 4 describes how to validate a block B in our system. All of the transactions in the block should be checked using some predefined validation rules. In line 2, this algorithm checks whether the block B has solved the PoW puzzle or not. If the block has been redacted before, the old state of B should be checked to judge whether the block has solved the PoW puzzle or not.

5. Security Analysis

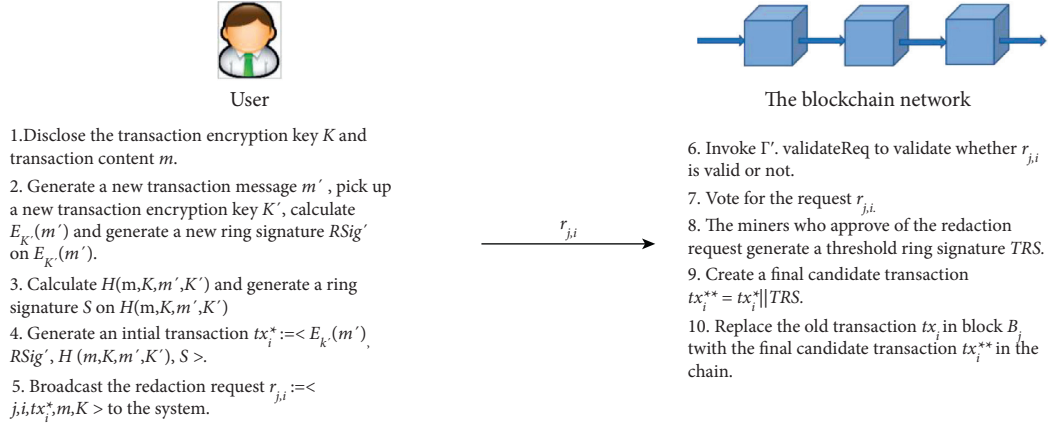
The following contents provide security analysis of our privacy-preserving redactable blockchain protocol Γ' . The original public redactable blockchain protocol Γ satisfies two security requirements: chain quality and common prefix. Supposing that the protocol is a secure and reliable system, the proposed protocol Γ' also satisfies these two requirements. Most importantly, Γ' can protect the users' identities and transaction data simultaneously. The following description will clearly illustrate that the proposed protocol is provably secure based on threshold ring signature, symmetric encryption, and the blockchain protocol proposed in [11].

5.1. Identity Privacy. The subsection states that the users' identities are private in our protocol.

Theorem 1. *The redactable blockchain scheme realizes the privacy of the users' identities if the threshold ring signature is anonymous.*

Proof. Proof. The game is executed between an adversary \mathcal{A} and a simulator \mathcal{B} . At the beginning of the game, \mathcal{B} is given a tuple $(Y, V_0, V_1, M, TRS(M, Y, V_b))$, where $b \in \{0, 1\}$, V_0, V_1 are both subsets of Y with same size, M is the redaction message including the redaction reasons and timestamp, and $TRS(M, Y, V_b)$ is the threshold ring signature generated by V_b . Assuming that adversary \mathcal{A} obtains the users' identities with a nonnegligible probability in the proposed protocol, simulator \mathcal{B} can obtain the identities of the real signers with a nonnegligible probability in the threshold ring signature:

Setup. The simulator sets up the blockchain system and sends the public parameters $(TRS(\cdot), E(\cdot), H(\cdot))$ to the adversary, where $TRS(\cdot)$ is a threshold ring signature

FIGURE 3: Proposing a redaction for the transaction tx_i of block B_j .

Input: Blockchain \mathcal{C} , redaction request $r_{j,i} = \langle j, i, tx_i^*, m, K \rangle$

Output: 0/1

- (1) Parse the original j -th block $B_j^* = \langle s_j, x_j^*, ctr_j, o_j \rangle$;
- (2) Parse the old block data $x_j = \{tx_1, tx_2, \dots, tx_p\}$;
- (3) Replace the old transaction tx_i with the initial candidate transaction tx_i^* ;
- (4) Parse the new block data $x_j^* = \{tx_1, tx_2, \dots, tx_i^*, \dots, tx_p\}$;
- (5) Parse the redacted j -th block $B_j^* = \langle s_j, x_j^*, ctr_j, o_j \rangle$;
- (6) If **vali da teBlock**(B_j^*) = 0, return 0;
- (7) Parse the initial candidate transaction $tx_i^* = \langle E_{K'}(m'), RSig', H(m, K, m', K'), S \rangle$.
- (8) Calculate $E_K(m)$ using the disclosed transaction data m and encryption key K .
- (9) Compare $E_K(m)$ with the ciphertext in the original transaction. If not equal, return 0;
- (10) Validate the ring signature S contained in the redacted transaction tx_i^* and compare the message signed by S with the digest $H(m, K, m', K')$, if not equal return 0;
- (11) Return 1.

ALGORITHM 1: **vali da teReq** (implements $\Gamma'.\text{vali da teReq}$).

Input : Chain \mathcal{C} , a final candidate transaction tx_i^{**} , an index j of the block

Output : 0/1

- (1) Parse $x_j^* = \{tx_1, tx_2, \dots, tx_i^*, \dots, tx_p\}$;
- (2) Parse $B_j^* = \langle s_j, x_j^*, ctr_j, o_j \rangle$;
- (3) Parse $tx_i^{**} = tx_i^* || TRS$.
- (4) if **vali da teBlock**(B_j^*) = 0 then
- (5) return 0;
- (6) if **TRS** is invalid then
- (7) return 0;
- (8) Compare the message signed by **TRS** and the hash of the redaction request, If not equal, return 0;
- (9) Parse $B_{j-1} = \langle s_{j-1}, x_{j-1}, ctr_{j-1}, o_{j-1} \rangle$;
- (10) Parse $B_{j+1} = \langle s_{j+1}, x_{j+1}, ctr_{j+1}, o_{j+1} \rangle$;
- (11) if $s_j^* = H(ctr_{j-1}, o_{j-1}, o_{j-1}) \wedge s_{j+1} = H(ctr_j, o_j, o_j)$ then
- (12) return 1;
- (13) else
- (14) return 0;

ALGORITHM 2: **vali da teCan dT x**(implements $\Gamma'.\text{vali da teCan dT x}$).

```

Input: Chain  $\mathcal{C} = (\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_n)$  of length  $n$ .
Output: 0/1
(1)  $j := n$ ;
(2) if  $j = 1$  then
(3) return vali da teBlock( $\mathbf{B}_1$ );
(4) while  $j \geq 2$  do
(5) Parse  $\mathbf{B}_j = \langle s_j, x_j, \text{ctr}_j, o_j \rangle$ ;
(6) if vali da teBlock( $\mathbf{B}_j$ ) = 0 then
(7) return 0
(8) if  $s_j = \mathbf{H}_1(\text{ctr}_{j-1}, \mathbf{H}_2(s_{j-1}, x_{j-1}), o_{j-1})$  then
(9)  $j = j + 1$ ;
(10) else if  $s_j = \mathbf{H}_1(\text{ctr}_{j-1}, o_{j-1}, o_{j-1})$  then
(11)  $j = j + 1$ ;
(12) else
(13) return 0;
(14) return 1;
(15) return 1

```

ALGORITHM 3: **vali da teChain**(implements Γ .**vali da teChain**).

```

Input: Block  $\mathbf{B} = \langle s, x, \text{ctr}, o \rangle$ 
Output: 0/1
(1) if all the transactions in  $x$  are valid then
(2) if  $\mathbf{H}_1(\text{ctr}, \mathbf{H}_2(s, x), o) < D \vee \mathbf{H}_1(\text{ctr}, o, o) < D$  then
(3) return 1;
(4) else
(5) return 0;
(6) else
(7) return 0;

```

ALGORITHM 4: **vali da teBlock**(implements Γ .**vali da teBlock**).

scheme, $E(\cdot)$ is a symmetric encryption algorithm, and $H(\cdot)$ is a secure hash function.

Oracle Queries. The adversary simulates the following oracles:

Transaction Generation Oracle. Adversary \mathcal{A} sends an index j , some transactions collected from the blockchain system. The simulator \mathcal{B} returns a block $tx_i := \langle E_K(m_i), \text{RSig}_i \rangle$.

Transaction Redaction Oracle. Adversary \mathcal{A} sends a block index j , a transaction index i , an initial candidate transaction tx_i^* , an old transaction message m and an old transaction encryption key K , a redaction message M and two sets Y and V to the simulator, where Y is a public key set, V is a subset of Y and $V \neq V_0$ or V_1 . The simulator returns the corresponding final candidate transaction $tx_i^{**} = tx_i^* \parallel \text{TRS}(M, Y, V)$.

Challenge. The simulator \mathcal{B} generates the final candidate transaction $\widehat{tx}_i^* = \widehat{tx}_i^* \parallel \text{TRS}(M, Y, V_b)$, where M contains the disclosed transaction data m and encryption key K included in transaction tx_i^* . Then, he transmits \widehat{tx}_i^{**} to the adversary \mathcal{A} . \mathcal{A} outputs $b \in \{0, 1\}$ and then \mathcal{B} submits $b = b$. Assuming that \mathcal{A} can guess the users' identities successfully with a probability of more than $1/2$, the simulator \mathcal{B} can

distinguish the identities of the real signers with a probability of more than $1/2$, and it contradicts the anonymity of the threshold ring signature. Therefore, it is impossible for the adversary \mathcal{A} to guess the identities of the signers in a threshold ring signature successfully with a probability of more than $1/2$, and the proposed redactable protocol can protect the users' identities effectively. \square

5.2. Data Privacy. The subsection states that the transaction data in our blockchain system is private.

Theorem 2. *The redactable blockchain scheme realizes the privacy of transaction data if the symmetric encryption algorithm is indistinguishable against chosen plaintext attack (IND-CPA) secure.*

Proof. The game is executed between an adversary \mathcal{A} and a simulator \mathcal{B} . Assuming that adversary \mathcal{A} can guess the transaction data with a nonnegligible probability in the proposed protocol, the simulator \mathcal{B} can distinguish the plaintext with a nonnegligible probability in the symmetric encryption scheme:

Setup. The simulator sets up the blockchain system and generates a symmetric encryption algorithm E and an encryption key K and then sends the public parameters E to the adversary. The encryption key K is kept secret itself.

Oracle Queries. The adversary simulates the following oracles:

Encryption Oracle. The adversary sends the plaintext of transaction data m to the simulator. The simulator returns the corresponding ciphertext tx : $= \langle E_K(m), \text{RSig} \rangle$, where $E_K(m)$ is the ciphertext of m by using the symmetric encryption algorithm, and RSig is the ring signature of the simulator on the ciphertext $E_K(m)$.

Challenge. Adversary \mathcal{A} sends a tuple $(m_0, m_1, K_0, K_1, E_{K_b}(m_b))$ to the simulator. Then, \mathcal{B} randomly chooses $b \in \{0, 1\}$ and computes $E_{K_b}(m_b)$. Note that $E_{K_b}(m_b)$ has never been queried in the Encryption oracles. Next, the simulator \mathcal{B} generates the ring signature RSig^* on the ciphertext $E_{K_b}(m_b)$ and transmits the transaction ciphertext $tx^* = \langle E_{K_b}(m_b), \text{RSig}^* \rangle$ to adversary \mathcal{A} .

\mathcal{A} outputs $br \in \{0, 1\}$ and then \mathcal{B} submits $b = br$. Assuming that \mathcal{A} can guess the transaction data successfully with a probability of more than $1/2$, the simulator \mathcal{B} can distinguish the ciphertext of two plaintexts with a probability of more than $1/2$, and it contradicts the IND-CPA security of the symmetric encryption schemes. Therefore, it is impossible for the adversary \mathcal{A} to guess the plaintext of transaction data successfully with a probability of more than $1/2$, and the proposed redactable protocol can protect the transaction data effectively. \square

5.3. Chain Quality. This property points out that the proportion of adversary blocks in a blockchain is limited by the computing capabilities held by the adversaries.

Definition 1. The property of (μ, ℓ) chain quality is parameterized by $\mu \in \mathbb{R}$ and $\ell \in \mathbb{N}$, which states that, for any part of the chain composed by ℓ continuous blocks held by the honest party, the proportion of adversarial blocks is at most μ , where μ is used to evaluate the computing capabilities of the adversaries.

Theorem 3. Γ' satisfies (μ, ℓ) -chain quality if the public immutable blockchain satisfies (μ, ℓ) -chain quality for any (t, n) threshold ring signature where $t/n > \mu$.

Proof. Different from the public immutable blockchain, we provide the function of transaction redaction in an anonymous environment. In order to increase the proportion of malicious blocks in the chain and break the chain quality property, adversary \mathcal{A} could redact an honest transaction tx_i of block B_j into a malicious transaction tx_i^{**} . We prove that the probability of replacing an honest transaction with a malicious one by the adversary \mathcal{A} is negligible.

Suppose that adversary \mathcal{A} proposes an initial candidate transaction tx_i^* to an honest transaction tx_i of block B_j . Limited by computing resources, the adversary \mathcal{A} can only mine μ ratio of blocks during the phase of threshold ring signature generation. However, the proposed scheme requires that the ratio of redaction supporters to generate the threshold ring signature has to be at least t/n , where $t/n > \mu$. The adversary \mathcal{A} cannot be approved by t honest users and generate a valid (t, n) threshold ring signature TRS. Therefore, the adversary \mathcal{A} needs to create an ‘‘honest looking’’ initial candidate transaction $\widetilde{tx}_i^* \neq tx_i^*$ such that $H(\widetilde{tx}_i^*) = H(tx_i^*)$, where H is a collision-resistant hash function. Next, the adversary \mathcal{A} deceives the honest users, and the honest users ensure the initial candidate transaction \widetilde{tx}_i^* by generating a threshold ring signature TRS. Then, the adversary \mathcal{A} creates the final candidate transaction $\widetilde{tx}_i^{**} = \widetilde{tx}_i^* \parallel \text{TRS}$ successfully and redacts the chain with the ‘‘honest looking’’ final candidate transaction tx_i^{**} . However, the chance of creating such a transaction tx_i^* is negligible, since it violates the collision-resistance property of the hash function H . To sum up, the proof of Theorem 3 is established. \square

5.4. Common Prefix. The property of common prefix is parameterized by $k \in \mathbb{N}$, which states that if there are two honest users, and they hold two different chains C_1 and C_2 respectively, the number of the far right different blocks of these two chains is at most k . The game is also executed between an adversary \mathcal{A} and a simulator \mathcal{B} .

Definition 2. For any two different chains C_1 and C_2 held by two honest users U_1 and U_2 respectively, it holds that

$$\begin{aligned} C_2^k &\preceq C_1, \\ C_1^k &\preceq C_2. \end{aligned} \quad (3)$$

As shown in (3), C_1^k and C_2^k are denoted as the chain resulting from deleting the k far right blocks of chain C_1 and C_2 .

However, Definition 2 cannot be used directly in our new redactable protocol Γ' . Consider that if a redaction request $r_{i,j}$ and the corresponding final candidate transaction tx_i^{**} were approved, an honest user U_1 updates the chain C_1 at time t_1 and he replaces the old transaction tx_i with tx_i^{**} in his local chain. However, another honest user U_2 may not update the chain C at time t_1 timely. In this case, $C_1^k \not\preceq C_2$ and $C_2^k \not\preceq C_1$, which violates Definition 3.

Therefore, a new definition was described as follows for the new privacy-preserving redactable protocol Γ' . Note that the redacted block is denoted as B_j^* .

Definition 3. For any two different chains C_1 and C_2 held by two honest users U_1 and U_2 respectively, it holds that

- (1) $C_1^k \preceq C_2$ and $C_2 \preceq C_1$
- (2) For any redacted block $B_j^* \in C_1^{(l_2-l_1)+k}$ and $B_j^* \notin C_2^k$, or block $B_j^* \in C_2^{(l_2-l_1)+k}$ and $B_j^* \notin C_1^k$, it satisfies $\Gamma'.\text{validateCandTx}(\mathcal{C}, tx_i^{**}, j) = 1$

Theorem 4. Γ' satisfies the property of k -editable common prefix if the public immutable blockchain satisfies the property of k -common prefix.

Proof. Proof. Assume that the adversary \mathcal{A} proposes a final candidate transaction tx_i^{**} to redact an honest block B_j in chain C_2 . The chain C_2 is redacted by an honest user U_1 later. However, the adversary \mathcal{A} cannot create another final candidate transaction $tx_i^{**} \neq tx_i^{**}$ such that $H(tx_i^{**}) = H(tx_i^{**})$ because of the collision-resistance property of hash function H . Since the final candidate transaction tx_i^{**} is accepted by the honest user U_1 , it must be the case that tx_i^{**} contains a valid threshold ring signature TRS and is approved by most of the honest users in the system. To sum up, the proof of Theorem 4 is established. \square

6. Experiments

Several experiments are executed to test the efficiency of our work. We give the time of block redaction and block generation and compare the efficiency of the proposed protocol with that of the previous ones. We mainly focus on the time-consuming operations in these protocols, and the most expensive operations for generating and redacting a block are separately solving the PoW puzzle and generating a threshold ring signature.

We conduct our experiments on a computer with a 64-bit Win 10 operating system, 8.0 GB RAM and Intel(R) Core(TM) i7-5500 CPU @ 2.4 GHz processor. We use the IntelliJ IDEA 2020 compiler and Java language to implement our programs. The JPBC 2.0.0 library is used to generate elliptic curve groups and pairings. The elliptic curve groups and pairings are used in the threshold ring signature scheme [22]. The AES algorithm is used to encrypt the transaction data.

We test multiple times to get the average values as the final measurement results from each experiment. It should be noted that the number of the consecutive zero of the hash prefixes is defined as the difficulty level. For example, the difficulty level is 5 when the number of consecutive zeros of the hash prefixes is 5.

6.1. Time Consumption of Generating a Block and Redacting a Transaction. We first test the computational overhead for generating a block. As shown in Figure 4, the overhead of generating a block is almost a constant, i.e., 2.894 s and 15.127 s, when the difficulty level is set as 4 and 6, respectively. The reason is that the overheads of generating a block depend on the difficulty of the POW puzzle instead of the percentage of the users participating in the redaction operation. We also test the overheads for redacting a transaction, which is much more important, and the time to redact a transaction actually determines the efficiency of our proposed scheme. As shown in Figure 5, the time consumption increases when the percentage of the users participating in a redaction operation grows, and the time consumption of redacting a transaction is not affected by the difficulty. The reason is that the time of redacting a

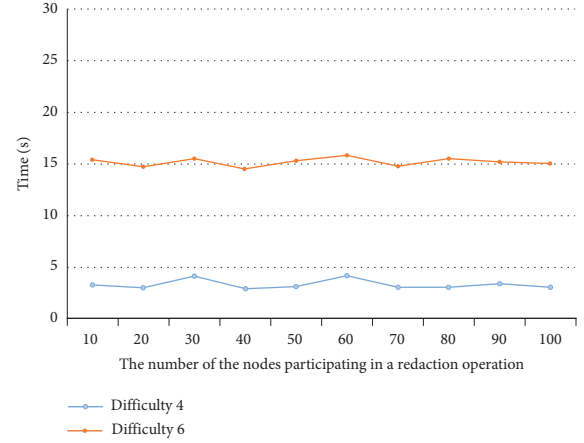


FIGURE 4: Time consumption of generating a block.

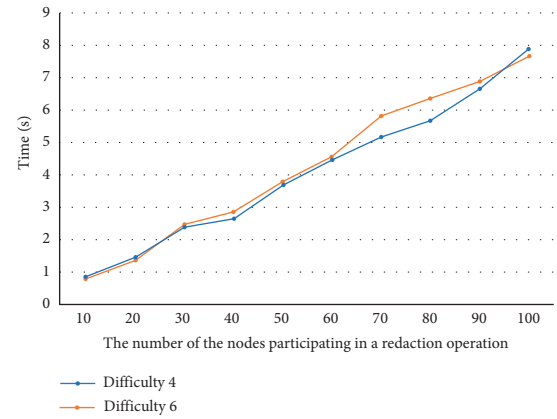


FIGURE 5: Time consumption of redacting a transaction.

transaction is mainly determined by the time of generating a (t, n) threshold ring signature, which depends on the value of threshold t .

In Table 1, we give the time of redacting a transaction when the difficulty level is 5 and the percentage of the users participating in the redaction operation is 80%. The experimental results indicate that the average time of generating a block and redacting a transaction is separately 4580 ms and 7088 ms, and the time of redacting a transaction is about twice that of generating a block. Therefore, redacting data on a block is efficient in the proposed protocol.

In order to describe the change of the block structure before and after the block is redacted more clearly, the block 23 is used as an example to illustrate what happens when a block is redacted. The original block information from block 22 to block 24 in the current blockchain is shown in Figure 6. Suppose that the second transaction in block 23 needs to be redacted. In order to redact the transaction, the owner of transaction tx_2 discloses the encrypted transaction data and the encryption key as the redacting reasons, and then a new transaction tx_2^* is generated to replace the old transaction tx_2 . Moreover, the POW puzzle of the new block 23* is solved. The users in the blockchain reached a consensus to

TABLE 1: The time consumption of block generation and redaction.

Time (ms)	Block1	Block2	Block3	Block4	Block5	Average
Block generation	4875	4350	4586	4109	4217	4580
Block redaction	7016	6889	6875	6959	6830	7088

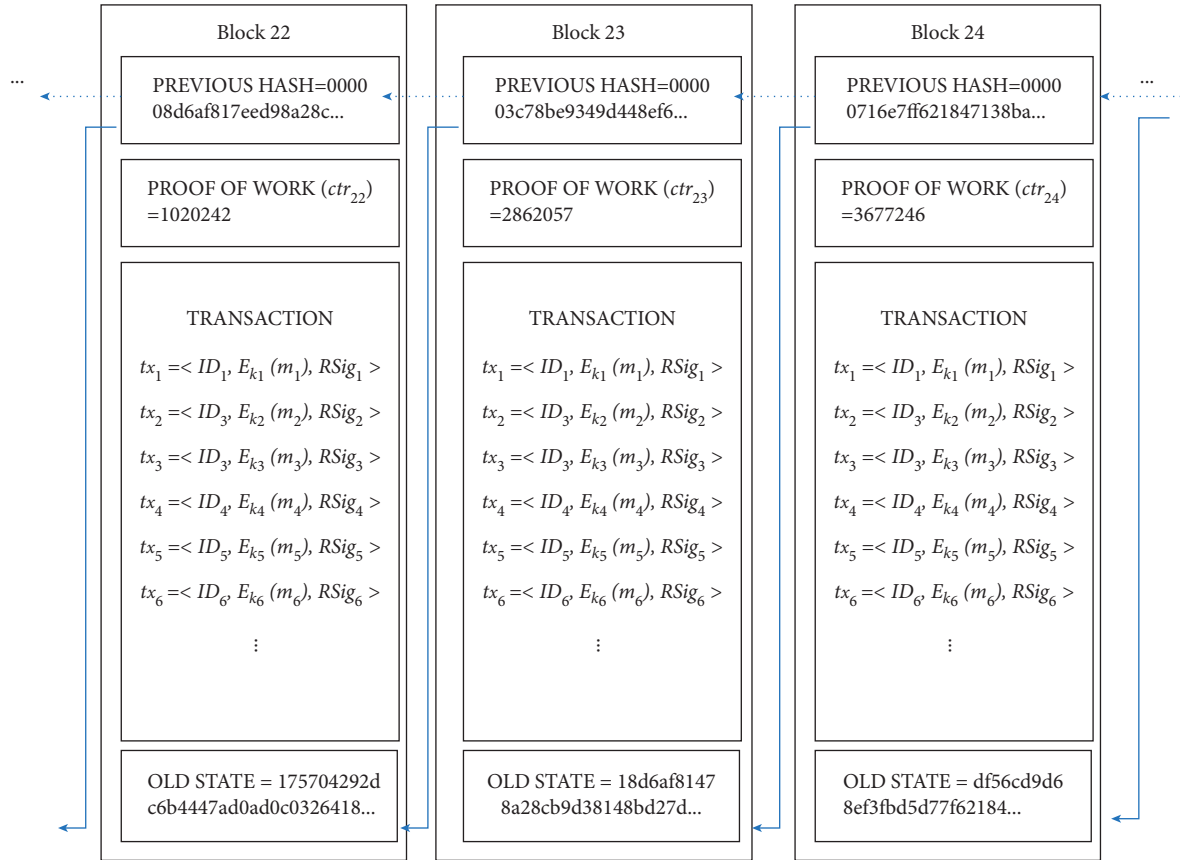


FIGURE 6: The original block information from block 22 to block 24.

redact the block and generate a threshold ring signature TRS to approve the redaction operation. The new redacted block information from block 22 to block 24 is shown in Figure 7.

6.2. Efficiency Comparison. This experiment is intended to compare the efficiency of our proposed redactable protocol with that of [11, 19], where a privacy-preserving deletable blockchain and a public redactable blockchain are proposed.

In the following experiments, we generate redactable and deletable chains. The length of these two chains and the number of block modifications are the same in each experiment. As shown in Figure 8, the number of block redactions or deletions ranges from 10 to 40, and the overhead ratio is from 76% to 84%, and thus the average time of a redaction in the proposed protocol is about 80 percent of that in the deletable chains [19]. The reason is that our protocol needs to generate a ring signature only once, and the protocol in [19] needs to generate a traceable ring signature twice. Therefore, a transaction redaction is more

efficient in the proposed protocol than that of deleting a block in [19].

To compare the efficiency of our work and that of [11], we both generate privacy-preserving redactable and public redactable chains. The length of these two chains and the number of block modifications are kept the same. As shown in Figure 9, the number of redactions ranges from 10 to 40, and the overhead ratio is from 30% to 41%, and thus the average time of redacting a transaction in the proposed protocol is about three times that in the public chains [11]. The reason is that our protocol uses more time-consuming operations, including the threshold ring signature and symmetric encryption in the redacting process to protect identities and data of the users. Therefore, the proposed protocol spends more time than the protocol in [11].

Next, we give a time comparison of redaction for these three protocols, and the results are shown in Table 2. We can see that, in the protocol of [19], it takes an average of 8904 ms to delete a block. However, the protocol can only delete the whole block by the data owner, and it cannot redact a single

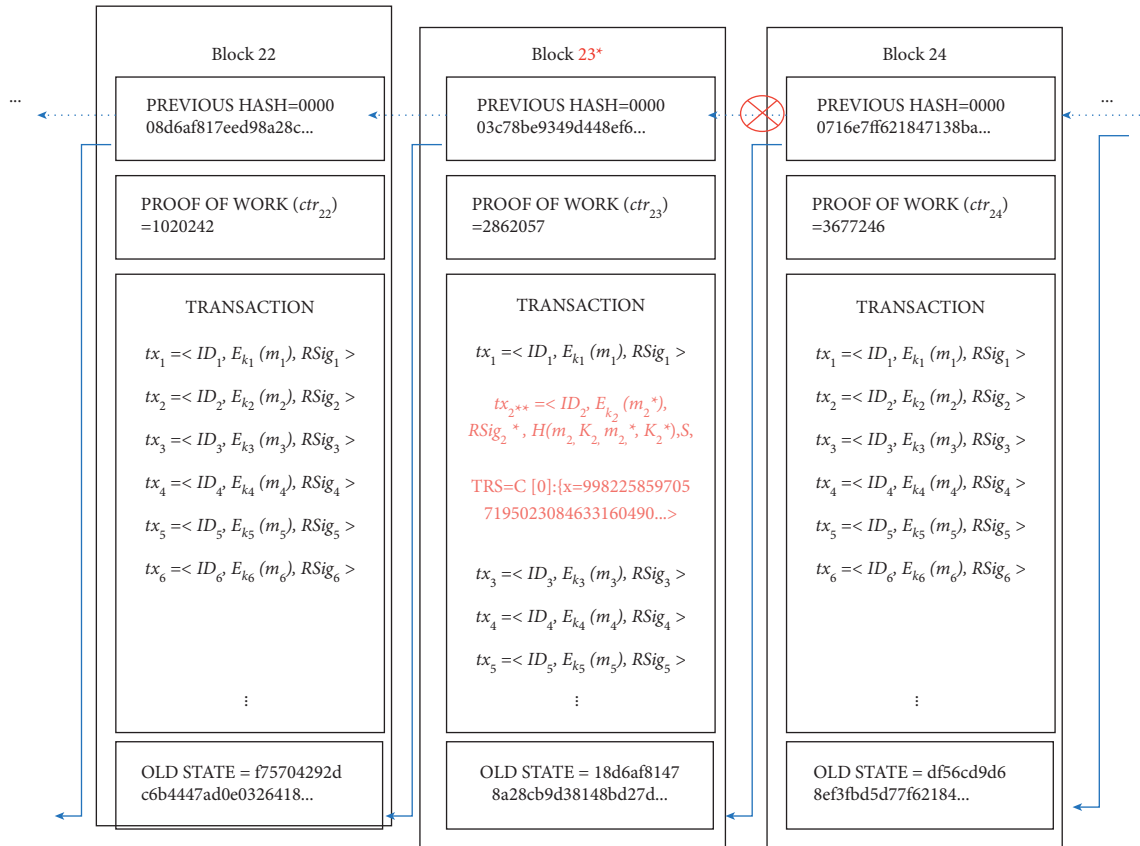


FIGURE 7: The new redacted block information from block 22 to block 24.

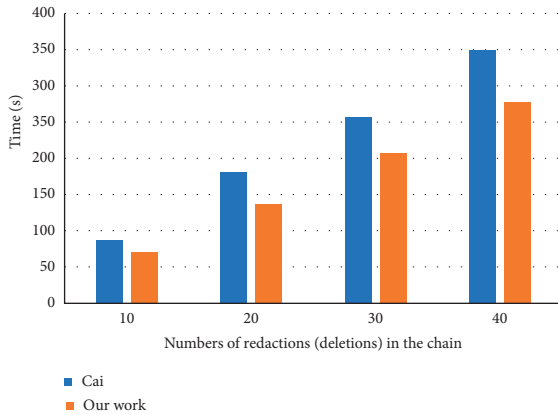


FIGURE 8: Block redaction overhead compared to deletable blockchain [19].

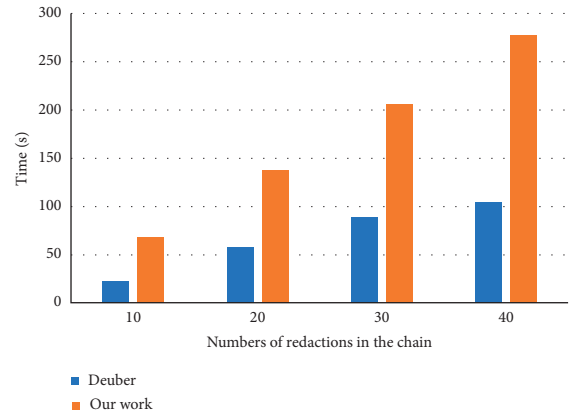


FIGURE 9: Transaction redaction overhead compared to the public redactable blockchain [11].

transaction in the block. In our protocol, it takes an average of 7088 ms to redact a transaction, which is about 79.6% of the time for deleting a block in [19]. In the protocol of [11], it takes an average of 2552 ms to redact a transaction, which is about 36.0% of the time for redacting a transaction in the proposed redactable blockchain. However, the identities of the users and the transaction data are all public in the protocol of [11]. Thus, the proposed protocol can realize the redaction on the transaction-level, while the identities of the users and transaction data are also private.

Finally, we compare the proposed protocol with those in [11, 19], and the results are shown in Table 3. It is concluded that our protocol provides the function of transaction redaction, transaction deletion without disclosing the identities of the users and the transaction data. Although the protocol in [19] realizes the deletion of block data, the transaction redaction is not allowed. In the protocol of [11], transaction redaction is allowed, but the identities of the users and transaction data are public. The proposed protocol constructs a transaction-level redactable blockchain, and the

TABLE 2: The time comparison for block redaction.

Time (ms)	Block 1	Block 2	Block 3	Block 4	Block 5	Average
Cai et al. [19]	8739	9069	8630	8754	9328	8904
Deuber et al. [11]	2194	2889	2970	2645	2072	2552
Our work	7016	6889	6875	6959	6830	7088

TABLE 3: The comparison of three redactable blockchain protocols.

	Redaction	Deletion	Identity privacy	Data privacy	Transaction level	Block level
Cai et al. [19]	×	✓	✓	✓	×	✓
Deuber et al. [11]	✓	✓	×	×	×	✓
Our work	✓	✓	✓	✓	✓	×

users only need to replace a single transaction to complete the data redaction instead of replacing the entire block. However, the protocol [11, 19] can only achieve block-level redactable blockchain.

7. Conclusion

In this paper, we propose a privacy-preserving redactable blockchain based on the old state of the block. A symmetric encryption algorithm and a threshold ring signature are separately used to protect the transaction data and the users' identities during the process of redaction. All of the users can check whether a redaction operation is valid or not according to the threshold ring signature and the old link of the redacted block. The experiments' results indicate that the proposed protocol is efficient and effective, and the identities of the users and the transaction data are also private.

Data Availability

All data are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. U1736120) and Natural Science Foundation of Shanghai (20ZR1419700).

References

- [1] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," manubot, Technology Report, 2019.
- [2] T. Jiang, H. Fang, and H. Wang, "Blockchain-based internet of vehicles: distributed network architecture and performance analysis," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4640–4649, 2018.
- [3] H. N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for internet of things: a survey," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8076–8094, 2019.
- [4] E. F. Jesus, V. R. L. Chicarino, C. V. N. D. Albuquerque, and A. A. D. A. Rocha, "A survey of how to use blockchain to secure internet of things and the stalker attack," *Security and Communication Networks*, vol. 2018, Article ID 9675050, 27 pages, 2018.
- [5] R. Kamath, "Food traceability on blockchain: walmart's pork and mango pilots with ibm," *The Journal of the British Blockchain Association*, vol. 1, no. 1, 2018.
- [6] T. T. Thwin and S. Vasupongayya, "Blockchain-based access control model to preserve privacy for personal health record systems," *Security and Communication Networks*, vol. 2019, Article ID 8315614, 15 pages, 2019.
- [7] T. McGhin, K. K. R. Choo, C. Z. Liu, and D. He, "Blockchain in healthcare applications: research challenges and opportunities," *Journal of Network and Computer Applications*, vol. 135, pp. 62–75, 2019.
- [8] J. Wang, L. Wu, H. Wang, K. K. R. Choo, and D. He, "An efficient and privacy-preserving outsourced support vector machine training for internet of medical things," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 458–473, 2020.
- [9] S. Zeng, Y. Yuan, X. C. Ni, and F. Y. Wang, "Scaling blockchain towards bitcoin: key technologies, constraints and related issues," *Acta Automatica Sinica*, vol. 45, no. 6, pp. 1015–1030, 2019.
- [10] G. Ateniese, B. Magri, D. Venturi, and E. Andrade, "Redactable blockchain—or—rewriting history in bitcoin and friends," in *Proceedings of the IEEE European Symposium on Security and Privacy*, pp. 111–126, IEEE, Paris, France, April 2017.
- [11] D. Deuber, B. Magri, and S. Thyagarajan, "Redactable blockchain in the permissionless setting," in *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 124–138, IEEE, San Francisco, CA, USA, May 2019.
- [12] Q. Feng, D. He, S. Zeadally, M. K. Khan, and N. Kumar, "A survey on privacy protection in blockchain system," *Journal of Network and Computer Applications*, vol. 126, pp. 45–58, 2019.
- [13] S.-F. Sun, M. H. Au, J. K. Liu, and T. H. Yuen, "Ringct 2.0: a compact accumulator-based (linkable ring signature) protocol for blockchain cryptocurrency monero," in *Proceedings of the European Symposium on Research in Computer Security*, pp. 456–474, Springer, Cham, Oslo, Norway, August 2017.
- [14] J. Bonneau, A. Narayanan, A. Miller, J. Clark, J. A. Kroll, and E. W. Felten, "Mixcoin: anonymity for bitcoin with accountable mixes," in *Proceedings of the International Conference on Financial Cryptography and Data Security*, pp. 486–504, Springer, Berlin, Heidelberg, November 2014.
- [15] E. B. Sasson, A. Chiesa, C. Garman et al., "Zerocash: decentralized anonymous payments from bitcoin," in *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 459–474, IEEE, Berkeley, CA, USA, May 2014.

- [16] T. H. Yuen, S. Sun, J. K. Liu et al., “Ringct 3.0 for blockchain confidential transaction: shorter size and stronger security,” in *Proceedings of the International Conference on Financial Cryptography and Data Security*, pp. 464–483, Springer, St. Kitts and Nevis, February 2019.
- [17] G. Ateniese and B. D. Medeiros, “Identity-based chameleon hash and applications,” in *Proceedings of the International Conference on Financial Cryptography*, pp. 164–180, Springer, FL, USA, February 2004.
- [18] D. Derler, K. Samelin, D. Slamanig, and C. Striecks, “Fine-grained and controlled rewriting in blockchains: chameleon-hashing gone attribute-based,” in *Proceedings of the 26th Annual Network and Distributed System Security Symposium, NDSS 2019*, CA, USA, April 2019.
- [19] X. Cai, Y. Ren, and X. Zhang, “Privacy-protected deletable blockchain,” *IEEE Access*, vol. 8, pp. 6060–6070, 2020.
- [20] J. Garay, A. Kiayias, and N. Leonardos, “The bitcoin backbone protocol: analysis and applications,” in *Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 281–310, Springer, Sofia, Bulgaria, April 2015.
- [21] S. S. Chow, L. C. Hui, and S.-M. Yiu, “Identity based threshold ring signature,” in *Proceedings of the International Conference on Information Security and Cryptology*, pp. 218–232, Springer, Shanghai, China, November 2004.
- [22] T. H. Yuen, J. K. Liu, M. H. Au, W. Susilo, and J. Zhou, “Threshold ring signature without random oracles,” in *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, pp. 261–267, Hong Kong, China, March 2011.

Research Article

A Fine-Grained IoT Data Access Control Scheme Combining Attribute-Based Encryption and Blockchain

Xiaofeng Lu ¹, Songbing Fu ¹, Cheng Jiang ¹ and Pietro Lio ²

¹Beijing University of Posts and Telecommunications, Beijing 100876, China

²Computer Laboratory, University of Cambridge, Cambridge, UK

Correspondence should be addressed to Xiaofeng Lu; luxf@bupt.edu.cn

Received 11 June 2021; Revised 30 July 2021; Accepted 23 August 2021; Published 16 September 2021

Academic Editor: AnMin Fu

Copyright © 2021 Xiaofeng Lu 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.

IoT technology has been widely valued and applied, and the resulting massive IoT data brings many challenges to the traditional centralized data management, such as performance, privacy, and security challenges. This paper proposes an IoT data access control scheme that combines attribute-based encryption (ABE) and blockchain technology. Symmetric encryption and ABE algorithms are utilized to realize fine-grained access control and ensure the security and openness of IoT data. Moreover, blockchain technology is combined with distributed storage to solve the storage bottleneck of blockchain systems. Only the hash values of the data, the hash values of the ciphertext location, the access control policy, and other important information are stored on the blockchain. In this scheme, smart contract is used to implement access control. The results of experiments demonstrate that the proposed scheme can effectively protect the security and privacy of IoT data and realize the secure sharing of data.

1. Introduction

At present, the Internet of Things (IoT) technology has been more and more widely used [1], such as in smart medicine and smart cars. IoT devices generate data in real-time and share data by connecting to the Internet. In most traditional IoT systems, data storage and computing adopt a centralized architecture, but the massive, heterogeneous data characteristics and multidimensional, real-time service requests pose great challenges to the centralized data management architecture [2]. Centralized storage and computing architectures are vulnerable to various network attacks, such as data tampering attack, single-point attacks, and distributed denial-of-service attacks [3]. IoT data is a true description of the physical world and often involves personal privacy. If these data are maliciously tampered, forged, or illegally accessed, it may cause disastrous consequences.

A blockchain is a decentralized, tamper-proof, traceable, and multiparty distributed database that integrates a P2P protocol, asymmetric encryption technology, a consensus mechanism, and other technologies [4]. The data on a blockchain are transparent to each participant, and as long

as an honest node holds most of the CPU computing power, the system is safe and reliable. Therefore, the security mechanism of blockchain can effectively solve the problem of data authenticity in the application of IoT, and to a certain extent, it can guarantee the confidentiality and availability of industrial IoT data [5]. Smart contracts—self-executing scripts that reside on the blockchain—allow for distributed, heavily automated workflows. This should make blockchains very suitable for combination with IoT applications [6]. Ge et al. [7] combined blockchain technology with IoT and proposed a decentralized security mechanism based on blockchain technology; the key data generated by IoT devices are stored in the blockchain, and the privacy and security problems that may be encountered by the centralized IoT are solved. Li et al. [8] applied blockchain technology to the medical field and proposed a data protection system (DPS) architecture that stores important medical records on the blockchain to achieve the purposes of data being tamper-proof and traceable.

On the contrary, a centralized storage architecture results in data producers having no control over the data. For instance, large websites collect not only personal

information, such as users' hobbies and browsing habits, but also personal privacy data, which introduces potential threats to personal privacy security. Therefore, it is necessary to strengthen the data owner's control over the data and decide whether other users have the right to access the data [9, 10]. In 2005, attribute-based encryption (ABE) was first proposed by Waters to support the integration of data privacy protection and access control [11]. In the ABE mechanism, the user's private key and ciphertext are associated with a group of attributes, and the user can successfully decrypt the plaintext only when the number of attribute intersections between the attribute set associated with the user's private key and that associated with the ciphertext reaches the threshold value set by the system.

However, there remain many defects and vulnerabilities in blockchain technology, among which storage is one of the most notable [12]. For instance, the Bitcoin system produced a total of 546,349 blocks by October 18, 2018, and the size of a block is generally 996.2 kB, and the size of the entire blockchain is 186.9 GB [13]. With the passage of time, the storage space required by blockchain systems will become increasingly larger. If all the data generated by IoT devices are stored on the blockchain, the size of the entire blockchain will be very large. However, IoT devices cannot store such a large blockchain, so storage optimization is an important challenge that must be overcome. The interplanetary file system (IPFS) is a point-to-point distributed file system [14] that aims to connect all computing devices with the same file system and can be used as a storage scheme for blockchains. Xu et al. [15] proposed a decentralized social network system based on Ethereum and IPFS, which uses IPFS to save large amounts of file data to reduce the Ethereum storage pressure. Klems et al. [16] proposed a decentralized service market system based on blockchain technology and IPFS. In this system, large amounts of data and service metadata are stored in the IPFS network, while only the hash values of these data are stored in the blockchain.

This paper proposes an IoT data access control scheme that combines an ABE algorithm and blockchain technology. In view of the privacy and security problems of IoT data, a symmetric encryption algorithm is combined with an ABE algorithm. Only users who meet the access control policy can successfully access the data and achieve fine-grained access control. To solve the problem of the low storage capacity of blockchain network nodes, the "off-chain" storage mode is adopted. The data are encrypted by a symmetric encryption algorithm, and the ciphertext of the data is stored in the IPFS network. The address hash value of encrypted data stored on IPFS is saved on the blockchain. In this way, the blockchain network is associated with IPFS, which reduces the burden of blockchain storage. Moreover, a consortium blockchain system was constructed based on the Hyperledger Fabric framework, and experiments were carried out to prove the feasibility of the scheme.

At present, there have been many access control schemes based on ciphertext policy attribute-based encryption (CP-ABE) [17] and blockchain, but these schemes often focus on the granularity of access control, but pay less attention to the efficiency of data sharing. Some researchers realize that

blockchain cannot store a large amount of data and adopt cloud storage, but the efficiency of cloud storage is low. The access control scheme proposed in this paper not only uses CP-ABE and blockchain to realize access control but also ensures the simplicity of the control process. Therefore, the scheme realizes the balance between fine-grained access control and efficient file sharing.

The main contributions of our work are as follows:

- (1) The scheme proposed in this paper can ensure the secure and efficient sharing of IoT data, provide fine-grained access to data, simple control process, and enhance the scalability of blockchain system storage. The scheme realizes the balance between fine-grained access control and efficient file sharing.
- (2) The blockchain is used to store the address hash value of a file on IPFS, data hash value, access control policy, timestamp, and other information. This scheme not only ensures that the data cannot be tampered with but also ensures that the access control strategy cannot be tampered with.

2. Related Work

There have been many studies on the application of blockchain technology to data sharing and access control of IoT. For example, Ge et al. and Li et al. [8] used blockchain technology to store important data generated by IoT devices, solve the privacy and security problems of centralized storage architectures, and realize the secure sharing of data. However, these strategies are weak in terms of data access control and do not optimize the blockchain storage; thus, they easily encounter the blockchain storage bottleneck.

2.1. Access Control and Blockchain. Access control technology is widely used in all types of information systems to control the access rights of users and avoid illegal access to data. However, the traditional access control scheme mainly relies on a third-party trusted server, which is characterized by problems such as a single point of failure and low efficiency.

Many researchers have proposed attribute-based data access control methods [9, 10]. In order to ensure the confidentiality of data, an application level fine-grained data access is designed by using the attribute-based encryption method [9]. Fan et al. proposed a fine-grained access-control scheme based on CP-ABE and Trusted Execution Environment (TEE) [10]. TEE is employed as a trusted computing environment to protect encrypted data [9, 10]. This paper uses blockchain technology to ensure the integrity of encrypted data. In the work by Jemel and Serhrouchni [18], a dynamic access control strategy based on blockchain technology and ciphertext policy attribute-based encryption (CP-ABE) was proposed. The time attribute is introduced to realize the dynamic access of data, and only a user whose attribute meets the access control policy within the specified time can access the data.

Blockchain technology is a new technology that can effectively solve the problems existing in traditional access

control. Many researchers have explored the combination of a blockchain and IoT access control technology. For instance, Novo [19] proposed an extensible access management architecture for IoT based on blockchain technology, which defines and executes access control rules via smart contracts. However, this scheme is based on a private blockchain network and sacrifices decentralization to improve performance.

Li et al. [20] proposed an authentication and security scheme for IoT based on blockchain technology, in which the hash values of important data are stored in the blockchain and the unique ID of the device is recorded for authentication, which can effectively avoid single-point-of-failure attacks on the certificate authority (CA) server. However, this scheme ignores the storage bottleneck of blockchains. Ding et al. [21] proposed an attribute-based access control framework, which uses a set of attributes to describe devices and records the distribution of attributes on the chain to avoid a single point of failure and data tampering. Moreover, this scheme uses signature technology and hash operations to simplify the access control protocol.

To a certain extent, while the methods proposed in these studies can achieve data access control, even fine-grained access control, there remain blockchain storage bottlenecks.

2.2. Blockchain and Storage Optimization. Compared with the traditional centralized storage architecture, blockchain technology is characterized by decentralization and non-tamperability, which can effectively guarantee the security of data and improve the scalability of the system. However, the application of blockchain technology to a storage system also introduces new challenges, such as the increase of the storage space overhead.

To solve the blockchain storage problem, many investigations have been carried out. In the work by Zhang and Wang [12], a blockchain fragmentation storage model based on threshold secret sharing was proposed. By improving the Shamir threshold, instead of storing complete transaction data in each node, the data on the chain are segmented and stored, which can effectively reduce the storage capacity of each node. Dai et al. [22] proposed a blockchain storage architecture called NC-DS, which uses network coding to encode data and save the storage space of the blockchain.

In these previous studies, data were dealt with directly. The method proposed by Zhang and Wang [12] stores the data in pieces, while the method proposed by Dai et al. [22] encodes the data. Both schemes can reduce the storage cost of blockchains, but they are both improved on the basis of storing a complete ledger, which saves little space and weakens access control. Cheng et al. [23] proposed a data management scheme for IoT based on blockchain technology and edge computing, which uses the Advanced Encryption Standard (AES) encryption algorithm to protect data security and personal privacy, stores the hash values and some important files on the chain, and stores the encrypted data on the edge server by using a distributed algorithm (Kademlia) to solve the storage bottleneck problem of the blockchain system.

Some data sharing schemes combine the CP-ABE algorithm with blockchain and cloud storage. Wang and Song [24] used attribute-based encryption (ABE) and identity-based encryption (IBE) to encrypt medical data and used identity-based signature (IBS) to implement digital signatures. Wang et al. proposed a personal health records sharing scheme based on blockchain [25]. Wang's scheme used searchable symmetric encryption and attribute-based encryption techniques to achieve fine-grained access control. Wang's scheme allowed patients to distribute attribute private key for users. Both [24, 25] uses the cloud to store medical data, but the security of data privacy depends on cloud service providers.

However, in each data access process, data consumers must communicate with data owners to obtain access rights, which takes a considerable amount of time. In the scheme proposed in this paper, the encrypted data are stored in an IPFS distributed network, and the CP-ABE algorithm is used to achieve fine-grained data access control. The blockchain only stores the hash values of data, the content hash values generated by IPFS, the access control policy, timestamps, and other metadata information, which greatly reduces the storage overhead.

3. Data Access Control Scheme

3.1. Scheme Architecture. The architecture of the data access control scheme that combines ABE and blockchain technology consists of five layers, namely, the consumption layer, interaction layer, access control layer, data layer, and IoT devices layer, as shown in Figure 1. The consumption layer contains all kinds of data consumers, software, or hardware. The interaction layer provides good access protocols and services for data consumers. The access control layer is the concrete realization of business logic, including the realization of smart contracts and the data control access algorithm. Finally, the data layer consists of the blockchain network and IPFS distributed network. The bottom layer is IoT devices layer. IoT devices generate all kinds of data.

The blockchain network is used to store the hash values of data, the content hash values generated by IPFS, the access control policy, timestamps, and other information. Data consumers must meet the stipulations of the access control policy on the blockchain to access the data. After successful access, the consistency and integrity of the data can also be verified through the blockchain. The IPFS distributed network is composed of several server devices with good performance. The data generated by IoT devices will be encrypted and stored in IPFS, and the hash content values generated by IPFS will be stored on the blockchain. In this way, the blockchain network is associated with IPFS, thereby reducing the burden of blockchain storage. The third-party authorization server mainly generates and transmits the public key (PK) and master key (MK) generated by the CP-ABE initialization algorithm and the private key (SK) generated by the CP-ABE key generation algorithm. In this architecture, users do not need to join the blockchain network, let alone consider the specific implementation details of the whole access control scheme. IoT devices only

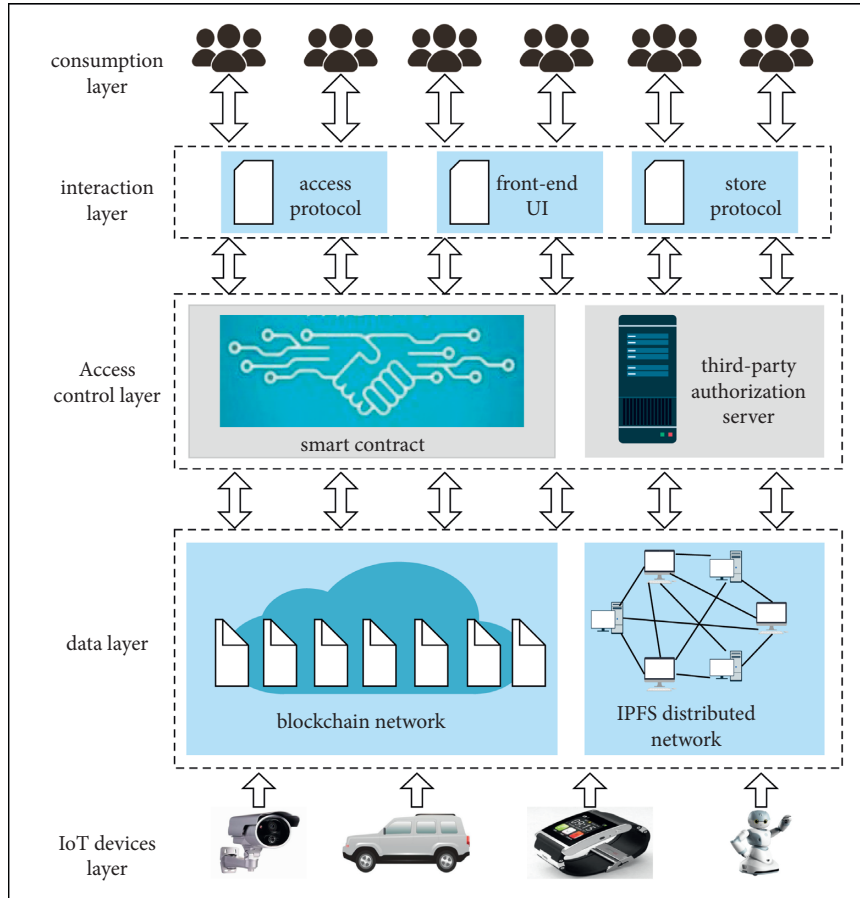


FIGURE 1: Architectural diagram of the scheme.

produce data or files, and data consumers only need to access data or files. In this scheme, the owner of the IoT device is called the data owner.

3.2. Blockchain Design. A blockchain is composed of many blocks, each of which contains a block header and block body. The structure of a blockchain is presented in Figure 2. The block header stores the version number (VersionId), the hash value of the previous block (PreBlock Hash), the Merkel root, and the timestamp (TimeStamp). PreBlock Hash links isolated blocks into chains, TimeStamp refers to the generation time of the block, and the Merkle root is the hash of multiple transaction data. The tampering of transaction data will lead to the inconsistency of the Merkle root, which can be used for transaction integrity verification. The block body primarily stores the hash values of the source data (hashfile), the hash values generated by IPFS (hashipfs), and the access control strategy (policy). Among them, hashfile is generated by the SHA256 algorithm and occupies 32 bytes, hashipfs is the hash value returned after uploading the file to the IPFS network and also occupies 32 bytes, and policy is the access control strategy of the data owner; different owners have different policies, and the upper limit is 1000 bytes.

3.3. Attribute-Based Encryption Mechanism. In the ABE mechanism, the sender uses a set of attributes W to encrypt the message, and the receiver uses a set of attributes W' to describe the identity corresponding to the private key. Only when the intersection number of W' and W exceeds the threshold value t set by the system can the message receiver decrypt the ciphertext. However, this mechanism is limited by access control structures that can only support the threshold policy.

To solve this problem, Goyal et al. [26] proposed a key policy attribute-based encryption (KP-ABE) scheme, which supports fine-grained data access control. In this scheme, the ciphertext is associated with the attribute set of the system, and the key is associated with the access control structure. Only when the attribute set of the user satisfies the access control policy can the decryption be performed. Additionally, Bethencourt et al. proposed CP-ABE [17], which is different from KP-ABE. The ciphertext of CP-ABE is associated with access control, and its key is associated with the attribute set. Only when the user's attribute set satisfies this access control structure can it be decrypted. Moreover, the access control authority of CP-ABE is controlled by the message sender. Therefore, the CP-ABE encryption scheme was adopted in the present work to ensure the data owner's control over the data and realize the fine-grained access control of the data. The

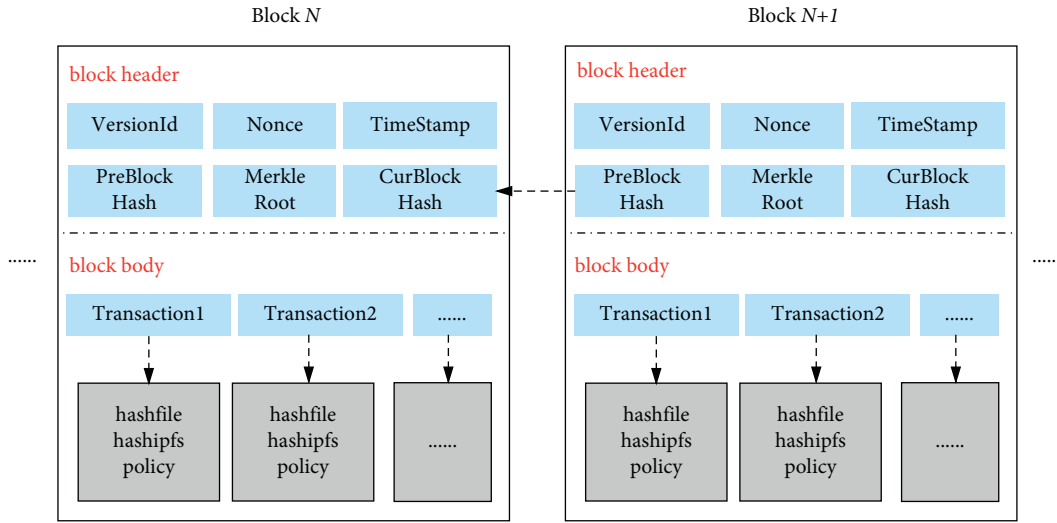


FIGURE 2: Blockchain structure.

algorithm flow of CP-ABE is presented in Figure 3 and is composed of four polynomial algorithms.

- (1) The initialization phase: the trusted key distribution center executes the random initialization algorithm, as shown in equation (1); the input is the security parameter r , and the output includes the public key (PK) and master key (MK):

$$(PK, MK) = \text{Setup}(r). \quad (1)$$

- (2) The key generation stage: the trusted key distribution center executes the key generation algorithm, as shown in equation (2); the inputs are PK and MK generated by equation (1) and the user-defined attribute set A , and the output includes a private key (SK):

$$SK = \text{KeyGen}(PK, MK, A). \quad (2)$$

- (3) The data encryption stage: the data owner executes the encryption algorithm, as shown in equation (3); the inputs include PK, the message m to be encrypted, and the access structure T , and the output is the ciphertext c :

$$c = \text{Encrypt}(PK, m, T). \quad (3)$$

- (4) The data decryption stage: the data requester executes the decryption algorithm, as shown in equation (4); the inputs include PK, SK, and c , and the output is the plaintext message m :

$$m = \text{Decrypt}(PK, c, SK). \quad (4)$$

Table 1 presents the description of the symbols used in this scheme.

3.4. IPFS Distributed Storage. IPFS combines the distributed hash table (DHT), incentive block exchange, self-authentication namespace, and other technologies. Moreover, the

data of IPFS are distributed on different devices, and there exist multiple backups to avoid a single point of failure. Different from the existing web system, in which resources are accessed through URLs, IPFS allows for the retrieval of files by obtaining a unique hash value from the file content. Therefore, once the content of the file changes, the address of the file will change, thereby achieving tamper-proof data. With the passage of time, the storage space required by the blockchain will become increasingly larger. In the proposed method, the ciphertext of the file is stored in the IPFS network, which can alleviate the rapid expansion of the blockchain caused by too much data.

3.5. Data Storage and Access

3.5.1. Data Storage. As shown in Figure 4, the data storage procedure of this scheme includes five participants, namely, the data owner, system server side, IPFS distributed network, blockchain network, and third-party authorization server. The detailed process is as follows.

- (1) The data owner (Owner) selects the file to be stored and sets the access control policy of the file (policy). The data consumer (Consumer) can successfully access the file only if the set of attributes of the data consumer meets the stipulations of the access control policy:

$$\text{policy} \leftarrow (\text{Owner}, \text{file}). \quad (5)$$

- (2) The data owner has a unique AES key (key). If the data owner has not generated the key before, the server side calls the AES key generation algorithm to generate the key, as shown in equation (6). Then, the server side calls the AES encryption algorithm to encrypt the file and obtain the encrypted file (encfile), as shown in equation (7). Finally, the server side calls the IPFS storage algorithm to store the encrypted file in the IPFS distributed network, as

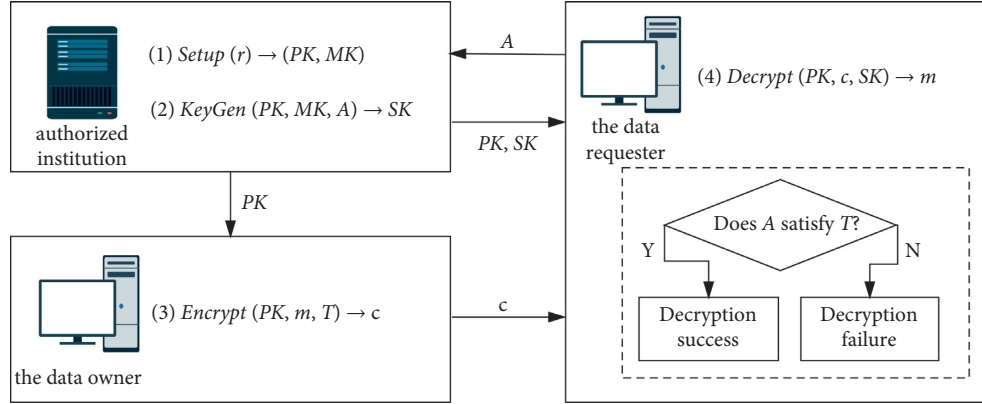


FIGURE 3: CP-ABE algorithm flow.

TABLE 1: Symbol description.

Symbol	Description
Owner	Data owner
Consumer	Data consumer
file	File
policy	Access control policy set by data owner
key	Key generated by AES algorithm
encfile	Ciphertext of a file
hashipfs	The address of the ciphertext of the file on IPFS
hashfile	The hash value of the file
PK	Public key generated by CP-ABE algorithm
MK	Master key generated by CP-ABE algorithm
SK	Private key generated by CP-ABE algorithm
A	Attribute set of data consumers
enckey	Ciphertext of key
deckey	Decrypted key
decfile	Decrypted file
dechash	Hash value of encrypted file after decryption

shown in equation (8), and records the hash value (hashipfs) used to access the ciphertext:

$$\text{key} = \text{AES.Gen}(\text{Owner}), \quad (6)$$

$$\text{encfile} = \text{AES.Enc}(\text{key}, \text{file}), \quad (7)$$

$$\text{hashipfs} = \text{IPFS.Store}(\text{encfile}). \quad (8)$$

- (3) The server side calls the SHA256 algorithm to hash the file to get the file hash value (hashfile), as shown in equation (9). Then, the previously generated hashfile, hashipfs, and policy are sent to the blockchain network:

$$\text{hashfile} = \text{SHA256.Hash}(\text{file}). \quad (9)$$

- (4) The blockchain network receives the data storage request and triggers the storage smart contract (StoreCont) to store the hashfile, hashipfs, and policy on the blockchain:

$$\text{StoreCont}(\text{hashfile}, \text{hashipfs}, \text{policy}). \quad (10)$$

- (5) The server side requests the public key (PK) from the third-party authorization server for the later encryption of the file.

- (6) The data owner has a unique public key (PK) and a unique master key (MK). If the data owner has not generated the public key and master key before, the third-party authorization server will call the initialization algorithm (Setup) of the CP-ABE algorithm to generate and store PK and MK, as shown in equation (11), and will then send PK to the server side:

$$\text{PK}, \text{MK} = \text{CPABE.Setup}(r). \quad (11)$$

- (7) The server side calls the encryption algorithm (Encrypt) of the CP-ABE algorithm, takes the policy and PK as the input of the encryption algorithm, encrypts the *key* to get the ciphertext of the *key* (enckey), and stores it, as shown in the following equation:

$$\text{enckey} = \text{CPABE.Encrypt}(\text{PK}, \text{key}, \text{policy}). \quad (12)$$

The data storage algorithm is described in Algorithm 1.

3.5.2. Data Access. As shown in Figure 5, the data access process of this scheme includes five participants, namely, the data consumer, system server side, IPFS distributed network, blockchain network, and third-party authorization server. The detailed process is as follows.

- (1) The data consumer (Consumer) sends out a request to access the file, which contains the attribute set *A* of the data consumer.
- (2) After receiving the request from the data consumer, the server side requests the hash value of the file (hashfile) and the hash value used to access the ciphertext of the file to the IPFS network (hashipfs) from the blockchain network.
- (3) The blockchain network receives the data access request, triggers a query smart contract

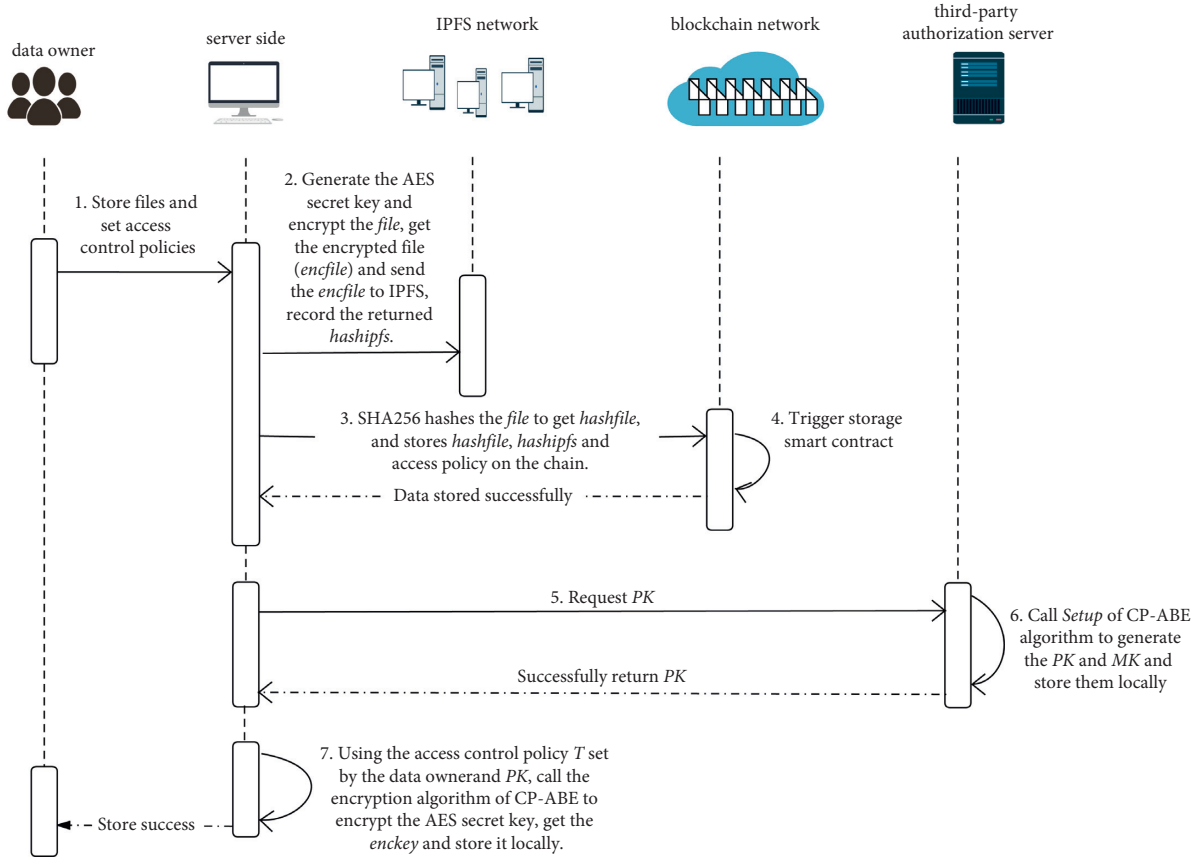


FIGURE 4: Data storage flow.

```

Begin
  policy ← (Owner, file)
  if key not exist then
    key = AES. Gen(Owner)
  end if
  encfile = AES. Enc(key, file)
  hashipfs = IPFS. Store(encfile)
  hashfile = SHA256. Hash(file)
  StoreCont(hashfile, hashipfs, policy)
  if PK, MK not exist then
    PK, MK = CPABE. Setup(r)
  end if
  enckey = CPABE. Encrypt(PK, key, policy)
End
    
```

ALGORITHM 1: Data Storage.

(QueryCont), gets the hashfile and hashipfs, and sends them to the server side:

$$\text{hashfile, hashipfs} \leftarrow \text{QueryCont}(\text{file}). \quad (13)$$

(4) The server side requests the public key PK and private key SK from the third-party authorization server to later decrypt the file.

(5) According to PK, MK, and attribute set A of the data consumer, the third-party authorization server executes the key generation algorithm (KeyGen) of the CP-ABE algorithm to generate the private key (SK), as shown in equation (14), and sends PK and SK to the server side:

$$\text{SK} = \text{CPABE.KeyGen}(\text{PK}, \text{MK}, A). \quad (14)$$

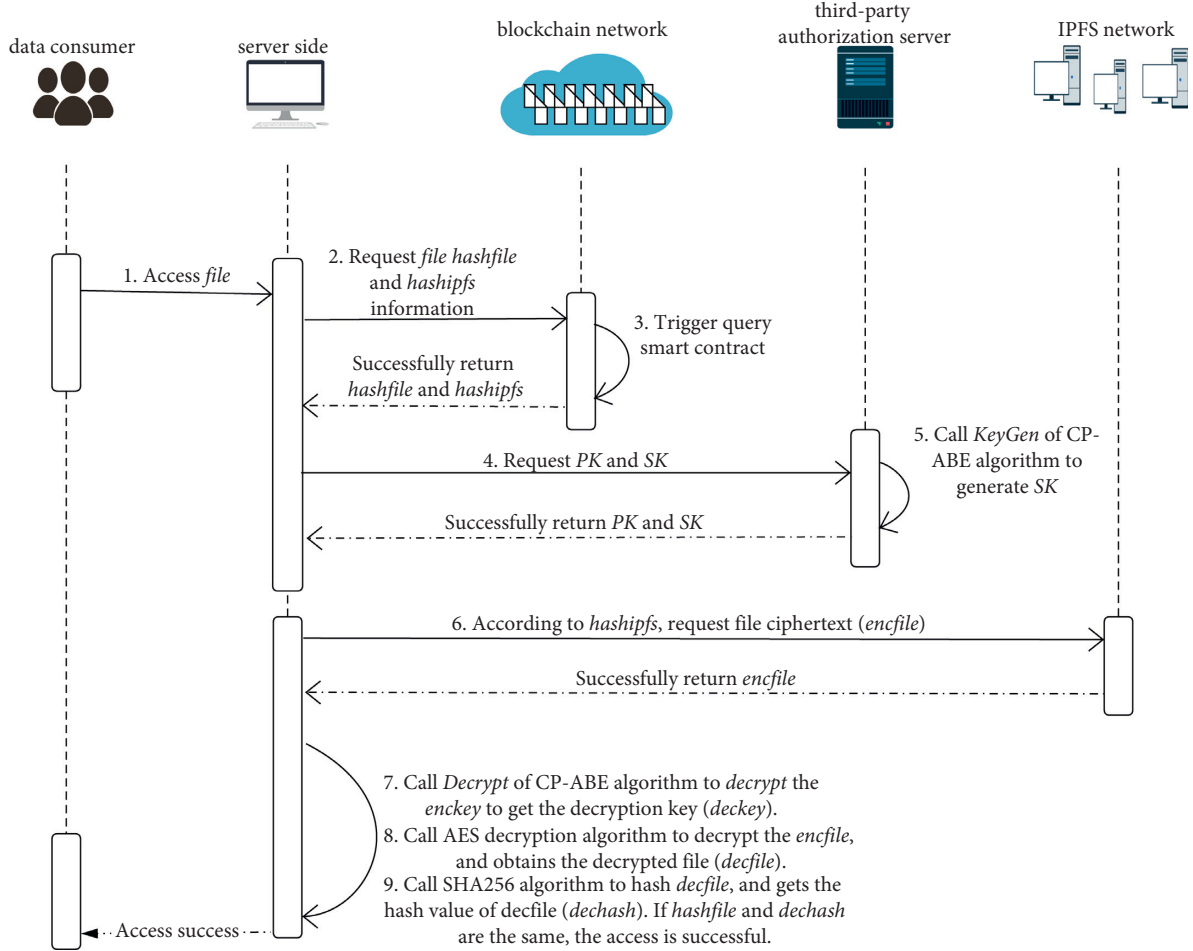


FIGURE 5: Data access flow.

(6) According to the hashipfs obtained from the chain, the server side calls the IPFS query algorithm to obtain the ciphertext of the file (*encfile*) from the IPFS network, as shown in the following equation:

$$\text{encfile} = \text{IPFS.Query}(\text{file}, \text{hashipfs}). \quad (15)$$

(7) The server side obtains the AES key ciphertext (*enckey*) that encrypts the file locally and calls the decryption algorithm (*Decrypt*) of the CP-ABE algorithm to decrypt *enckey* and obtain the decryption key (*deckey*), as shown in the following equation:

$$\begin{aligned} \text{deckey} &= \text{CPABE.Decrypt}(\text{PK}, \text{enckey}, \text{SK}), \\ \text{enckey} &= \text{Get}(\text{file}). \end{aligned} \quad (16)$$

(8) According to *deckey*, the server side calls the AES decryption algorithm to decrypt *encfile* and obtains the decrypted file (*decfile*), as shown in the following equation:

$$\text{decfile} = \text{AES.Dec}(\text{deckey}, \text{encfile}). \quad (17)$$

(9) The server side calls the SHA256 algorithm to hash *decfile* and gets the hash value of *decfile* (*dechash*),

as shown in equation (18). If *hashfile* and *dechash* are the same, the access is successful:

$$\text{dechash} = \text{SHA256.Hash}(\text{decfile}). \quad (18)$$

The data access algorithm is shown in Algorithm 2.

4. Security Model

4.1. Tamper-Proof. In environmental monitoring, equipment detection, and other scenarios, a large number of IoT devices will be used. For example, a variety of sensors are used to detect the quality of large-scale equipment. Dishonest device manufacturers can make their unqualified products pass the detection by tampering the detection data, which is the threat of data tampering in the IoT.

The data block on the blockchain contains a timestamp and the hash value of the previous block. The blocks linked in the chronological order ensure that the transaction messages cannot be modified arbitrarily unless 51% of the nodes in the whole network are tampered, which is almost impossible in a large blockchain network. Therefore, this scheme can avoid the threat of data tampering.

```

Begin
  hashfile, hashipfs ← QueryCont(file)
  if SK not exist then
    SK = CPABE.KeyGen (PK, MK, A)
  end if
  encfile = IPFS.Query (file, hashipfs)
  enckey = Get (Owner)
  deckey = CPABE.Decrypt (PK, enckey, SK)
  decfile = AES.Dec (deckey, encfile)
  dechash = SHA256.Hash (decfile)
  If dechash equal hashfile then
    return success
  Else
    return failure
End

```

ALGORITHM 2: Data access.

4.2. Confidentiality and Integrity. Some IoT applications will store users' private information. For example, the medical IoT system will store patients' condition, and the traffic monitoring system will store vehicles' video. Hackers will attack this kind of IoT applications to illegally obtain the private data, resulting in the threat of illegal access and data leakage.

The data access control scheme proposed in this paper ensures the confidentiality and integrity of system data. Via the combination of the AES symmetric encryption algorithm and CP-ABE algorithm, the data owner formulates the control access policy. Only when the attribute set owned by the data visitor meets the stipulations of the policy can the data be accessed. The data access permission of this scheme is determined by the data owner, and the fine-grained access control and sharing of data are realized by setting different access policies. During the process of data storage and access, the data are transmitted in an encrypted state, and visitors outside the data access authority cannot obtain the real data, which ensures data confidentiality. Moreover, the hash algorithm is used to hash the original data, and the hash values are stored on the blockchain. The data visitor calculates the hash value after receiving the data and compares it with the hash value on the chain; if the data has been tampered with, the two hash values will be inconsistent, which allows for the verification of the integrity of the data.

4.3. Scalability and Reliability. IoT devices usually run for a long time, and many devices work in the field or unmanned environment for a long time, such as water quality monitoring sensors. IoT data servers are prone to various failures due to long hours of work or are subject to DOS attacks, resulting in data loss or system crashes. This is called the single-point-of-failure threat.

The data access control scheme proposed in this paper ensures the reliability and robustness of the system. Traditional centralized storage and computing architectures are prone to a single point of failure. If a device is damaged, all devices may be affected, thereby causing irreparable losses.

In this work, decentralized blockchain technology and IPFS distributed technology were combined to store the ciphertext of data on the P2P distributed IPFS network, which alleviates the storage pressure of the blockchain and enhances the scalability of the blockchain system. In the IPFS network, large-capacity ciphertext data will be divided into multiple 256 kB packets and stored on different nodes with multiple copies. As long as a node on the IPFS network has a copy of the data content, the data visitors can access it according to the retrieval hash value of the corresponding data. In addition, once the content of the data changes, the address to retrieve the data will change to achieve the antitampering purpose.

5. Experiment and Analysis

5.1. Experimental Environment. To verify the feasibility of the proposed data access control scheme that combines ABE and blockchain technology, a prototype system was constructed for verification. The configuration of the prototype system environment mainly included a blockchain network and IPFS distributed network. This system used the Hyperledger Fabric framework and Docker container technology to build the blockchain network. Due to the limited number of servers, the CentOS 7 virtual machine was used as the running environment of Fabric. This blockchain network consisted of an Orderer node and four Peer nodes. The details are reported in Table 2. The system used go-ipfs to build a private IPFS distributed network, which was composed of seven devices in the same LAN. In addition, a local server was compared with the IPFS distributed network. The specific configuration information of the equipment is presented in Table 3.

5.2. Operation Process of the Smart Contract. In the Hyperledger Fabric framework, a smart contract is called a Chaincode, which is a piece of code written in the Go programming language. It runs in an independent and secure Docker container and initializes and manages the

TABLE 2: Fabric network configuration.

Node name	Docker image file	Simulation object
orderer.example.com	Hyperledger/Fabric-orderer	Orderer
peer0.org1.example.com	Hyperledger/Fabric-peer	Peer1 node
peer1.org1.example.com	Hyperledger/Fabric-peer	Peer2 node
peer0.org2.example.com	Hyperledger/Fabric-peer	Peer3 node
peer1.org2.example.com	Hyperledger/Fabric-peer	Peer4 node
ca.example.com	Hyperledger/Fabric-ca	CA server
Couchdb	Hyperledger/Fabric-couchdb	Couchdb
Cli	Hyperledger/Fabric-tools	Client

TABLE 3: Performance parameters of experimental equipment.

Equipment name	Configuration
Local server	Ubuntu Linux release 18.04
Virtual machine	CentOS Linux release 7.8.2003(Core)
IPFS_1	Windows10; CPU i7-8550U; RAM 16.0 GB
IPFS_2	Windows10; CPU r7-4800U; RAM 16.0 GB
IPFS_3	Windows10; CPU i5-9300H; RAM 16.0 GB
IPFS_4	Windows10; CPU i7-8650U; RAM 16.0 GB
IPFS_5	Windows10; CPU i5-6200U; RAM 4.0 GB
IPFS_6	Windows10; CPU i5-4210H; RAM 8.0 GB
IPFS_7	Windows10; CPU i5-9300H; RAM 8.0 GB

ledger state via the transaction submitted by the application. A smart contract works automatically. Once the smart contract is verified, the verified result set is sent to the Orderer nodes, and the changes in the running results will be shared or synchronized to all Peer nodes in the Fabric network. Hyperledger Fabric provides four basic commands to manage the life cycle of smart contracts, namely, package, install, instantiate, and upgrade.

In this experiment, there were two main smart contracts, namely, storage and query smart contracts. The storage smart contract primarily stores the file hash result encrypted by SHA256 (hashfile), the file hash value generated by IPFS (hashipfs), and the access control policy on the blockchain. The query smart contract mainly extracts metadata, such as hashfile and hashipfs. The specific steps are as follows:

- (1) Start the Fabric network: use the Docker-Compose tool to read the configuration file to start the Orderer, Peer1, Peer2, Peer3, Peer4, CA, and Couchdb services.
- (2) Create a channel: based on the Orderer node, establish a channel named mychannel, and add four Peer nodes to the channel.
- (3) Install and initialize the smart contracts: use the install command to install the storage and query smart contracts for Peer nodes, and set the endorsement policy. In our experiment, there are two organizations, and a transaction is valid only when the members of the two organizations jointly endorse.
- (4) Use smart contracts: use the fabric-sdk-java tool to call the smart contracts to meet the needs of storage and query.

5.3. Experimental Results

5.3.1. Access Control Structure. The efficiency of the access control scheme is limited by the access control structure and file storage mode. The access control structure in this study consists of the process of attribute-based encryption and file encryption. Therefore, comparative experiments were carried out under different numbers of attributes. All the files in the experiment are numerical data files, which are randomly generated by computer software. The numerical value has no effect on the experimental results. The reported experimental results are the average results of 50 experimental runs.

This experiment was conducted to investigate the time taken by the access control scheme to encrypt and decrypt files under different numbers of attributes. The fixed file size in this experiment was 500 kB. The comparison method was C-AB/IB-ES which was proposed by Hao Wang and Song [24]. C-AB/IB-ES uses attribute-based encryption (ABE) and identity-based encryption (IBE) to encrypt data and uses identity-based signature (IBS) to implement digital signatures. The access control structure of C-AB/IB-ES is more complex than our scheme. As shown in Figure 6, the experimental results reveal that, with the increase of the number of attributes, the encryption and decryption time of the two algorithms increased linearly, which demonstrates that the computational overhead of these schemes were controllable. The encryption and decryption time of our method is much shorter than that of [24]. This indicates that the more complex the access control structure of the CP-ABE algorithm, the more time it takes for access control. Therefore, while pursuing fine-grained data access control, the fewer the number of attributes, the lesser the computational overhead of the system.

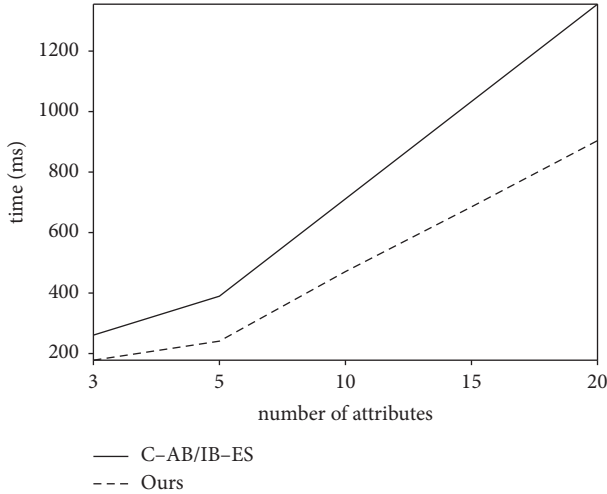


FIGURE 6: Encryption and decryption times with different numbers of attributes.

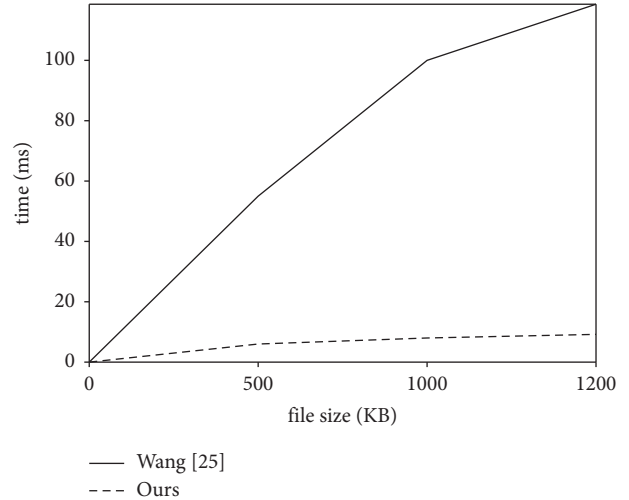


FIGURE 7: Upload times of the two schemes.

5.3.2. *File Storage Mode.* This experiment was a storage performance comparison between different file storage modes. The access control scheme proposed by Wang et al. [25] uses a cloud server to store files. A local LAN server was used for cloud server storage, and the cluster devices of other IPFS distributed networks were also in the same LAN; the detailed configuration is presented in Table 3. This experiment was conducted to test the time taken to upload and download different-sized files under two storage modes. As shown in Figures 7 and 8, the experimental results indicate that, under the condition of the same file size, the upload and download times of the IPFS storage scheme were, respectively, about 8% and 11% those of the cloud storage scheme. Therefore, the adoption of the proposed IPFS storage scheme can improve the computational efficiency and reduce the system overhead.

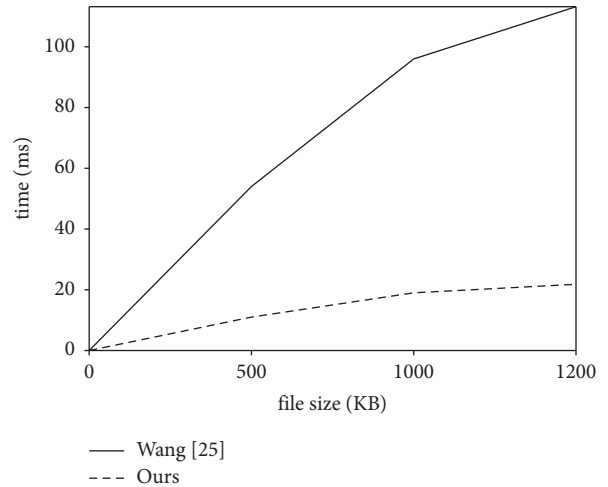


FIGURE 8: Download times of the two schemes.

5.3.3. *Overall Encryption Efficiency.* In this experiment, we compare the overall encryption efficiency. The overall encryption time includes file storage time, time spent writing data to the blockchain, CP-ABE encryption time, and AES encryption time. The number of attributes was unchanged, and the encryption and decryption time of different schemes were calculated by changing the size of the files. As shown in Figure 9, with the increase of the file size, the encryption time of Wang’s increased rapidly. Both Wang’s and our scheme use the CP-ABE and AES algorithm to archive access control, but the two access control processes are different. The overall encryption efficiency of Wang’s was lower than that of our scheme.

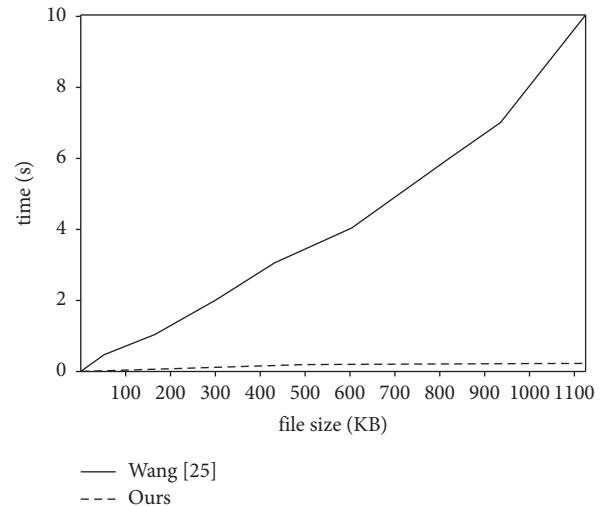


FIGURE 9: Overall encryption efficiency.

5.4. *Security Analysis.* In our scheme, the owner generates and distributes the attribute private key for the user. It not only solves many security risks caused by untrusted attribute authorities in attribute-based encryption schemes but also fine-grained access control is implemented for users without relying on any third party.

TABLE 4: Scheme comparison.

Literature	Time	Confidentiality	Integrity	Scalability	Reliability/robustness	Access control
[18]	2017	Strong	Strong	Weak	Weak	Strong
[8]	2018	Strong	Strong	Weak	Weak	Weak
[19]	2018	Strong	Strong	Weak	Weak	Strong
[20]	2018	Strong	Strong	Weak	Weak	Strong
[22]	2018	Strong	Strong	Strong	Weak	Weak
[12]	2019	Strong	Strong	Strong	Weak	Weak
[21]	2019	Strong	Strong	Weak	Weak	Strong
[7]	2020	Strong	Strong	Weak	Weak	Weak
[23]	2020	Strong	Strong	Strong	Strong	Strong
Our	2021	Strong	Strong	Strong	Strong	Strong

Table 4 presents a comparison between the proposed scheme and methods put forward in other similar studies. Ge et al. [7] and Li et al. [8] only applied blockchain technology to their respective schemes, and stored important data on the blockchain to ensure data security and personal privacy. Novo [19], Li et al. [20], Ding et al. [21], and Jemel and Serhrouchni [18] did not only store important data on the blockchain but also demonstrated strong access control ability. Similar to the work reported in the present article, Jemel introduced CP-ABE encryption technology to achieve the fine-grained access control of blockchain data, and the built-in encryption technology ensures the confidentiality and integrity of the data. However, the storage bottleneck of blockchain systems was not considered, and the scalability is weak.

Zhang and Wang [12], Dai et al. [22], and Cheng et al. [23] considered the storage bottleneck of blockchain systems. Zhang and Dai reduced the storage space of blockchain data and improved the scalability of the blockchain system by compressing the data itself. The method proposed by Cheng et al. [23] was also demonstrated to achieve high performance in many aspects. It uses AES to encrypt data and then stores the hash values on the chain to ensure the confidentiality and integrity of the data. Moreover, it stores encrypted data on edge servers with multiple copies, thereby improving the scalability, reliability, and robustness of the system. However, in this scheme, data consumers need the consent of the data owner to access the data, and the fine-grained access ability is weak. With the increase of the amount of access, the communication overhead increases significantly, and the feasibility of practical application is low. Via comparative analysis, the scheme proposed in the present study demonstrated advantages in ensuring data confidentiality, integrity, and scalability.

6. Conclusion

To achieve secure IoT data sharing and fine-grained access control, this paper proposed a data access control scheme based on the CP-ABE algorithm and blockchain technology. In this scheme, the hash value of the data, the location information of the encrypted data, and the access control strategy are stored on the blockchain. The blockchain ensures the integrity and tamperability of these data, so as to achieve fine-grained access to data efficiently. Our scheme

not only ensures that the data cannot be tampered with but also ensures that the access control strategy cannot be tampered with.

The proposed scheme alleviates the storage pressure and effectively improves the scalability of the blockchain. The results of a contrast experiment with the cloud storage scheme demonstrate that, as compared to the cloud storage scheme, the proposed scheme consumed 8% of the time to store files and 11% of the time to access files. In addition, a fine-grained access control mechanism that combines the symmetric encryption algorithm and CP-ABE was proposed. In this mechanism, the symmetric encryption algorithm is first used to encrypt the data, and the CP-ABE algorithm is then used to encrypt the symmetric encryption key. This mechanism ensures the security of IoT data, enables data owners to control their data visitors in a fine-grained way, and ensures that the data cannot be tampered with.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by National Key R&D Program of China (Grant no. 2020YFB2104700) and National Natural Science Foundation of China (Grant no. 62136006).

References

- [1] Y. Li, Y. Guo, and S. Chen, "A survey on the development and challenges of the Internet of things (IoT) in China," in *Proceedings of the International Symposium in Sensing and Instrumentation in IoT Era*, September 2018.
- [2] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and privacy for cloud-based IoT: challenges," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 26–33, 2017.
- [3] F. Liu, G. Qian, Y. Yarom, F. Mckeen, and R. B. Lee, "CATalyst: defeating last-level cache side channel attacks in cloud computing," in *Proceedings of the IEEE International*

- Symposium on High Performance Computer Architecture*, Barcelona, Spain, March. 2016.
- [4] Q. Shao, C. Jin, Z. Zhang, W. Qian, and A. Zhou, "Blockchain technology: architecture and progress," *Chinese Journal of Computers*, vol. 18, no. 8, p. 2449, 2018.
 - [5] W. Liang, Y. Fan, K. C. Li, D. Zhang, and J. L. Gaudiot, "Secure data storage and recovery in industrial blockchain network environments," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 99, p. 1, 2020.
 - [6] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.
 - [7] C. Ge, Z. Liu, and L. Fang, "A blockchain based decentralized data security mechanism for the Internet of Things," *Journal of Parallel and Distributed Computing*, vol. 141, pp. 1–9, 2020.
 - [8] H. Li, L. Zhu, M. Shen, F. Gao, X. Tao, and S. Liu, "Blockchain-based data preservation system for medical data," *Journal of Medical Systems*, vol. 42, no. 8, p. 141, 2018.
 - [9] Y. Fan, S. Liu, X. Lei, K. C. Li, and G. Tan, "One enhanced secure access scheme for outsourced data," *Information Sciences*, pp. 230–242, 2020.
 - [10] Y. Fan, S. Liu, G. Tan, and F. Qiao, "Fine-grained access control based on trusted execution environment," *Future Generation Computer Systems*, pp. 551–561, 2018.
 - [11] A. Sahai and B. R. Waters, "Fuzzy identity-based encryption," in *Proceedings of the 24th Annual International Conference on Theory and Applications of Cryptographic Techniques*, Berlin, Heidelberg, May 2005.
 - [12] G. Zhang and R. Wang, "Blockchain shard storage model based on threshold secret sharing," *Journal of Computer Applications*, vol. 39, no. 9, pp. 2617–2622, 2019.
 - [13] Blockchain, *The Blockchain Data of Bitcoin [EB/OL]*, Blockchain, Luxembourg, UK, 2018, <https://www.blockchain.com/>.
 - [14] B. Confais, A. Lebre, and B. Parrein, "An object store service for a fog/edge computing infrastructure based on IPFS and scale-out NAS," in *Proceedings of the IEEE International Conference on Fog & Edge Computing*, 2017.
 - [15] Q. Xu, Z. Song, R. Goh, and Y. Li, "Building an Ethereum and IPFS-based decentralized social network system," in *Proceedings of the 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)*, Singapore, December 2018.
 - [16] M. Klems, J. Eberhardt, S. Tai, S. Härtle, S. Buchholz, and A. Tidjani, "Trustless intermediation in blockchain-based decentralized service marketplaces," in *Proceedings of the International Conference on Service-Oriented Computing*, Malaga, Spain, November 2017.
 - [17] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *Proceedings of the IEEE Symposium on Security & Privacy*, 2007.
 - [18] M. Jemel and A. Serhrouchni, "Decentralized access control mechanism with temporal dimension based on blockchain," in *Proceedings of the IEEE International Conference on E-business Engineering*, November 2017.
 - [19] O. Novo, "Blockchain meets IoT: an architecture for scalable access management in IoT," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1184–1195, 2018.
 - [20] D. Li, P. Wei, W. Deng, and F. Gai, "A blockchain-based authentication and security mechanism for IoT," in *Proceedings of the 2018 27th International Conference on Computer Communication and Networks (ICCCN)*, Hangzhou, China, October 2018.
 - [21] S. Ding, J. Cao, C. Li, K. Fan, and H. Li, "A novel attribute-based access control scheme using blockchain for IoT," *IEEE Access*, vol. 7, p. 1, 2019.
 - [22] M. Dai, S. Zhang, H. Wang, and S. Jin, "A low storage room requirement framework for distributed ledger in blockchain," *IEEE Access*, vol. 6, pp. 22970–22975, 2018.
 - [23] G. Cheng, Z. Huang, and S. Deng, "Data management of Internet of things based on blockchain and edge computing," *Chinese Journal on Internet of Things*, vol. 4, no. 2, pp. 2–10, 2020.
 - [24] H. Wang and Y. Song, "Secure cloud-based EHR system using attribute-based cryptosystem and blockchain," *Journal of Medical Systems*, vol. 42, no. 8, p. 152, 2018.
 - [25] S. Wang, D. Zhang, and Y. Zhang, "Blockchain-based personal health records sharing scheme with data integrity verifiable," *IEEE Access*, vol. 7, pp. 102887–102901, 2019.
 - [26] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, CCS 2006, Alexandria, VA, USA, November 2006.

Research Article

Attribution Classification Method of APT Malware in IoT Using Machine Learning Techniques

Shudong Li ¹, Qianqing Zhang ¹, Xiaobo Wu ², Weihong Han ¹ and Zhihong Tian ¹

¹Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China

²School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China

Correspondence should be addressed to Shudong Li; lishudong@gzhu.edu.cn

Received 10 June 2021; Accepted 23 August 2021; Published 7 September 2021

Academic Editor: Shui Yu

Copyright © 2021 Shudong 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 recent years, the popularity of IoT (Internet of Things) applications and services has brought great convenience to people's lives, but ubiquitous IoT has also brought many security problems. Among them, advanced persistent threat (APT) is one of the most representative attacks, and its continuous outbreak has brought unprecedented security challenges for the large-scale deployment of the IoT. However, important research on analyzing the attribution of APT malware samples is still relatively few. Therefore, we propose a classification method for attribution organizations with APT malware in IoT using machine learning. It aims to mark the real attacking organization entities to better identify APT attack activity and protect the security of IoT. This method performs feature representation and feature selection based on APT behavior data obtained from devices in the Internet of Things and selects the features with a high degree of differentiation among organizations. Then, it trains a multiclass model named SMOTE-RF that can better deal with imbalance and multiclassification problems. Our experiments on real dynamic behavior data are combined to verify the effectiveness of the method proposed in this paper for attribution analysis of APT malware samples and achieve good performance. Our method could identify the organization behind complex APT attacks in IoT devices and services.

1. Introduction

As IoT applications and services spread to every corner of our lives, the number of Internet of Things devices is rapidly increasing. However, most of the devices were developed without considering security issues, as well as cannot be updated, which makes cybercriminals vulnerable to attack when they find some errors or security problems. Ubiquitous IoT has brought many security problems [1, 2]. The VPNFilter incident was one of the most serious IoT device attacks of 2018, and the US Department of Justice has since linked the incident to APT28. The incident affected 50W devices in at least 54 countries and regions worldwide, affecting and damaging the ubiquitous IoT environment. In August 2019, Microsoft reported that its Threat Intelligence Center had detected an attack on IoT devices, including VoIP phones, printers, and video decoders. Two of the three devices affected by the attack had factory security settings, while the software in the third was not updated. Microsoft blamed the attack on a Russia-based group

commonly known as APT28. As AI advances, hackers are also using it to launch more sophisticated attacks on computer systems. Among the attacks of the IoT, advanced persistent threat (APT) is one of the most representative attacks, and its continuous outbreak has brought unprecedented security challenges. Therefore, the APT attack has attracted much attention of various researchers and many governments. An APT attack is a form of long-term and persistent network attack by individuals or organizations, who use advanced attack techniques against specific targets. The difference between the APT attack and the traditional network attack is that the APT attack has the characteristics of concealment, pertinence, persistence, and organization [3]. Its attack means are changeable, the attack effect is remarkable, and it is difficult to prevent, such as the famous APT attack, the "Stuxnet" virus [4]. The virus broke out in 2010. The technology of the virus is complex and hidden, which makes the discovery and analysis process take long time. Its infection was targeted mainly at Iran's nuclear facilities, which had a huge impact on Iran's nuclear program. This

incident is considered an organized state act. Moreover, in 2016, hackers launched DDOS attacks by manipulating IoT devices infected with malware known as Mirai. Behind APT attacks, there are usually organizations with government background or intelligence institutional background that provide funding with political or economic purpose [5]; the threat to national and enterprise information security systems is becoming more and more serious, and the number of APT reports is increasing year by year. Security agencies of various countries have disclosed hundreds of APT organizations, commonly active ones being Russia's APT28 and APT29, North Korea's Lazarus, and so on. Attribution analysis of APT samples has always been one of the most important links in the analysis of APT attacks, and it is also a method to detect APT attacks [6]. At present, industrial analysis on the attribution of APT samples mainly relies on the manual analysis by safety experts, which are greatly affected by the expert experience. Besides, it cannot meet the need of a large number of samples, which are low in efficiency and time-consuming. There are relatively few studies on the attribution of attack samples in academia. With the rapid increase in the number of polymorphic viruses and deformed Trojans, malware has become one of the usual methods of APT attacks [7]. FireEye is proposed to perform APT organization clustering based on the similarity of malicious code samples [8]. The characteristics of malware are mainly divided into static features (binary file characteristics, disassembly features, etc.) [9] and dynamic features (execution behavior features, etc.) [10]. Static features are generally disassembled, etc. It is usually difficult to extract effective features due to polymorphism, deformation, and shelling. Dynamic features are generally obtained by monitoring the behavior of the program during runtime, which is not affected by confusion technology [11–14].

APTs of the same organization have certain similarities in their behavior, to realize the automatic classification of APT malware samples, that is, to classify and identify the samples of the same organization. Based on the behavioral data of APT attack malware obtained from the Internet of Things devices, this paper proposes a classification method of APT attack organization based on machine learning. The main contributions of this paper are as follows:

- (i) We propose an APT organization classification method based on machine learning and malware. The method that aims to effectively identify APT attack activity has been verified by experiments in that it has stable performance and high efficiency, which can mark real attack organization entities to protect the security of the Internet of Things.
- (ii) We carry out feature representation and selection filtering in that to get the features with a higher distinguishing degree in different organizations based on the acquired behavior data of malware, which reduces the feature dimension and improves the calculation speed.

- (iii) Due to the imbalance of the APT organization data set, we designed the SMOTE-RF model to solve this multiclassification problem.

1.1. Related Works. The APT attack is a complex network attack with a very obvious purpose. It attacks the target network step by step through multiple stages and maintains long-term access to the target [15]. With the aid of APT malware, an attacker can remotely control the infected machine and steal sensitive information. APT malware, such as Trojan horses or backdoors, is a firewall dedicated to antivirus software and target networks. It is not only used to remotely control the infected machine in APT attacks but also used to steal sensitive information from the infected host for a long period [16].

Currently, commonly used detection methods related to APT are mainly researched from the aspects of malicious code detection, attack detection, and network traffic detection. Abomhara and Kien [17] proposed threats and attacks faced by the IoT infrastructure. In addition to analyzing and describing intruders and attacks faced by IoT devices and services, they also tried to classify threat types. Sung et al. [18] proposed a new and practical security architecture model that protects each layer and interface. The protection includes data protection, access control, prevention of threats, and protection against network attacks. By mapping these analytical protection controls to the risks in each department and its resources, companies can apply robust multilayer defenses against any attack, including advanced persistent threats. Lee and Lewis [19] proved that it is possible to use undirected graphs to associate attacks based on shared targets between different attacks. Based on this information, a map of APT activities can be built and clusters that may represent the activities of a single team of malware writers can be identified. In addition, there are some other detection methods [20, 21].

For the detection of malicious software, malware can be identified by intelligent analysis of the characteristics of malicious samples [22]. As malware has become a necessary strategic tool for APT attacks, the characteristics of malware can also be used as the characteristics of APT attack organizations [23]. The methods of extracting malware features mainly include static feature extraction and dynamic feature extraction [24]. The static feature extraction method uses file structure analysis, decompilation, disassembly, control flow, and data flow analysis techniques to extract static features such as component instructions, control flow, and function call sequence of the program without running the program. For example, Qiao et al. [25] proposed an automatic malware homology identification method based on API calls. This method obtains its API set through static analysis of malicious samples, then uses the Jaccard similarity coefficient to calculate the homology degree of different malware types based on the six calling behaviors defined by

programming habits, establishes a threshold to compare with the homology degree through experience, and draws a conclusion about whether the samples are similar or not. This method can be used to determine the degree of homology between APT samples and determine the organization of the samples. The dynamic feature extraction is used to monitor the behavior of the program when it is running and then extract the dynamic behavior characteristics of the code such as API operations, file system operations, function access, and system calls. For example, Chen et al. [5] proposed a new genetic model combined with a knowledge map of malware behavior. Their method is to build a genetic model based on the content of the node, extract the gene sequence of all malware belonging to each APT organization, and then calculate the similarity between the malware and the gene library and compare it with a preset threshold like which APT organization of the malware belongs to.

In the industry, APT organizational identification is more inclined to analyze the correlation between malicious code structure and its attack chain. For example, FireEye Lab [26] analyzed 11 APT attacks in 2013 and found the malicious code used in the attack based on the same code segment, timestamp, digital certificate, etc. Based on these collected characteristics, the correlation analysis is carried out, and it is believed that the attacks are all manipulated by the same organization. Beijing Venustech Inc. [27] analyzes the shellcode function and code similarity of some samples of vulnerabilities as the characteristics of correlation analysis and then traces the source of the Hedwig organization. Lockheed Martin proposed an advanced continuous threat kill chain model [28], which divided APT attack activities into Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control, and Actions on Objectives, They performed these 7 steps and pointed out that as long as the defending party detects and blocks one of these steps, the attack can be prevented from occurring. MITRE Company adopted a similar idea and proposed a more detailed ATT&CK framework [29]. ATT&CK integrates known historical and actual advanced threat attack tactics and technologies to form a common language for hacking descriptions and an abstract knowledge base framework for hacking attacks. Its official website has published descriptions of 87 APT attacking organizations, including the software, tactics, techniques, and procedures (TTP) used by the attacking organizations.

2. The Proposed Method

This paper proposes a classification method for attribution organizations with APT malware based on machine learning. Based on malicious software samples in APT attacks, this method first dynamically analyzes samples, preprocesses the acquired behavior data, constructs a behavioral data set of malware samples, then uses the TF-IDF method to perform the feature representation forms a vector matrix, and calculates the chi-square value of the high-latitude feature vector to perform feature selection. Based on the SMOTE-RF model designed in this paper, the multiclass model is trained and finally, the test set is predicted and output. The overall design framework of this article is shown in Figure 1.

2.1. Feature Representation. In this paper, we use the behavior feature data set that is the APT data set provided by NSFOCUS. They collected and obtained the dynamic information of a large amount of malware in the sandbox and marked the APT organization to which it belongs. This experiment selected sample data of 7 APT organizations to form the original data set, and the information is shown in Table 1.

The behavioral data of the samples in this dataset contain a lot of redundant data, including path data generated when the malware executes operations, various files called by the malware, APIs, operation object data, and other information (see Figure 2).

The behavior data of a malware sample is in the form of text, as shown in the diagram of a sample behavior data of a malware sample. Therefore, before model training, the text data must be quantified. According to our statistics, the text character length of most samples is below 10,000, so the first 10,000 characters are intercepted for the text data of each sample (see Figure 3). Then, we choose to use TF-IDF (term frequency-inverse document frequency), that is, the word frequency and inverse document frequency method to weight each word to vectorize the text. The TF-IDF method consists of two parts: term frequency (TF) and inverse document frequency (IDF). The former represents the frequency of a term in a document and is used to describe the importance of a term to a document. The latter represents the proportion of documents that contain a certain term and is used to measure whether the term is common or rare in all documents.

If the $n_{i,j}$ represents the frequency of the term v_i in the document d_j , the word frequency of the term v_i represents as

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{i,k}}. \quad (1)$$

Inverse text frequency of term $IDF_{i,j}$ is expressed as

$$IDF_{i,j} = \log \frac{|D|}{\|d_j: v_i \in d_j\| + 1}, \quad (2)$$

where $|D|$ represents the total number of all documents in the corpus and $\|d_j: v_i \in d_j\|$ represents the total number of documents in the corpus containing the term v_i . Therefore, the term frequency and inverse text frequency $TF - IDF_{i,j}$ of the term v_i in the document d_j are expressed as

$$TF - IDF_{i,j} = TF_{i,j} \times IDF_{i,j}. \quad (3)$$

If the TF value of a word extracted from the behavior data is very high but the IDF value is very low, it indicates that the word may be important to the attack.

When using the TF-IDF algorithm to identify keywords in behavioral data, treat all the data extracted from the same sample as an independent document d_j , and all d_j constitute a corpus. Calculate the TF-IDF value for each word v_i in d_j . To save costs, improve efficiency, and reduce false alarms, a set of "stop words" is constructed as a white list. By default, these words are used very widely and do not affect classification. This set includes common English words in the

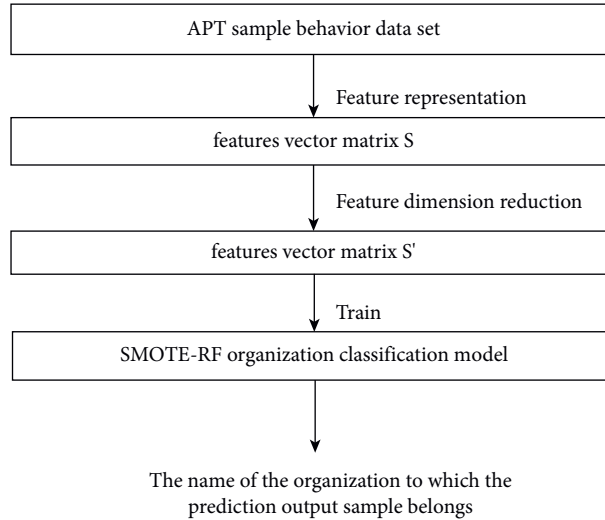


FIGURE 1: Overall design framework diagram.

TABLE 1: Distribution of the APT organization.

	APT organization name	Numbers
1	Lazarus	1060
2	APT28	343
3	Operation C-Major	276
4	APT29	273
5	Dropping Elephant	156
6	Sandworm	154
7	Naikon	127

```

['KERNEL32.dll',
'ADVAPI32.dll',
'HKCU\\SESSIONINFORMTION\\ProgramCount',
'WS2_32.dll',
'USER32.dll',
'C:\\WINDOWS\\explorer.exe',
'MSVCRT.dll']
  
```

FIGURE 2: A sample diagram of the behavior data of a malware sample.

operating system, such as “microsoft, documents, desktop.” The words in this set will be automatically eliminated, and the TF-IDF value will no longer be calculated. Besides, filter out some data that appear too frequently and too much.

Finally, after the above calculations, the behavior data of the data set are represented as a feature matrix S , which includes more than one thousand features. In addition, we analyzed the top 20 features TF-IDF value size in each organization and the size distribution of some features is different in each organization, indicating that our feature representation is effective, and this method can obtain some features with a degree of difference (see Figures 4–10).

2.2. Feature Dimensionality Reduction. Since the feature representation generates many feature dimensions and sparse feature vector values in the previous step, dimensionality reduction of feature vectors is a more feasible method to increase speed and efficiency of detection and

improve the model fitting effect. Here, the chi-square test is used for feature dimensionality reduction. The chi-square test (CHI) is also called χ^2 statistic, which is used to test the independence of two variables. The chi-square test feature selection algorithm is mainly used to determine the correlation between the feature item t_i and the category c_j . Time obeys the χ^2 distribution, and the value χ^2 reflects the degree of correlation between t_i and c_j . If a word has a higher χ^2 relative to a certain category, it indicates that the word has a great correlation with that category. The calculation method is shown in

$$\chi^2(t_i, c_j) = \frac{N \times (A D - B C)^2}{(A + C) \times (B + D) \times (A + B) \times (C + D)} \quad (4)$$

Among them, A means the number of texts belonging to the category c_j and containing the feature item t_i ; B means the number of texts containing t_i but not belonging to c_j ; C means the number of texts that do not contain t_i but belong

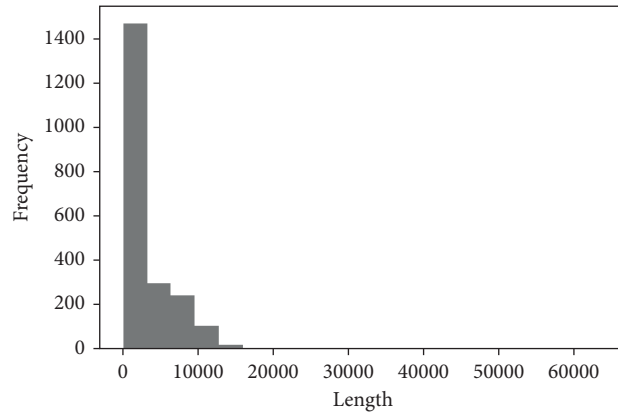


FIGURE 3: The length histogram of the sample’s behavior text data.

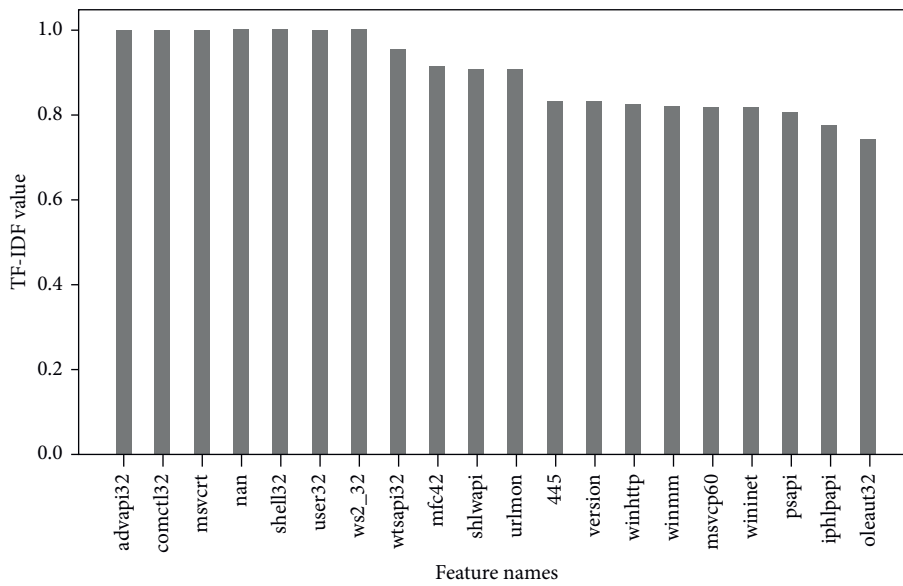


FIGURE 4: The top 20 features of the Lazarus group sample.

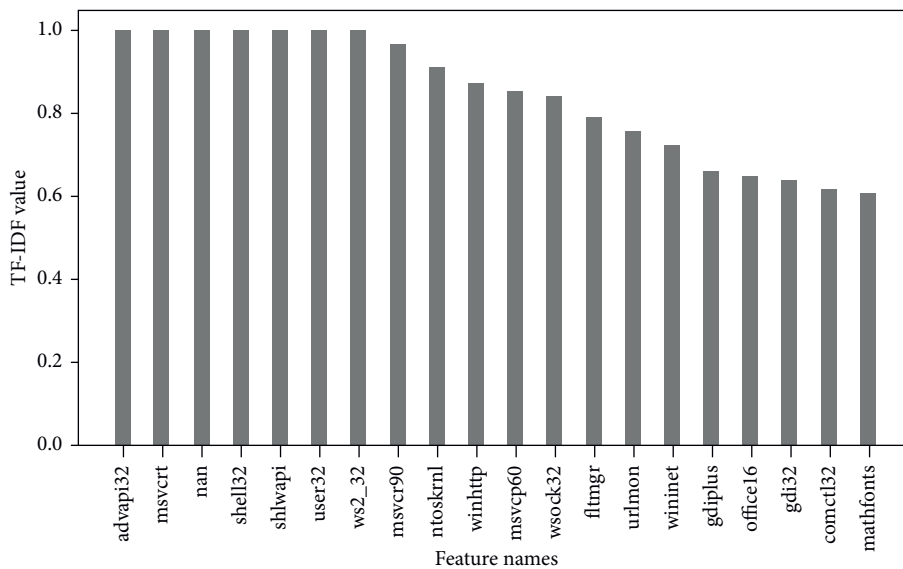


FIGURE 5: The top 20 features of the APT28 sample.



FIGURE 6: The top 20 features of the Operation C-Major sample.

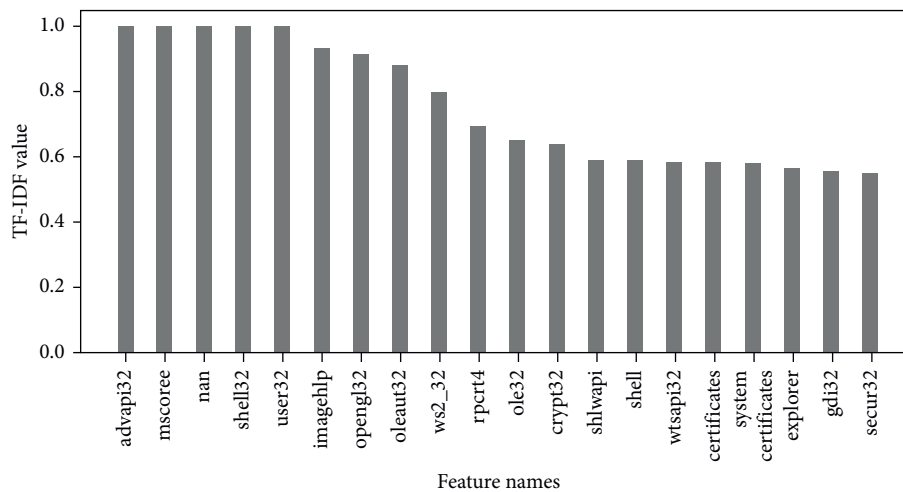


FIGURE 7: The top 20 features of the APT29 sample.

to c_j ; D represents the number of texts that do not belong to c_j and do not contain t_i feature items, and $N = A + B + C + D$. The feature items selected by the chi-square test have a strong correlation with the text category, and the feature items represent category information. According to the calculation formula, the first k words of each category are selected as features according to certain requirements. Calculate the χ^2 statistics between each feature and the category standard, and finally select the k features with the highest score χ^2 .

We calculate the chi-square value of more than one thousand features generated after feature representation. The larger the chi-square value, the better the ability of the feature to distinguish the sample. The top 20 characteristics of chi-square value size are shown in Figure 11.

2.3. Classification Model. To deal with the trouble of unbalanced classification and multiclassification in APT data sets, this paper designs the SMOTE-RF model. The model integrates SMOTE and random forest algorithms. The SMOTE algorithm is a simple and effective oversampling method proposed by Chawla et al. [30]. This method randomly selects k nearest neighbors among minority samples and increases the number of minority samples through interpolation with k nearest neighbors to improve imbalanced data set distribution. Random forest is an integrated algorithm based on decision tree learners. It uses bootstrap resampling technology of the self-service method to randomly select k samples from the original training sample set N with replacement to generate a new training sample set. Some samples may be selected multiple times, and some may

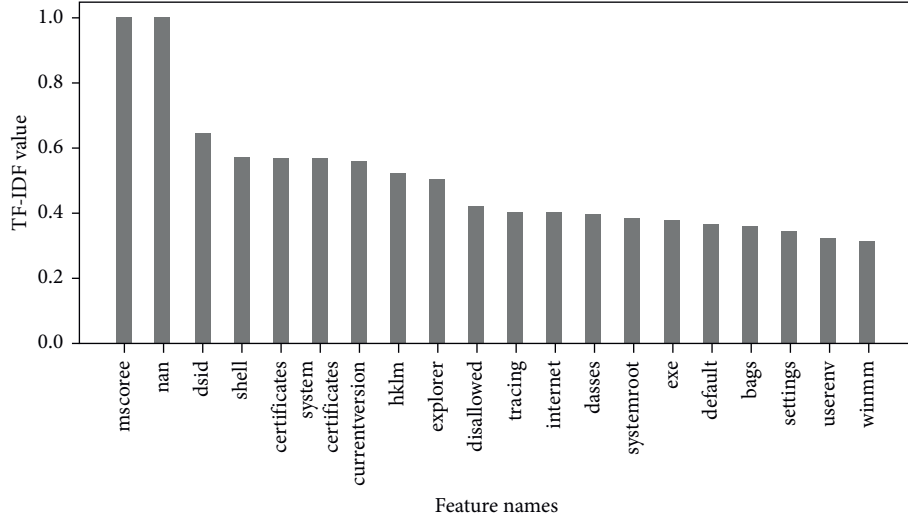


FIGURE 8: The top 20 features of the Dropping Elephant sample.

not be selected once and then generate k classification trees to form a random forest based on the self-service sample set. The final classification result is voted by all classification trees in the forest, and the algorithm has good generalization ability.

The SMOTE-RF model is first based on the number of samples N of the category with the largest number of samples in the data set S' and uses the SMOTE algorithm to generate N new samples for the samples of each other category. Then, multiclassification training is performed based on the random forest algorithm to obtain the classification model, and finally, the output category is predicted.

The SMOTE-RF model construction process is divided into seven steps.

The original training set is S_{train}' , the class with the largest number of samples is N , other classes are classified as minority classes, the minority class sample set is M , i is a sample of the minority class, and its feature vector is x_i , $i \in (1, \dots, M)$:

Step 1: calculate the Euclidean distance between each sample x_i in the minority M samples and the minority sample M . Get the k nearest neighbors of the sample.

Step 2: randomly select N samples from the k nearest neighbors, and each sample and its selected N nearest neighbor samples are combined into N new samples according to the following equation:

$$x_i(\text{new}) = x_i + \text{rand}(x_{ij} - x_i). \quad (5)$$

Among them, $x_i(\text{new})$ represents the newly added minority sample and x_{ij} indicates the j the nearest neighbor sample of x_i , where $j = 1, \dots, N$.

Step 3: put the newly synthesized samples into the original training set S'_{train} to form a new balanced training set S'_{train} .

Step 4: use bootstrap resampling technology to randomly select T samples from all T samples in the new

training set S'_{train} , and select T samples to train a decision tree.

Step 5: assuming that each sample has F features, when each node of the decision tree is split, randomly select f features ($f < F$) from these F features as candidate features and then select from candidate features. The feature that yields the best value splits the nodes of the decision tree.

Step 6: follow steps 4-5 to generate T decision trees to construct a random forest.

Step 7: vote the classification target through all the trees, and the classification with the most votes is the final classification result.

3. Experimental Results and Analysis

3.1. Model Evaluation Index. The experiment in this paper is a multiclassification. In order to comprehensively investigate various classifications, the performance indicators choose precision, recall, and F1-score. Also, a confusion matrix is used to represent the results of this classification, as shown in Table 2.

For the i -th category ($1 \leq i \leq n$), precision (P_i), recall rate (R_i), and F-score (F_score_i), respectively, are

$$P_i = \frac{c_{ii}}{\sum_j c_{ji}}, \quad (6)$$

$$R_i = \frac{c_{ii}}{\sum_j c_{ij}}$$

$$F_score_i = 2 \times \frac{P_i * R_i}{P_i + R_i}. \quad (7)$$

Finally, the arithmetic average of the indicators of each category is calculated to obtain the macro average, which is used to measure the overall effect of each algorithm classification:

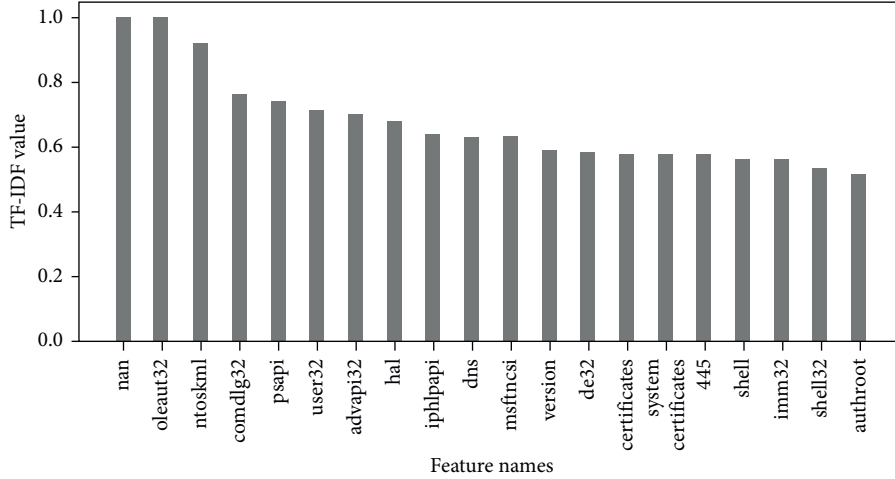


FIGURE 9: The top 20 features of the Sandworm sample.

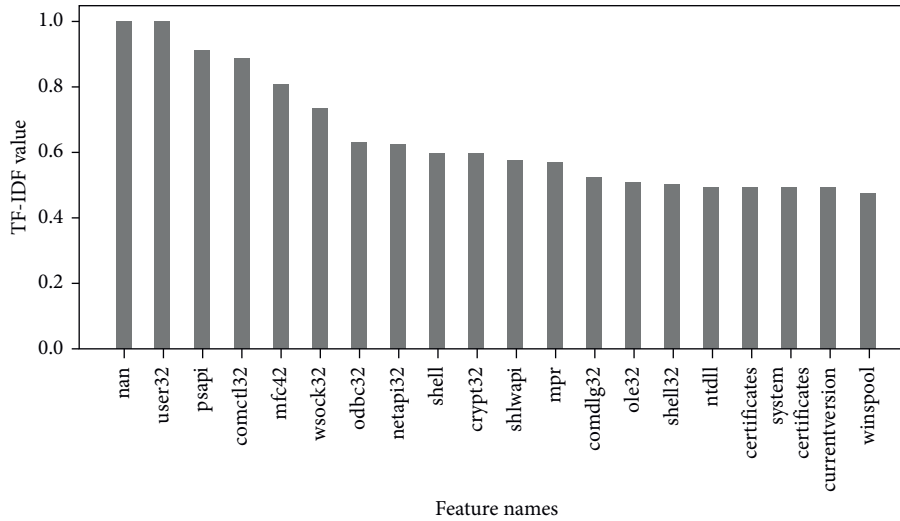


FIGURE 10: The top 20 features of the Naikon sample.

$$P_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n P_i,$$

$$R_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n R_i, \quad (8)$$

$$F_{\text{score}}_{\text{macro}} = 2 \times \frac{P_{\text{macro}} * R_{\text{macro}}}{P_{\text{macro}} + R_{\text{macro}}}.$$

3.2. Experimental Results. To compare the prediction results, the algorithms that often perform well on classification tasks, such as KNN algorithm, DT algorithm, and XGBoost algorithm, are selected here and the SMOTE-RF model of this article is compared and verified by experiments. The prediction results of each model in each category are shown in Table 3. It can be seen that APT29,

Dropping Elephant, and Sandworm have better classification effects on KNN, DT, XGB, and SMOTE-RF models, respectively. Operation Sandworm has the best classification effect on the SMOTE-RF model, with an F-score reaching 0.939. The classification effect of the Dropping Elephant organization on the DT model is the best, with an F-score reaching 0.903. The classification effect of Operation C-Major organization on the four models is the same. Lazarus, APT28, APT29, and Naikon organizations all achieved the highest F-score on the SMOTE-RF model. The F-score of Lazarus Group and APT 28 is relatively low. After data analysis, the main reason is that some samples have fewer text data, which leads to very few effective features extracted. In addition, combining the performance of each model on the training set (see Figure 12) and test set (see Figure 13), it can be seen that the DT model is the most unstable and the SMOTE-RF model has the best overall classification effect and stability. Our

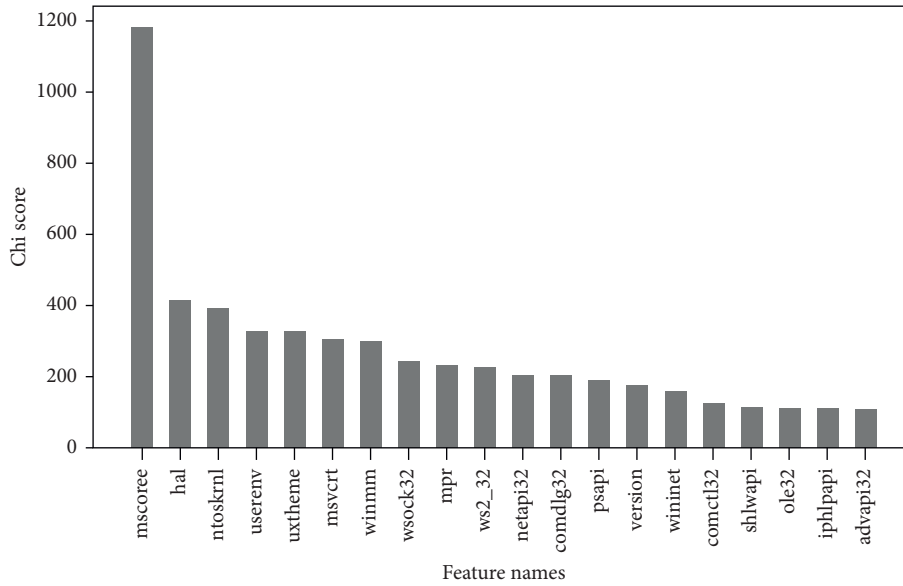


FIGURE 11: Top 20 features of the chi-square value.

TABLE 2: Confusion matrix representation of classification results.

True value	Predictive value		
	Group 1	Group 2	Group 3
Group 1	c_{11}	c_{12}	c_{13}
Group 2	c_{21}	c_{22}	c_{23}
Group 3	c_{31}	c_{32}	c_{33}

TABLE 3: Classification results of each model on each APT organization.

APT organization	Evaluation	KNN	DT	XGB	SMOTE-RF
Lazarus group	Precision	0.791	0.750	0.800	0.845
	Recall	0.507	0.493	0.478	0.567
	F-score	0.618	0.595	0.598	0.644
APT28	Precision	0.360	0.351	0.355	0.366
	Recall	0.854	0.833	0.792	0.854
	F-score	0.506	0.494	0.490	0.513
Operation C-Major	Precision	0.889	0.889	0.889	0.889
	Recall	0.828	0.828	0.828	0.828
	F-score	0.857	0.857	0.857	0.857
APT29	Precision	0.937	0.938	0.912	0.968
	Recall	0.825	0.839	0.857	0.834
	F-score	0.877	0.886	0.884	0.896
Dropping Elephant	Precision	0.927	0.980	0.944	0.927
	Recall	0.836	0.836	0.836	0.836
	F-score	0.879	0.903	0.887	0.879
Sandworm	Precision	0.840	0.917	1.0	1.0
	Recall	0.808	0.846	0.846	0.885
	F-score	0.824	0.880	0.917	0.939
Naikon	Precision	0.913	0.957	0.917	0.957
	Recall	0.700	0.733	0.733	0.733
	F-score	0.792	0.830	0.815	0.830

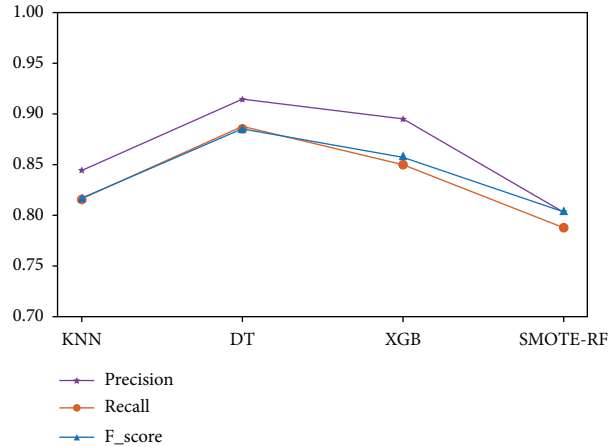


FIGURE 12: Performance indicators of each model on the train set.

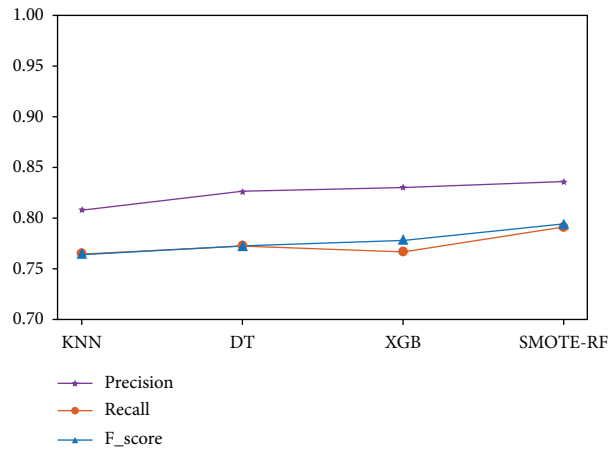


FIGURE 13: Performance indicators of each model on the test set.

experimental results prove the effectiveness of our feature extraction method and the superiority of our model.

4. Conclusions

In recent years, cyber-attacks are being used by various countries and intelligence agencies as one of the important means to achieve their political, diplomatic, military, and other purposes. The detection of APT has aroused widespread concern in information security and academic research circles. The classification of the attribution of APT malware samples is conducive to constructing attack scenarios, tracking attackers, and effectively identifying APT attack organizations of subsequent incidents. This paper proposes a classification method of APT organizations based on machine learning and malware. This method is based on the behavior data with APT organization tags obtained from dynamic analysis of APT malicious software acquired from the Internet of Things devices, and relatively strong feature vectors are obtained through feature representation and feature dimensionality reduction. Considering the sample imbalance in the data set, this paper designs a SMOTE-RF model that integrates SMOTE and random forest

algorithms. Finally, the effectiveness of the proposed method for the attribution analysis of APT malware is verified by multiple sets of experiments. Among them, our method of feature extraction can achieve more than 80% accuracy in general models and the SMOTE-RF model performs well and has stable performance in the classification of APT malware. Next, we will combine non-APT malware samples to further study the features of APT attacks and each organization and to better identify APT attack activities and protect the security of the next generation of complex networks.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Shudong Li and Qianqing Zhang contributed equally to this work.

Acknowledgments

This research was funded by the Key R D Program of Guangdong Province (No. 2019B010136003), NSFC (Nos. 62072131 and 61972106), Science and Technology Projects in Guangzhou (No. 202102010442), National Key Research and Development Program of China (No. 2019QY1406), Open Project of National Engineering Laboratory for Mobile Internet System and Application Security, and Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2019). The authors thank the data provided by the NSFOCUS company.

References

- [1] L. Xiao, X. Wan, C. Dai, X. Du, X. Chen, and M. Guizani, "Security in mobile edge caching with reinforcement learning," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 116–122, June 2018.
- [2] H. Yang, S. Li, X. Wu, H. Lu, and W. Han, "A novel solution for malicious code detection and family clustering based on machine learning," *IEEE Access*, vol. 7, no. 1, pp. 148853–148860.
- [3] I. Ghafir and V. Prenosil, "Advanced persistent threat attack detection: an overview," *International Journal Of Advances In Computer Networks And Its Security*, vol. 4, no. 4, pp. 154–158, 2014.
- [4] T. M. Chen, "Stuxnet, the real start of cyber warfare? [Editor's Note]," *IEEE Network*, vol. 24, no. 6, pp. 2–3, 2010.
- [5] W. Chen, X. Helu, C. Jin et al., "Advanced persistent threat organization identification based on software gene of malware," *Transactions on Emerging Telecommunications Technologies*, vol. 31, 2020.
- [6] S. Li, D. Zhao, X. Wu, Z. Tian, A. Li, and Z. Wang, "Functional immunization of networks based on message passing," *Applied Mathematics and Computation*, vol. 366, Article ID 124728, 2020.
- [7] A. S. Bist and S. Jalal, "Identification of metamorphic viruses [C]," in *Proceedings of the 2014 IEEE International Advance Computing Conference (IACC)*, pp. 1163–1168, Gurgaon, India, February 2014.
- [8] "Going ATOMIC: clustering and associating attacker activity at scale," 2019, <https://www.fireeye.com/blog/threat-research/2019/03/clustering-and-associating-attacker-activity-at-scale.html>.
- [9] F. Turkmen, S. Foley, B. O'Sullivan, W. Fitzgerald, T. Hadzic, and M. S. Basagiannis, "Explanations and relaxations for policy conflicts in physical access control," in *Proceedings of the 25th IEEE International Conference on Tools with Artificial Intelligence*, pp. 330–336, IEEE Press, Herndon, VA, USA, November 2013.
- [10] R. Koike, N. Nakaya, and Y. Koi, "Development of system for the automatic generation of unknown virus extermination software," in *Proceedings of the 2007 International Symposium on Applications and the Internet*, Hiroshima, Japan, January 2007.
- [11] S. Li, L. Jiang, X. Wu, W. Han, D. Zhao, and Z. Wang, "A weighted network community detection algorithm based on deep learning," *Applied Mathematics and Computation*, vol. 401, Article ID 126012, 2021.
- [12] S. Shen, H. Zhou, F. Sheng et al., "HSIRD: a model for characterizing dynamics of malware diffusion in heterogeneous WSNs," *Journal of Network and Computer Applications*, vol. 146, 2019.
- [13] S. Li, Y. Li, W. Han, X. Du, M. Guizani, and Z. Tian, "Malicious mining code detection based on ensemble learning in cloud computing environment," *Simulation Modelling Practice and Theory*, Article ID 102391, 2021.
- [14] M. Fan, S. Li, X. Wu, W. Han, Z. Gu, and Z. Tian, "A novel malware detection framework based on weighted heterograph," in *Proceedings of the CIAT 2020: 2020 International Conference on Cyberspace Innovation of Advanced Technologies*, vol. 4-6, pp. 39–43, Guangzhou China, December 2020.
- [15] P. Zhu, X. Wang, D. Jia, Y. Guo, S. Li, and C. Chu, "Investigating the co-evolution of node reputation and edge-strategy in prisoner's dilemma game," *Applied Mathematics and Computation*, vol. 386, Article ID 125474, 2020.
- [16] G. Zhao, K. Xu, L. Xu, and B. Wu, "Detecting APT malware infections based on malicious DNS and traffic analysis," *IEEE Access*, vol. 3, pp. 1132–1142, 2015.
- [17] M. Abomhara and G. M. Kien, "Cyber security and the Internet of Things: vulnerabilities, threats, intruders and attacks," *Journal of Cyber Security and Mobility*, vol. 4, no. 1, pp. 65–88, 2015.
- [18] Y. Sung, P. K. Sharma, E. M. Lopez, and J. Park, "Fs-opensecurity: a taxonomic modeling of security threats in sdn for future sustainable computing," *Sustainability*, vol. 8, no. 9, 2016.
- [19] M. Lee and D. Lewis, "Clustering disparate attacks: mapping the activities of the advanced persistent threat," 2013, https://www.virusbulletin.com/uploads/pdf/conference/_slides/2011/Lee-VB2011.pdf Accessed.
- [20] 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.
- [21] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine learning techniques," *IEEE Internet of Things Journal*, vol. 8, 2020.
- [22] S. Shen, H. Ma, E. Fan et al., "A non-cooperative non-zero-sum game-based dependability assessment of heterogeneous WSNs with malware diffusion," *Journal of Network and Computer Applications*, vol. 91, pp. 26–35, 2017.
- [23] S. Yu, G. Gu, A. Barnawi, S. Guo, and I. Stojmenovic, "Malware propagation in large-scale networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 1, pp. 170–179, 2015.
- [24] S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: a survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 925–941, 2014.
- [25] Y. Qiao, X. Yun, and Y. Zhang, "How to automatically identify the homology of different malware," in *Proceedings of the 2016 IEEE Trustcom/BigDataSE/ISPA*, pp. 929–936, Tianjin, China, August 2016.
- [26] N. Moran and J. Bennett, "Supply chain analysis: from quarter master to sunshop," Technical Report <https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-maware-supply-chain.pdf>, Fire Eye Labs, 2013.
- [27] <http://www.freebuf.com/news/92945.html>.
- [28] E. Hutchins, M. Cloppert, and R. Amin, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," *Leading Issues in Information Warfare and Security Research*, vol. 1, 2011.

- [29] “The MITRE Corporation, ATT&cK matrix[EB/OL],” 2020, <https://attack.mitre.org/>.
- [30] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research*, vol. 16, no. 1, pp. 321–357, 2002.

Research Article

V-LDAA: A New Lattice-Based Direct Anonymous Attestation Scheme for VANETs System

Liquan Chen ^{1,2}, Tianyang Tu,¹ Kunliang Yu,¹ Mengnan Zhao,¹ and Yingchao Wang¹

¹School of Cyber Science and Engineering, Southeast University, Nanjing, China

²Purple Mountain Laboratories for Network Communication and Security, Nanjing, China

Correspondence should be addressed to Liquan Chen; lqchen@seu.edu.cn

Received 10 June 2021; Accepted 12 August 2021; Published 2 September 2021

Academic Editor: Jinguang Han

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

Privacy protection and message authentication issues in VANETs have received great attention in academia. Many authentication schemes in VANETs have been proposed, but most of them are based on classical difficult problems such as factorization in RSA setting or Elliptic Curve setting and are therefore not quantum resistant. If a quantum computer becomes available in the next few decades, the security of these schemes will be at stake. This paper presents a vehicular lattice-based direct anonymous attestation (V-LDAA) scheme adopting an optimized signature scheme based on automorphism stability which achieves postquantum security. A distributed pseudonym update and vehicle revocation mechanism based on the lattice is introduced in this paper, which means vehicles can update their pseudonyms and revoke the identity certificate by themselves without the need for pseudonym resolutions or CRLs checking. Compared with the existing lattice-based attestation schemes in VANETs, computation costs during signing and verification operations in V-LDAA are no longer related to the number of users, which makes it suitable for large-scale VANETs. Security analysis shows that V-LDAA resists TPM theft attacks and provides users with user-controlled anonymity, user-controlled unlinkability, and unforgeability against quantum adversaries. Experimental results show that V-LDAA reduces the blind signature size by 18%. The speed of blind signing is increased by 30%, and blind verification operation is accelerated 3 times compared with the existing lattice-based direct anonymous attestation (LDAA) scheme.

1. Introduction

The Intelligent Transportation System (ITS) provides vehicles with intelligent and efficient services, such as collision avoidance, traffic condition reports, and entertainment services, etc. Messages are sent to various network nodes through vehicular ad hoc networks (VANETs) [1]. VANET is a key facility of an intelligent transportation system, which is composed of Certification Authority (CA), roadside units (RSUs), and on-board units (OBUs) [2]. Among them, the OBU is responsible for supporting the V2I communication between the roadside units and the vehicle and V2V communication between vehicles. These nodes are connected to each other to form a network, and the communications in the entire network are achieved through the information transferring among adjacent nodes. The key issue that needs to be solved in the implementations for ITS

is how to protect the security and privacy of users in VANETs. Vehicle users in ITS need to send information about their location, speed, and other driving conditions, or traffic jams, icy roads, and other surrounding road conditions to adjacent users. If this information is maliciously tracked or tampered with by an adversary, it will cause serious privacy leakage accidents and even threaten the life of the driver. For example, the adversary can obtain the real location information of the vehicle by tracing the navigation route information or modify the traffic information, which may lead to traffic paralysis or even serious traffic accidents. Therefore, an anonymous attestation protocol in VANETs needs to be established to ensure the anonymity of users and the integrity and untraceability of messages.

In addition, with the development of quantum computing technology, the security of traditional public key cryptosystems has received an impact. Most of the existing

authentication protocols in VANETs have their security supported on classic difficult problems such as factorization in RSA setting or Elliptic Curve setting. Under traditional computing conditions, these difficult problems can only be solved in exponential or subexponential time. However, according to Shor's algorithm, quantum computers can efficiently solve these problems, leading to the failure of traditional cryptosystems. Thus, there is a need to introduce quantum-resistant authentication schemes in VANETs.

We have proposed the following major contributions in this paper.

- (1) A vehicular lattice-based direct anonymous attestation scheme that achieves postquantum security is proposed in this paper. In this scheme, a lattice-based distributed pseudonym update and certificate revocation mechanism is introduced. By embedding a trusted platform module (TPM) in each vehicle, trust is distributed from Certification Authority (CA), pseudonym provider (PP), Revocation Authority (RA), and other authoritative institutions to each legitimate user, transforming a centralized trust system into a distributed trust system. "Distributed trust" is reflected in the processes of pseudonym update and vehicle revocation. Users can generate pseudonyms by themselves without the need for regular updates and distributions by PP. TPM performs the revocation operation independently, without RA performing pseudonym resolution operations, and there is no need to maintain the certificate revocation lists CRLs. Moreover, the calculation costs in signing operations are no longer related to the number of members. Thus, it is more suitable for large-scale VANETs.
- (2) V-LDAA optimizes the signature scheme based on automorphism stability which is used in the *Blind-Sign* and *BlindVerify* protocols of the original LDAA scheme. The optimized signature scheme reduces the number of automorphisms that need to be proven stable, which simplifies the processes of signing and verification and reduces the signature size. Based on the experimental implementation of the V-LDAA scheme, the high computation and storage efficiency of the proposed scheme is confirmed.
- (3) V-LDAA binds TPM and Host to jointly generate an identity certificate in Join protocol to resist TPM theft attacks. This is important in VANETs because it prevents TPM from being transplanted to a new vehicle platform by an adversary and signed with the replaced identity certificate.

The rest of this paper is organized as follows. We first introduce related works, the background knowledge, an optimized signature scheme used in V-LDAA and VANET architectures based on V-LDAA. Then, the construction of the proposed V-LDAA scheme is described. After that,

security and performance analysis are detailed. Finally, the conclusion of this paper is presented.

2. Related Works

In recent years, research studies on authentication schemes mainly focused on the following aspects. The first is based on a symmetric key mechanism [2]. The sender uses a shared key to generate the message authentication code (MAC), while the receiver verifies it before accepting the message. However, because both parties need to share the private key, the mechanism based on message authentication code cannot withstand a large number of node tampering attacks in the network. In addition, the adversary can cheat any individual node to obtain the private key, which can be used for message authentication. The second is an identity-based encryption system [3, 4], where the trusted authority is responsible for the generation and distribution of public and private key pairs for legitimate members. However, under this mechanism, the adversary can easily obtain the user's real identity from the signature and track the signature. The third one is an authentication scheme based on vehicle public key infrastructures (VPKIs), which is also the design idea of this paper. CA is responsible for registering and managing long-term identity certificates of members, while members sign messages through short-term pseudonym certificates. The VPKIs scheme can meet the anonymity property and provide a pseudonym mechanism, but there are still many shortcomings. In this scheme, the security risk and computation burden are caused by different pseudonym update strategies. In order to prevent users from being maliciously tracked, CA needs to change pseudonyms for all users regularly [1]. In the case of unconditional security, the pseudonym should be changed every time the signature is signed, which causes a huge computational and storage burden when PP generates new pseudonym certificates and distributes them to every legitimate user periodically. In [5], an optimized pseudonym update scheme is proposed, but its computation costs still burden the vehicle and the Pseudonym Provider (PP). In addition, in order to revoke the identity certificate of an illegal vehicle, the Revocation Authority (RA) needs to resolve the user's long-term identity ID value from the user's pseudonym and save it to certificate revocation lists (CRLs) for all users to query. The update, query, maintenance, and storage of CRLs cause heavy computation and storage costs.

The existing authentication schemes for VANETs which achieve postquantum security are mainly lattice-based ring signature schemes [6–8]. In the lattice-based ring signature scheme, each member needs to use its private key and the public keys of all other members to sign the message, and the members in a ring need to change with the specific driving position of the vehicle. In recent years, several lattice-based direct anonymous attestation (LDAA) schemes are proposed by updating the cryptographic primitives to be quantum resistant in direct anonymous attestation (DAA) [9–11]. The

first LDAA in [9] is based on a lattice-based MAC scheme and a CMA-secure digital signature scheme, but it suffers from high computation costs in signing protocol. LDAA in [10] adopts a noninteractive sigma protocol construction and a modified Boyen's signature scheme, which can improve signing and storage efficiency compared to LDAA in [9]. Among them, the lattice-based direct anonymous attestation in [11] is most suitable for a future quantum-resistant TPM for its high efficiency. LDAA becomes an interesting candidate for the postquantum secure authentication protocol in VANETs because of its balance in authentication and anonymity.

3. Preliminaries

3.1. Notation. Symbols used in this paper are illustrated in Table 1 with their definitions.

3.2. Trapdoor Sampling. Sample two short vectors s_1, s_2 satisfying

$$[\mathbf{a} \mid \mathbf{b} + i[1\sqrt{q}]] \cdot \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \mathbf{u} + \mathbf{a}_2 \cdot \mathbf{e}, \quad (1)$$

where i is a nonzero element in \mathbb{Z}_q . According to [12], there is a set of basis $\mathbf{S} \in \mathbb{Z}^{4d \times 4d}$ for $\Lambda^\perp = \{x \in \mathbb{R}^4 \mid [\mathbf{a} \mid \mathbf{b} + i[1\sqrt{q}]] \cdot x \equiv 0 \pmod{q}\}$. The Gram-Schmidt orthogonalization of $\mathbf{S} \in \mathbb{Z}^{4d \times 4d}$ satisfies $\|\tilde{\mathbf{S}}\| \leq (s_1(\mathbf{R} + 1))\sqrt{\delta^2 + 1}$ with $\delta = \sqrt{q}$. To sample s_1, s_2 , first calculate an arbitrary solution (not necessarily short solutions) that satisfies (1). Then express it in basis \mathbf{S} , and use the randomized nearest plane discrete Gaussian sampling algorithm in [13] to get solutions distributed as the discrete Gaussian distribution with $s = 2 \cdot \|\tilde{\mathbf{S}}\| \leq 2(3\sqrt{d} + 1)\sqrt{\delta^2 + 1}$. The algorithm is called MP-Sampler.

3.3. Lattice-Based Commitment Scheme. We use the commitment scheme from [14] with M-LWE based hiding property and M-SIS based binding property. Define public parameters $\mathbf{A}_1 \in \mathcal{R}_{q_1}^{1 \times k}$, $\mathbf{A}_2 \in \mathcal{R}_{q_2}^{l \times k}$, where $\mathbf{A}_1 = [1 \mid \mathbf{A}'_1]$, $\mathbf{A}'_1 \xleftarrow{\$}$ and $\mathbf{A}_2 = [0^l \mid \mathbf{I}_l \mid \mathbf{A}'_2]$, $\mathbf{A}'_2 \xleftarrow{\$}$. To commit to a message $m \in \mathcal{R}_{q_2}^{l \times l}$, sample $\mathbf{r} \xleftarrow{\$}$ and compute $\text{Com}(m; r) = \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{r} + \begin{bmatrix} 0 \\ \mathbf{m} \end{bmatrix}$. If there exists $\mathbf{r} \leq B_{\text{com}}$ and $c \in \mathcal{R}$ satisfying $c \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{r} + c \begin{bmatrix} 0 \\ \mathbf{m} \end{bmatrix}$, then the opening $\begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix}$ is valid.

3.4. Lattice-Based Zero-Knowledge Proof. Lattice-based encryption schemes usually include a public A and small coefficient secret value e , which satisfies $Ae = t$. In order to prove that t is a legal ciphertext, a zero-knowledge proof about e needs to be generated, which satisfies $Ae = t$. There are several protocols to achieve zero-knowledge proof about e . The first one is based on a Stern-type protocol to prove a

norm bounded e satisfying exactly $Ae = t$, which is the most accurate but also the most expensive protocol. In V-LDAA, this method can be used in the zero-knowledge proof of TPM and Host secret values in the Join phase because each user only needs to perform it once in the entire certificate lifecycle. The second is to use rejection sampling and lattice-based Fiat-Shamir [15], which proves that $Ae^t = ct$, where c is the difference between two challenge values.

4. An Optimized Signature Scheme Based on Automorphism Stability of the Cyclotomic Field

The signature schemes of the LDAA schemes in [9, 10] both use Boyen's signature framework under the standard security model [16]. Although there are studies using polynomial lattices to improve the efficiency of Boyen's signature mechanism [17], the size of its group signature is still around 50 MB [18]. The LDAA framework proposed in [11] uses a selectively secure signature mechanism based on the lattice [19]. The so-called selective security refers to the security of messages that can be fixed in advance (fixed before the attacker communicates with the system). In the case of selectively secure, in order to prove the security of the message to be signed, we have to prove the invertibility of the signed message μ and its stability in a special subset. In [19], a Galois extension of the cyclotomic field was used to prove that μ belongs to a certain subset and is invertible. In this paper, we optimize the selective-secure signature scheme used in [11], reducing the number of automorphisms that need to be proven stable from two to one.

4.1. Galois Group of Cyclotomic Rings. If $T^m - 1$ is separable from K , then $K(\mu_n)$ is the splitting field of $T^m - 1$ on K and $K(\mu_n)/K$ is called a Galois extension. Suppose $K = \mathbb{Q}[X]/(\Phi_m(X))$ is a m -th cyclotomic field of degree $d = \varphi(m)$ with an integer ring $\mathcal{R} = \mathbb{Z}[X]/(\Phi_m(X))$ and its subring $\mathcal{S} \subset \mathcal{R}$. Then, the Galois group G is defined as $G = \text{Gal}(K/\mathbb{Q})$, which consists of all automorphisms of K . The Galois group on the cyclotomic field is isomorphic to \mathbb{Z}_m^\times , that is $j \mapsto \sigma_j: \mathbb{Z}_m^\times \rightarrow \text{Gal}(K/\mathbb{Q})$ where $\sigma_j(X) = X^j$. For the subfield $L \subset K$, there must be a subgroup $H < G$ which is the Galois group K on L , that is $H = \text{Gal}(K/L) = \{\sigma \in G \mid \sigma(x) = x \forall x \in L\}$. According to [19], if $\mu \in \mathcal{R}_q$ satisfies $\sigma(\mu) \equiv \mu \pmod{q\mathcal{R}}$ for all $\sigma \in H$, then μ is in the subfield \mathcal{S}_q of \mathcal{R}_q . Thus, in order to prove $\mu \in \mathcal{S}_q$, we need to prove the stability of μ by all Galois automorphisms in H . In other words, we need to prove the stability of μ under the generators of H .

4.2. Power-of-Two Cyclotomic Rings. Suppose $K = \mathbb{Q}[X]/(X^d + 1)$ is a power-of-two cyclotomic fields, we get $G = \text{Gal}(K/\mathbb{Q}) \cong \mathbb{Z}_{2^d}^\times \cong \mathbb{Z}_2 \times \mathbb{Z}_{d/2}$, which is generated by σ_{-1} and σ_5 , that is $G = \langle \sigma_{-1}, \sigma_5 \rangle$. Consider a subgroup $H = \langle \sigma_{-1}, \sigma_5^k \rangle$, according to [19], the fixed field L of H is generated by $\alpha = X^{d - (d/2k)} - X^{d/2k}$. Consider the parameter used in [11] when $k = 1$, then $H = G = \langle \sigma_{-1}, \sigma_5 \rangle$ and the corresponding fixed field $L = \mathbb{Q}$ and $\mathcal{S}_q = \mathbb{Z}_q$. For every prime number q ,

TABLE 1: Notation.

Notation	Description
\mathbb{Z}_q	Quotient ring $\mathbb{Z}/q\mathbb{Z}$
\mathfrak{q}	The moduli used in the commitment scheme
$K = \mathbb{Q}[X]/(X^d + 1)$	A cyclotomic ring
d	The dimension of ring K
$y \leftarrow D$	y is drawn according to the distribution D
$G = \text{Gal}(K/\mathbb{Q})$	The Galois group of K over \mathbb{Q}
$G = \langle \sigma_{-1}, \sigma_5 \rangle$	Galois group G is generated by σ_{-1} and σ_5
N	The number of users
\mathbf{e}	Lowercase bold letters denote a vector of polynomials
\mathbf{A}	Capital bold letters denote a matrix whose entities are polynomials

\mathcal{S}_q is a field. In this case, it is enough to prove that the message $\mu \in \mathcal{R}_q$ remains unchanged under σ_{-1} and σ_5 . This means that every time the zero-knowledge proof of the identity certificate is performed, similar calculations have to be repeated twice (on σ_{-1} and σ_5), which increases the computational complexity of the protocol and the size of the commitments.

In this paper, we change the subfield to $k = 2$, which means $H = \langle \sigma_{-1}, \sigma_5^2 \rangle$ or $H = \langle \sigma_5 \rangle$. When $H = \langle \sigma_5 \rangle$, the generator of its fixed field L is $\alpha = X^{d/2}$ and the minimal polynomial is $Y^2 + 1$. In this case, only one automorphism stability σ_5 needs to be proved during zero-knowledge proof. We select $\mu \in \mathcal{S}_q$, where $\mathcal{S}_q = \{c_0 + c_1 X^{d/2} \in \mathcal{R}_q \mid c_0, c_1 \in \mathbb{Z}_q\}$ of size q^2 and $q \equiv 3 \pmod{4}$. When TPM chooses its identity value, it computes $\mu = c_0 + c_1 \alpha = c_0 + c_1 X^{d/2} \in \mathcal{R}_q$ with arbitrary $c_0, c_1 \in \mathbb{Z}_q$ and proves that μ remains unchanged under σ_5 ($\mu = \mu$). The process of signing and verification is shown in Table 2.

5. VANET Architectures Based on V-LDAA

The traditional VPKI is shown in Figure 1, which is composed of a Certification Authority (CA), a pseudonym provider (PP), a vehicle Revocation Authority (RA), and user vehicles. The vehicle registers its identity with CA, and CA signs the long-term identity certificate VID to the vehicle after confirming that the vehicle is in a trustworthy state. After the vehicle shows VID to the pseudonym provider PP, PP generates a pseudonym certificate based on VID and issues it to the vehicle user. During V2V communication, the illegal behavior of the vehicle will be reported to PP, and PP will determine whether to revoke the user certificate. When deciding to revoke the user certificate, RA cooperates with PP and CA to resolute the pseudonymous certificate to obtain the user's real identity ID. The violation ID is updated to the certificate revocation lists (CRLs). Every time before the user verifies the signature, it needs to first check whether the sender is in the CRLs. The main shortcomings of the traditional VPKI architecture are high storage and calculation consumption for updating, maintaining and querying CRLs; pseudonym resolution is required when certificate revocation, computing efficiency, and security issues are brought by PP's regular update of pseudonym certificates, etc.

VANET architecture based on V-LDAA is shown in Figure 2. Compared with the traditional VPKI system, a hardware chip TPM is embedded in each user's vehicle

platform. Through the identity certificate, we distribute trust from CA to TPM embedded in each legitimate user, transforming a centralized trust system into a distributed trust system. "Distributed Trust" is reflected in the processes of pseudonym update and vehicle revocation. Users can generate pseudonyms by themselves without the need for regular updates and distribution by PP. During certificate revocation, RA only needs to broadcast the revocation instruction of a certain vehicle, while the target vehicle will check its identity, perform the revocation operation, and return the revocation certificate to RA. The whole process does not involve any pseudonym resolution or operations related to the revocation list CRLs.

6. Proposed V-LDAA Scheme

Based on the LDAA scheme in [11], we propose a V-LDAA scheme in VANETs. The overall V-LDAA scheme includes Setup, Join, Create, Sign/verify, Revoke protocols. The structure of the DAA protocol is redesigned. After the Join phase, each user needs to pass through the Create phase to generate identity credentials $\text{PSCert} = (\text{nym} \parallel \text{sig}_1 \parallel \text{sig}_2)$, where nym is a pseudonym public key, sig_1 is the certificate used to determine the identity when the certificate is revoked, and sig_2 is a blind signature on VID used to verify the identity of its legitimate users. Users can complete the anonymous authentication of the message and the self-revocation of the certificate by holding PSCert . TPM executes the destruction operation of the identity certificate and the pseudonymous certificate, generates the revocation certificate, and returns it to RA. RA verifies the identity certificate and the revocation certificate and confirms that the target vehicle has revoked its identity certificate.

Moreover, we optimize the signature scheme based on automorphism stability of the power-of-two cyclotomic fields. When the user interacts with the CA to generate the VID, the identity ID is selected in the more optimal $k = 2$ cyclotomic field, where $\mathcal{S}_q = \{c_0 + c_1 X^{d/2} \in \mathcal{R}_q \mid c_0, c_1 \in \mathbb{Z}_q\}$. At this time, it is enough to prove automorphism stability once instead of twice as in [11], which optimizes the computational efficiency and signature size during *Blind-Sign*. Finally, in the Join phase, the platform secret value sent to CA is changed to be generated by TPM and Host together instead of TPM alone. This is very important in VANETs, because the TPM chip embedded in the vehicle may be in an unmanned environment, and the adversary can directly steal

TABLE 2: An optimized signature scheme based on automorphism stability of the cyclotomic field.

Message: $M \in \{0, 1\}^*$	Signer	Public: \mathbf{t}, σ_5	Private: \mathbf{r}	Verifier	Message: $M \in \{0, 1\}^*$	Public: \mathbf{t}, σ_5
$y, y_5 \leftarrow D_\xi^3$ $\mathbf{W}_1 = \mathbf{a}_1^T \mathbf{y}$ $\mathbf{W}_{1,5} = \sigma_5^{-1}(\mathbf{a}_1^T) \mathbf{y}_5$ $\mathbf{W}_{2,5} = \mathbf{a}_2^T \mathbf{y} - \sigma_5^{-1}(\mathbf{a}_2^T) \mathbf{y}_5$ $c = \mathbf{H}(\mathbf{t}, \sigma_5, \mathbf{W}_1, \mathbf{W}_{1,5}, \mathbf{W}_{2,5}, M)$ $\mathbf{z} = \mathbf{r}c + \mathbf{y}$ $\mathbf{z}_5 = \sigma_5^{-1}(\mathbf{r})c + \mathbf{y}_5$ if $\text{rej}(\ \mathbf{z}\ , \ \mathbf{z}_5\ , \ \mathbf{r}c\ , \xi) = 1$, abort						
				$\xrightarrow{z, z_5}$		
				$\mathbf{W}'_1 = \mathbf{a}_1^T \mathbf{z} - t_1 c$ $\mathbf{W}'_{1,5} = \sigma_5^{-1}(\mathbf{a}_1^T) \mathbf{z}_5 - \sigma_5^{-1}(t_1) c$ $\mathbf{W}'_{2,5} = \mathbf{a}_2^T \mathbf{z} - \sigma_5^{-1}(\mathbf{a}_2^T) \mathbf{z}_5$ if $\ \mathbf{z}\ , \ \mathbf{z}_5\ \leq \beta_z$ and $c = \mathbf{H}(\mathbf{t}, \sigma_5, \mathbf{W}'_1, \mathbf{W}'_{1,5}, \mathbf{W}'_{2,5}, M)$ Output 1 else Output 0		

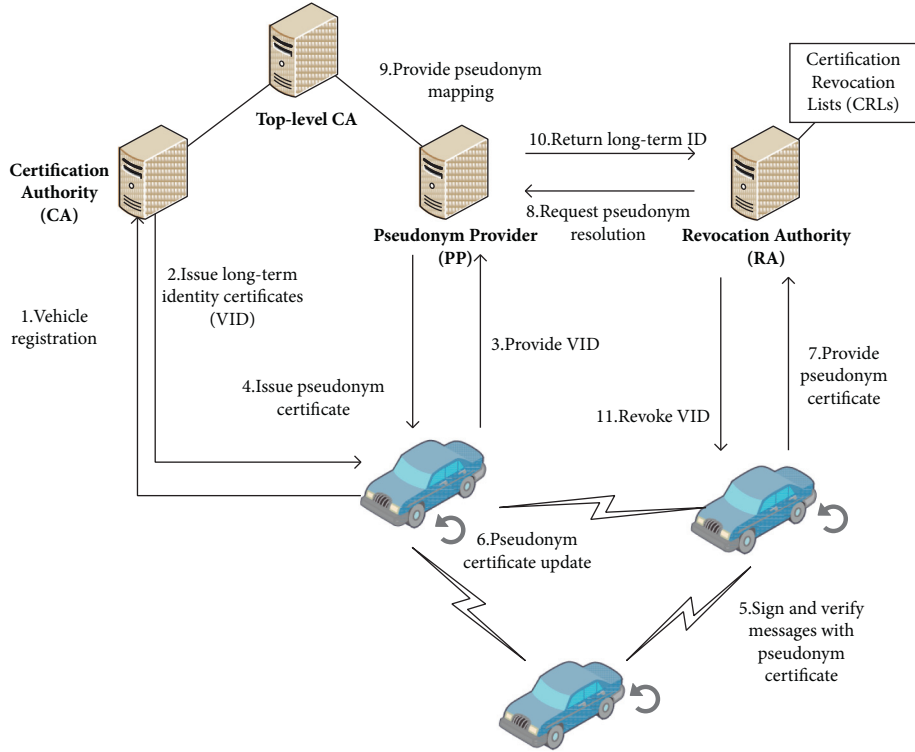


FIGURE 1: Traditional VPKI architectures.

the TPM chip and transplant it to another Host platform to cheat the verifier.

6.1. Setup. We consider a cyclotomic ring $\mathcal{R} = \mathbb{Z}[X]/(X^d + 1)$, $k = 2$ and identity ID in VID $i \in \mathcal{S}_q = \{c_0 + c_1 X^{d/2} \in \mathcal{R}_q | c_0, c_1 \in \mathbb{Z}_q\}$, which keeps stable under σ_5 . Randomly choose $\mathbf{a}_t = [a_1 \ a_2]$ as TPM public parameters, $\mathbf{a}_h = [a_3 \ a_4]$ as Host public parameters and $u \leftarrow \mathcal{R}_q$ as CA public parameter. The private key of CA is a trapdoor $\mathbf{R} \leftarrow \mathcal{R}^{2 \times 2}$

while the public key is $a \leftarrow \mathcal{R}_q$, $\mathbf{b} = [a \ 1] \mathbf{R}$. By Ring-LWE assumption, (a, \mathbf{b}) is indistinguishable from uniform. Thus, we write CA public key as $[\mathbf{a} | \mathbf{b}]$, where $\mathbf{a} = [a \ 1]$.

6.2. Join. TPM randomly select a secret value $\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \leftarrow \mathcal{R}_3^2$ and a private key $sk \in \{0, 1\}^{256}$. Compute $u_t = \mathbf{a} \cdot \mathbf{e} = a_1 e_1 + a_2 e_2$ and send u_t to the Host. Similarly, the Host chooses its secret $\mathbf{e}' = \begin{bmatrix} e'_1 \\ e'_2 \end{bmatrix} \leftarrow \mathcal{R}_3^2$ and computes $u_h = a_3 e'_1 + a_4 e'_2$. Then,

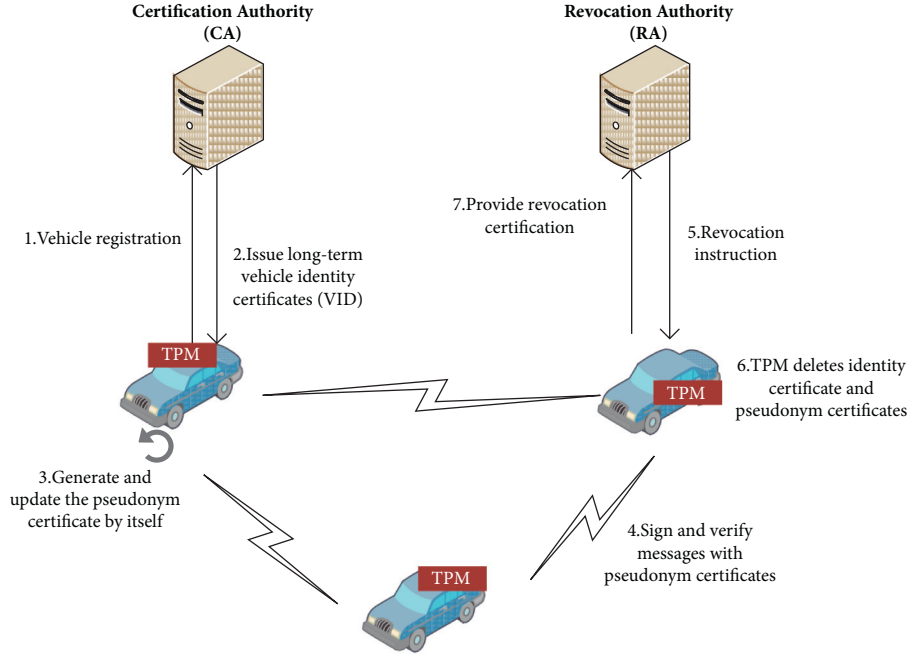


FIGURE 2: VANET architecture based on V-LDAA.

the Host adds u_t and u_h to generate u_1 . TPM and Host jointly give a zero-knowledge proof π_1 of short \mathbf{e} and e' . (u_1, π_1) is sent by Host to CA. Because the Join protocol only needs to be executed once, the calculation of zero-knowledge proof has little effect on the efficiency of the entire protocol. We can choose the ‘‘Stern-type’’ protocol with the largest amount of calculation but the most accurate. CA first confirms the zero-knowledge proof and then uses MP-sampler algorithm to sample $s = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$ satisfying $[\mathbf{a}|\mathbf{b} + i \cdot [1\sqrt{q}]] \cdot \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = u + u_1$. Note that $i \in \mathcal{S}_q = \{c_0 + c_1 X^{d/2} \in \mathcal{R}_q | c_0, c_1 \in \mathbb{Z}_q\}$ instead of an arbitrary $i \in R_q$. Finally, CA sends the generated identity certificate (s, i) to the Host and the Host saves it as VID.

6.3. Create. The Create protocol generates $PScert$ for vehicles to send and receive messages in VANETs, including pseudonym key pairs, identity certificate sig_1 in revocation, and legal member certificates sig_2 . To generate pseudonym key pairs, TPM picks a basenane bsn and creates a value $d = H_{R_q}(bsn)$ as well as the pseudonym private key (e_1, e') , where e_1 is a part of the TPM secret value and $e' = H_{R_q}(sk, bsn)$. TPM outputs $\text{nym} = de_1 + e' \in R_q$ as pseudonym public key and creates $\text{sig}_1 = H_{R_q}(\text{nym}, \mathbf{e})$.

Using the *BlindSign* protocol in Table 3, TPM and Host jointly sign the message ‘‘certified’’ with TPM private

key \mathbf{e} and the pseudonymous private key (e_1, e') to generate a legal identity certificate sig_2 . *BlindSign* is a zero-knowledge proof of VID (s, i) completed by the Host and TPM interaction. That is, to prove that the Host has (s, i) satisfying $[\mathbf{a}|\mathbf{b} + i \cdot [1\sqrt{q}]] \cdot \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = u + \mathbf{a}_2 \cdot \mathbf{e}$. However, the verifier needs to know the value of $\mathbf{a}|\mathbf{b} + i \cdot [1\sqrt{q}]$ in the scheme, among which $[\mathbf{a}, \mathbf{b}, [1\sqrt{q}]]$ are all public parameters, so the identity can be easily deduced, and the user’s identity will be leaked. Therefore, the zero-knowledge proof is not directly performed on i , but the commitment value about i is first generated, and the zero-knowledge proof is generated by replacing i with the commitment value. Bring the commitment value into the trapdoor function to get the following:

$$[\mathbf{a}^T|\mathbf{b}^T + [t_2 t_2']|\mathbf{a}_2^T] \begin{bmatrix} s_1 \\ s_2 \\ \mathbf{e} - [\mathbf{r} \mathbf{r}'] s_2 \end{bmatrix} = u, \quad (2)$$

where $\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \text{Com}(i, \mathbf{r})$ and $\mathbf{t}' = \begin{bmatrix} t_1' \\ t_2' \end{bmatrix} = \text{Com}(i\delta, r')$. Suppose $\mathbf{v}^T = [\mathbf{a}^T|\mathbf{b}^T + [t_2 t_2']|\mathbf{a}_2^T]$ and $s' = \begin{bmatrix} s_1 \\ s_2 \\ \mathbf{e} - [\mathbf{r} \mathbf{r}'] s_2 \end{bmatrix}$,

then (2) can be expressed as $\mathbf{v}^T s' = u$.

In summary, the Host needs to generate three zero-knowledge proofs in parallel, that is,

π_1' : prove t, t' open to messages m, m' satisfying $m' = \delta m$

π_2' : prove t opens to message m satisfying $m = \sigma_5(m)$

π_3' : prove s' satisfying $\mathbf{v}^T s' = u$

TABLE 3: BlindSign.

TPM	Public: \mathbf{nym}, \mathbf{d} Private: e', \mathbf{e}_1	Host	Message: μ Public: $\sigma_5, \delta, \mathbf{v}^T, \mathbf{t}, \mathbf{t}'$ Private: $\mathbf{r}, \mathbf{r}', i$
$\mathbf{y}_{e_1}, \mathbf{y}_{e'} \leftarrow D_\xi$ $\mathbf{t} = \mathbf{d}\mathbf{y}_{e_1} + \mathbf{y}_{e'}$	\xrightarrow{t}	$y, y', y_5 \leftarrow D_\xi^3$ $y_{s_1} \leftarrow D_{\xi_1}^4$ $y_{s_2} \leftarrow D_{\xi_1}^4$ $y_s = (y_{s_1}, y_{s_2})$ $\pi_1': \mathbf{W}_1 = \mathbf{a}_1^T \mathbf{y}$ $\mathbf{W}'_1 = \mathbf{a}_1^T y'$ $\mathbf{W}_2 = \delta \mathbf{a}_2^T \mathbf{y} - \mathbf{a}_2^T y'$ $\pi_2': \mathbf{W}_{1,5} = \sigma_5^{-1} (\mathbf{a}_1^T) \mathbf{y}_5$ $\mathbf{W}_{2,5} = \mathbf{a}_2^T \mathbf{y} - \sigma_5^{-1} (\mathbf{a}_2^T) \mathbf{y}_5$ $\pi_3': \mathbf{W}_s = \mathbf{v}^T \mathbf{y}_s$	
$\mathbf{S}_{e_1} = \mathbf{y}_{e_1} + \mathbf{c}\mathbf{e}_1$ $\mathbf{S}_{e'} = \mathbf{y}_{e'} + \mathbf{c}\mathbf{e}'$ if $\text{rej}([\mathbf{S}_{e_1} \mathbf{S}_{e'}])$, $[\mathbf{c}\mathbf{e}_1, \mathbf{c}\mathbf{e}']$, $\xi) = 1$, rebort	\xleftarrow{c} $\xleftarrow{(S_{e_1}, S_{e'})}$	$c = \mathbf{H}(\delta, \sigma_5, \mathbf{W}_1, \mathbf{W}'_1, \mathbf{W}_2, \mathbf{W}_{1,5}, \mathbf{W}_{2,5}, \mathbf{W}_s, \mu, t, \mathbf{t}, \mathbf{t}', \mathbf{v})$ $\pi_1': \mathbf{z} = \mathbf{r}\mathbf{c} + \mathbf{y}$ $\mathbf{z}' = \mathbf{r}'\mathbf{c} + \mathbf{y}'$ $\pi_2': \mathbf{z}_5 = \sigma_5^{-1} (\mathbf{r})\mathbf{c} + \mathbf{y}_5$ $\pi_3': \mathbf{z}_s = \mathbf{s}\mathbf{c} + \mathbf{y}_s$ if $\text{rej}([\mathbf{z} \mathbf{z}' \mathbf{z}_5 \mathbf{z}_s])$, $[\mathbf{r}\mathbf{c}, \mathbf{r}'\mathbf{c}, \sigma_5^{-1} (\mathbf{r})\mathbf{c}, \mathbf{s}\mathbf{c}], \xi) = 1$, rebort	$(z, z', z_5, c, S_{e_1}, S_{e'})$

Finally, the identity credentials $\text{PSCert} = (\mathbf{nym} \parallel \text{sig}_1 \parallel \text{sig}_2)$ are generated and saved on the Host platform.

6.4. Sign/Verify. When the vehicle is moving in VANETs, the Host generates messages about the location and speed of the vehicle and transmits them to TPM. TPM signs messages using *Sign* protocol in Table 5 with pseudonym private key (e_1, e') and pseudonym public key $\mathbf{nym} = \mathbf{d}\mathbf{e}_1 + e' \in R_q$ and returns m_{sign} to Host. The Host creates $\text{msg} = \{m_{\text{plain}} \parallel m_{\text{sign}} \parallel \text{PSCert}\}$ and sends it to the receiver. After receiving msg , the receiver first calls *BlindVerify* Protocol as in Table 4 to verify sig_2 , confirming that the message comes from a legal user. Then use pseudonym public key \mathbf{nym} to verify m_{sign} as in Table 5.

6.5. Revoke. The revocation instruction $\text{msg} = \{\text{revoke} \parallel \mathbf{nym} \parallel \text{reason}\}$ generated by RA is encrypted with the RA private key sk_{ra} and broadcast in VANETs so that all legitimate users can receive it. After receiving the message, the Host passes it on to TPM. TPM uses RA public key pk_{ra} to decrypt msg and recognizes that the target of the instruction is itself according to \mathbf{nym} . Then TPM creates $\text{sig}_1^{ra} = H_{R_1}(\mathbf{nym}, \mathbf{e})$ and calls *BlindSign* to generate sig_2^{ra} on message “confirm,” which is used to prove that TPM has received the revocation instruction and completed the self-revocation. After that, TPM deletes its own public and private key pairs and all identity certificates independently. The Host sends $\text{sig}^{rvk} = \{\text{sig}_1^{ra} \parallel \text{sig}_2^{ra}\}$ to RA. Since RA has knowledge of the misbehaving vehicle’s PSCert , it checks whether $\text{sig}_1 = \text{sig}_1^{ra}$ and guarantees that the target vehicle has been revoked. Then, RA calls *BlindVerify* to confirm that sig^{rvk} is indeed issued by the revoked vehicle.

It can be seen from the entire revocation process that RA can correctly revoke the target vehicle without any pseudonym resolution operations. The vehicle provides RA with proof that the identity certificate has been forcibly revoked by TPM. If the vehicle wants to communicate with the users in VANETs again, it must rerun the Join phase to generate a new identity certificate.

7. Security Analysis

The security comparison between V-LDAA, the lattice-based ring signature schemes in [7, 8, 20], and the VPKE scheme in [1] are shown in Table 6. Compared with lattice-based ring signatures in VANETs, V-LDAA has the advantage of achieving user-controlled unlinkability and unforgeability. In contrast to the existing VPKE scheme, V-LDAA achieves postquantum security and realizes the user’s independent pseudonym update scheme and the distributed vehicle certificate revocation scheme.

7.1. Unforgeability. Suppose CA public parameters are set as follows: $[\mathbf{a} | \mathbf{b}], u, \mathbf{a}_2$, where $i^* \in \mathcal{S}_q$, $\mathbf{R} \in R_1^{2 \times 2}$, $R' \leftarrow R^{4 \times 2}$, $\mathbf{s}_u \in D_\sigma$, $\mathbf{g} = [1 \sqrt{q}]$, $\mathbf{b} = \mathbf{a} \cdot \mathbf{R} - i^* \mathbf{g}$, $\mathbf{a}_2 = [\mathbf{a} | \mathbf{a}\mathbf{R}] \cdot \mathbf{R}'$, $u = [\mathbf{a} | \mathbf{a}\mathbf{R}] \cdot \mathbf{s}_u$.

Suppose we have a fake sampling algorithm. The adversary chooses the identity $i \in \mathcal{S}_q$ and secret value \mathbf{e} . When $i \neq i^*$, use the original MP – Sampler to generate \mathbf{s} satisfying $[\mathbf{a} | \mathbf{a}\mathbf{R} + [i - i^*] \mathbf{g}] \cdot \mathbf{s} = u + \mathbf{a}_2 \cdot \mathbf{e}$ and output \mathbf{s} to the adversary. When $i = i^*$, the gadget matrix vanishes and $[\mathbf{a} | \mathbf{a}\mathbf{R}] \cdot \mathbf{s} = u + \mathbf{a}_2 \cdot \mathbf{e}$. Therefore, compute $\mathbf{s}^* = \mathbf{s}_u + R' \mathbf{e}^*$, which is also a valid signature and output \mathbf{s}^* to the adversary and the adversary verifies $[\mathbf{a} | \mathbf{a}\mathbf{R}] \mathbf{s}^* = u + \mathbf{a}_2 \cdot \mathbf{e}^*$. According to [11], based on Ring-LWE and NTRU assumptions, the adversary cannot distinguish whether it is generated by the

TABLE 4: BlindVerify.

Host	Verifier
$(z, z', z_5, z_e, c, S_{e1}, S_e')$	$\ (z, z', z_5, z_e, S_{e1}, S_e')\ \leq \beta_z,$
	$\pi_1': \mathbf{W}'_1 = \mathbf{a}'_1^T \mathbf{z} - t_1 c$
	$\mathbf{W}'_1 = \mathbf{a}'_1^T z' - t_1' c$
	$\mathbf{W}'_2 = \delta \mathbf{a}'_2^T \mathbf{z} - \mathbf{a}'_2^T z' - (\delta t_2 - t_2') c$
	$\pi_2': \mathbf{W}'_{1,5} = \sigma_5^{-1} (\mathbf{a}'_1^T) \mathbf{z}_5 - \sigma_5^{-1} (t_1) c$
	$\mathbf{W}'_{2,5} = \mathbf{a}'_2^T \mathbf{z} - \sigma_5^{-1} (\mathbf{a}'_2^T) \mathbf{z}_5 - (t_2 - \sigma_5^{-1} (t_2')) c$
	$\pi_3': \mathbf{W}'_s = \mathbf{v}^T \mathbf{z}_s - uc$
	$t = \mathbf{d} S_{e1} + S_e' - \text{cnym}$
	if $c = \mathbf{H}(\delta, \sigma_5, \mathbf{W}'_1, \mathbf{W}'_1, \mathbf{W}'_2, \mathbf{W}'_{1,5}, \mathbf{W}'_{2,5}, \mathbf{W}'_s, \mu, t, \mathbf{t}, t', \mathbf{v})$
	Output 1
	else
	Output 0

TABLE 5: Sign and verify.

TPM	Public: $\mathbf{a}, \mathbf{y} = \mathbf{a}\mathbf{s} + \mathbf{e}$ Private: \mathbf{s}, \mathbf{e}	Verifier Public: $\mathbf{y}, \mathbf{a}, \beta_z$
$\mathbf{r}_s, \mathbf{r}_e \leftarrow D_\xi, \mathbf{t} = \mathbf{a}\mathbf{r}_s + \mathbf{r}_e$		
$c = \mathbf{H}(\mathbf{t}, \mu)$		
$\mathbf{z}_s = \mathbf{c}\mathbf{s} + \mathbf{r}_s$		
$\mathbf{z}_e = \mathbf{c}\mathbf{e} + \mathbf{r}_e$		
if $\text{rej}([\mathbf{z}_s, \mathbf{z}_e], [\mathbf{c}\mathbf{s}, \mathbf{c}\mathbf{e}], \xi) = 1$, abort	(t, c, z_{s1}, z_e)	$\mathbf{t} = \mathbf{a}\mathbf{z}_s + \mathbf{z}_e - c\mathbf{y}$ if $\ \mathbf{z}_s\ , \ \mathbf{z}_e\ \leq \beta_z$ and $c = \mathbf{H}(\mathbf{t}, \mu)$
		Output 1 else Output 0

real public parameters and the real preimage sampling algorithm or generated by the above public parameters and the fake preimage sampling algorithm.

According to the above conclusion, we can prove the unforgeability of the V-LDAA signature.

During BlindSign, the Host needs to generate a zero-knowledge proof about \mathbf{r}, \mathbf{r}' such that

$$\begin{bmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \end{bmatrix} \cdot [\mathbf{r} \mathbf{r}'] + \begin{bmatrix} 0 & 0 \\ ci & ci\sqrt{q} \end{bmatrix} = c \cdot \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}. \quad (3)$$

In parallel, it will also prove that

$$[\mathbf{a}|\mathbf{b} + [t_{21} \ t_{22}]] \cdot \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} = cu + \mathbf{a}_2 \cdot \mathbf{e} + \mathbf{a}'_2 \cdot \tilde{\mathbf{r}}. \quad (4)$$

Combine (3) and (4) to get the following:

$$c[\mathbf{a}|\mathbf{b} + i \cdot \mathbf{g}] \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} + \mathbf{a}'_2 \tilde{\mathbf{r}} = c^2 u + c\mathbf{a}_2 \cdot \mathbf{e} + c\mathbf{a}'_2 \cdot [\mathbf{r} \mathbf{r}'] \mathbf{s}_2. \quad (5)$$

The adversary randomly selects $i \in \mathcal{S}_q$, and the probability of selecting $i = i^*$ is $1/q^2$. At this time i is vanished, that is,

$$c[\mathbf{a}|\mathbf{a} \cdot \mathbf{R}] \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} + \mathbf{a}'_2 \tilde{\mathbf{r}} = c^2 u + c\mathbf{a}_2 \cdot \mathbf{e} + c\mathbf{a}'_2 \cdot [\mathbf{r} \mathbf{r}'] \mathbf{s}_2. \quad (6)$$

Sampling algorithm outputs \mathbf{e}^* satisfying the following:

$$c^2 [\mathbf{a}|\mathbf{b}] \begin{bmatrix} \mathbf{s}'_1 \\ \mathbf{s}'_2 \end{bmatrix} = c^2 (u + \mathbf{a}_2 \cdot \mathbf{e}^*). \quad (7)$$

Subtract (6) and (7) to get the following:

$$\mathbf{a}(c\mathbf{s}_1 - c^2 s' + c\mathbf{R}\mathbf{s}_2 - c^2 \mathbf{R}\mathbf{s}'_2) + \mathbf{a}_2(c^2 \mathbf{e}^* - c\mathbf{e}) + \mathbf{a}'_2[\tilde{\mathbf{r}} - c[\mathbf{r} \mathbf{r}'] \mathbf{s}_2] = 0, \quad (8)$$

which can be written as follows:

$$[\mathbf{a}|\mathbf{a}\mathbf{R}|\mathbf{a}_2|\mathbf{a}'_2] \cdot \begin{bmatrix} c\mathbf{s}_1 - c^2 s' \\ c\mathbf{s}_2 - c^2 s'_2 \\ c^2 \mathbf{e}^* - c\mathbf{e} \\ \tilde{\mathbf{r}} - [\mathbf{r} \mathbf{r}'] c\mathbf{s}_2 \end{bmatrix} = 0. \quad (9)$$

Because $\mathbf{s}_1, s', c, \mathbf{e}, \mathbf{e}^*, \mathbf{s}_2$ are all polynomials with small coefficients, (9) is a nonzero Ring-SIS solution to $[\mathbf{a}|\mathbf{a}\mathbf{R}|\mathbf{a}_2|\mathbf{a}'_2]$ unless all multiplicands are 0. Therefore, if the adversary can successfully generate a zero-knowledge proof that satisfies (9), the Ring-SIS problem can be solved with a probability of $1/q^2$. To generate a zero solution, it requires $c^2 \mathbf{e}^* - c\mathbf{e} = 0$. That is, $c\mathbf{e}^* = \mathbf{e}$, which means every \mathbf{e} extracted from the zero-knowledge proof sig_2 in the blind signing phase must be equal to a certain $c\mathbf{e}^*$, where \mathbf{e}^* is a TPM secret value of a legal certificate VID generated in the Join phase. So far, the unforgeability of the V-LDAA signature can be proved. If the adversary wants to break the unforgeability, the difficulty of using the secret value of a platform

TABLE 6: Security comparison.

Security requirement	ECPB in [20]	DAPRS in [7]	LRMA in [8]	Scheme in [1]	V-LDAA
Anonymity	✓	✓	✓	✓	✓
Spontaneity	✓	✓	✓	✓	✓
Unforgeability	✗	✓	✓	✓	✓
Postquantum security	✗	✓	✓	✗	✓
User-controlled unlinkability	✗	✗	✗	✗	✓
Distributed revocation mechanism	—	—	—	✗	✓
Pseudonym update spontaneity	—	—	—	✗	✓

without a legal identity certificate to generate a legal signature can be reduced to solve the Ring-SIS problem.

7.2. Anonymity. Anonymity means the adversary cannot extract the user identity value i from the signature. Suppose the adversary knows the TPM private key sk_1, sk_2 and outputs the message \mathbf{m}^* to be signed and two identity values i_1, i_2 to the challenger. The challenger randomly selects an identity value i to sign and returns the signature to the adversary. After receiving the signature, the adversary guesses whether the identity value chosen by the challenger is i_1 or i_2 . According to [19], the commitment scheme used in this article has hiding property based on the difficulty of M-LWE. That is, the adversary cannot distinguish the commitment value of two different messages. When signing, the challenger can replace the identity value at will to calculate the commitment value, and the generated signature is completely independent of the identity value i , so the difficulty of the adversary's guessing the id value used from the blind signature can be reduced to the M-LWE problem. In VANETs, the identity certificate generated in the Create stage only contains pseudonym information and does not contain any real identity information, and the TPM signing key cannot be associated with the vehicle user, so the adversary cannot distinguish different vehicles from the signature unless the user reveals his or her identity information.

7.3. User-Controlled Unlinkability. During Create protocol, the user can choose whether to use the same secret key sk to generate the same or different pseudonym private key so as to control whether the generated signature is linked. Once a different pseudonym is selected, the adversary cannot determine whether the two signatures are from the same user. Since sig_1 is generated by hashing the TPM private key and the pseudonym private key, the adversary cannot determine which TPM private key is used. In addition, sig_2 is a blind signature and cannot be linked.

7.4. Unforgeability of Revocation Instruction. In order to prevent the adversary from maliciously revoking the legal vehicle, it should be ensured that the revocation instruction received by TPM is from the real RA and not forged by an adversary. Adding the signature of RA to each revocation instruction can meet this requirement. TPM can confirm the authenticity of the revocation instruction by verifying the RA signature.

7.5. Unforgeability of Revocation Certificate. When RA receives the revocation certificate returned by TPM, RA must ensure that it is from the correct target vehicle and has honestly performed certificate and key destruction operations. In V-LDAA, the credibility of the revocation operation is guaranteed by the trusted hardware chip TPM. By comparing the signatures in the revocation certificate, RA can confirm that the target vehicle has indeed performed the revocation operation. No other user can forge this signature as long as the TPM key is not leaked.

7.6. Consistency of Revocation Operation. When the revocation instruction is correctly delivered to TPM, TPM will perform a series of destruction operations. However, the revocation instruction needs to be passed through the Host. If the Host is controlled by an adversary and maliciously intercepts the transmission of the revocation instruction, TPM cannot receive the correct information from RA and cannot complete the revocation operation, which is a major challenge in the distributed revocation mechanism. In V-LDAA, TPM receives information from RA at fixed time intervals which include time stamps and RA's signature. If TPM stops receiving the time stamp information, it is considered that the communication between TPM and RA interferes, and corresponding countermeasures should be taken.

8. Experimental Results and Analysis

We compare the performance of V-LDAA from two aspects: theoretical analysis and experimental simulation. Firstly, we compare V-LDAA with existing lattice-based authentication schemes in VANETs in Section 8.1 to measure the advantages of V-LDAA in the scenario of the Internet of Vehicles. Secondly, the BlindSign protocol in V-LDAA is compared with that in existing LDAA in Section 8.2 to highlight the improvement of computing efficiency after adopting the optimized signature scheme as presented in Section 4.2. This article uses Python language and SageMath9.2 library to simulate V-LDAA, LRMA in [8], DAPRS in [7], and LDAA in [11], in which the polynomial multiplication is accelerated by the NTL library. Based on the Intel(R) Core (TM) i5-7500 CPU @3.40 GHz memory 8 GB processor, we tested the execution time and signature size of each scheme.

8.1. Comparison with Existing Lattice-Based Authentication Schemes in VANETs. We compare the proposed V-LDAA scheme with existing lattice-based authentication schemes in

TABLE 7: Comparison of costs.

	Sign	Verify	Signature size
Scheme in [21]	$mT_{\text{samp}} + m(N+1)T_{\text{mult}}$	$M(N+2)T_{\text{mult}}$	$(N+2)m$
Scheme in [6]	$5NT_{\text{mult}}$	$T_N + 5NT_{\text{mult}}$	$2(N+1)m$
DAPRS in [7]	$2NT_{\text{mult}}$	$2NT_{\text{mult}}$	$(N+1)m$
LRMA in [8]	NT_{mult}	T_{mult}	$(N+1)m$
V-LDAA	$2T_{\text{mult}}$	$T_N + 2T_{\text{mult}}$	$3m + \text{PSCert}$

```

Verification pass
Scheme: V-LDAA
degree= 128
q= 114356107
ParamGen Running time: 7.452999999999999 Seconds
BlindSign Running time: 5.922000000000001 Seconds
BlindVerify Running time: 0.0470000000000006 Seconds
Sign Running time: 0.03100000000000236 Seconds
Verify Running time: 0.0470000000000006 Seconds
/opt/sagemath-9.2/local/lib/python3.7/site-packages/sage/repl/ipython_kernel/__main__.py

```

FIGURE 3: V-LDAA protocol experimental results.

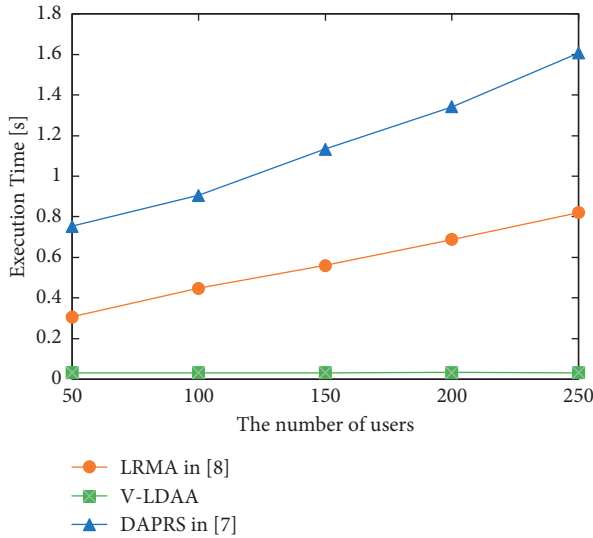


FIGURE 4: Signing performance for V-LDAA, DAPRS in [7], and LRMA in [8].

VANETs. Assuming that the time for a preimage sampling is T_{samp} , the time for a polynomial multiplication is T_{mult} , and the time for a zero-knowledge proof is T_N . The signing and verification calculation costs and signature length of each scheme are shown in Table 7. The studies in [6–8, 21] are all lattice-based ring signature schemes. In the ring signature scheme, users need to use their private key and all other users' public keys to sign messages. For a ring with numerous users, that is, when N is large, the computation burden is considerable. In addition, the members in a ring change as vehicles move. Thus, the member public key also needs to be updated consequently. However, in V-LDAA, users only need to sign with their pseudonym private keys each time, regardless of N .

The experimental results are shown in Figures 3–6. We implement *Sign*, *Verify*, *BlindSign*, and *BlindVerify* protocols

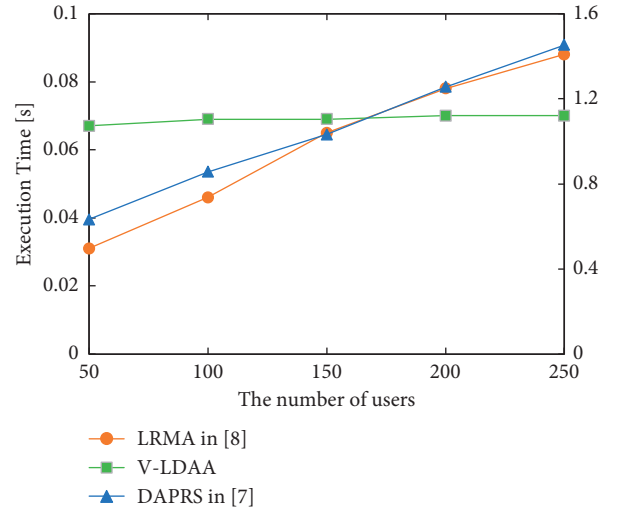


FIGURE 5: Verification performance for V-LDAA, DAPRS in [7], and LRMA in [8].

and measure the running time. The results are shown in Figure 3. The execution time is averaged after 10 runs of each protocol. We also compare the V-LDAA scheme with DAPRS in [7] and LRMA in [8]. A lattice-based double-authentication-preventing ring signature (DAPRS) is introduced in [7] using double-authentication-preventing signatures (DAPRSs) instead of conventional signatures. A lattice-based ring signature scheme for message authentication (LRMA) is presented in [8], providing unconditional privacy to vehicles. The number of users N varies from 50 to 200. The degree of cyclotomic $d=128$, and $q=114356107$. Since *BlindSign* protocol is called only when users want to update their pseudonyms and recreate *PSCert*, we ignore the cost of *BlindSign*. In Figure 4, the signing time required for LRMA and DAPRS increases tremendously as the number of users rises, while in V-LDAA the execution time in signing

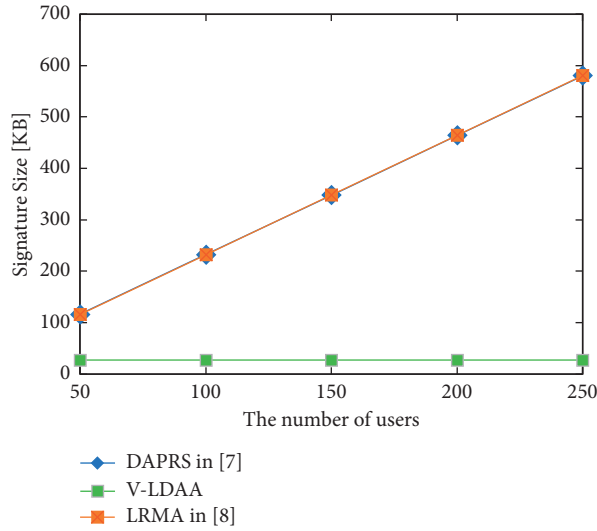


FIGURE 6: Signature size for V-LDAA, DAPRS in [7], and LRMA in [8].

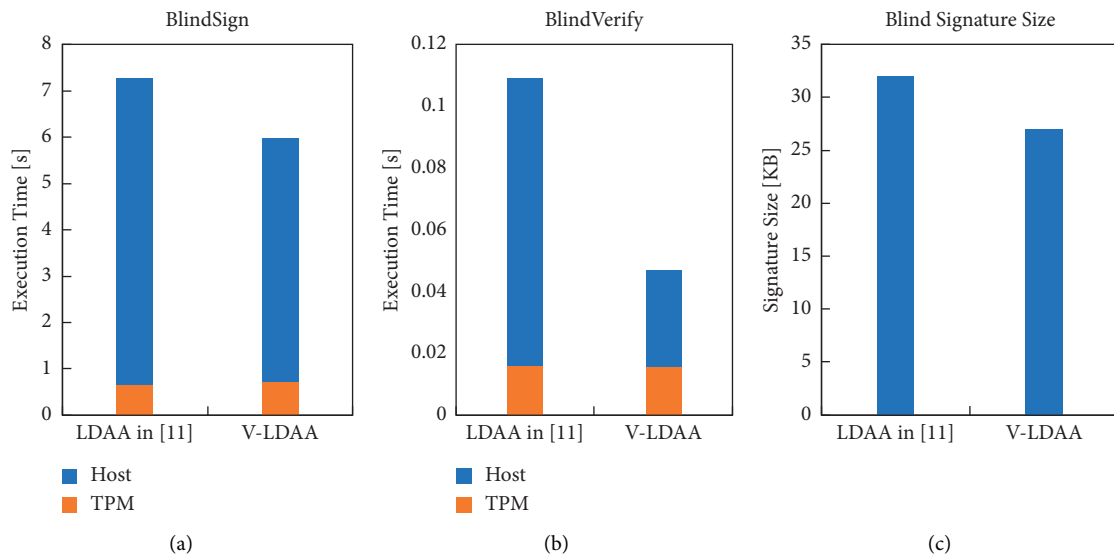


FIGURE 7: Comparison among V-LDAA and LDAA in [12] in terms of signing execution time (a), verification execution time (b), and signature size (c).

operations maintains at a low level with slight fluctuations. In Figure 5, additional verification of PScert is required in V-LDAA, so the verification execution time is longer than LRMA when N is small but is exceeded as N increases. The size of the certificate generated by V-LDAA is significantly smaller than that of LRMA and DAPRS, as shown in Figure 6, and it will not increase with the growth of the number of users.

8.2. Comparison with the Existing LDAA Scheme. We compare the performance of the proposed V-LDAA protocol with the existing LDAA protocol in [11] during the blind signing and blind verification on computation and storage resource consumption. In the blind signing phase,

V-LDAA adopts an optimized signature scheme which removes the proof for $\sigma_5(m) = m$ and thus reduces the number of response values to the challenge, so the number of polynomials that the generated signature contains is reduced from 40 in [11] to 36.

In the Joining phase, V-LDAA adds public and secret values to the Host and enables the Host's secret value to participate in the generation of the identity certificate. This change encourages TPM and Host to interact in the Joining phase to generate a zero-knowledge proof of their respective secret values. Although the amount of calculation is increased, considering that the long-term identity certificate of each legal user only needs to be generated once, it has little effect on the overall computing efficiency. In VANETs, the participation of TPM and Host in the generation of identity

certificates can effectively resist TPM chip theft attacks and prevent TPM from being transplanted to a new vehicle platform by the adversary and signed with the replaced identity certificate. The experimental results are shown in Figure 7, where $d = 128$, $\beta = 128$, and $q = 114356107$.

As shown in Figure 7(a), the speed of the Host blind signing operation is increased by 30% by reducing the number of proofs for automorphism stability. The Host operation during blind verification is accelerated 3 times, according to Figure 7(b). Also, V-LDAA reduces the signature size by 18%, as in Figure 7(c).

9. Conclusion

To solve the security and user privacy issues in VANETs, we propose a lattice-based direct anonymous attestation scheme in VANETs that achieves postquantum security. We introduce a lattice-based long-term certificate generation mechanism, a pseudonym certificate renewal mechanism, and a distributed certificate revocation mechanism. Users can update the pseudonym certificate by themselves and control the linkability of signatures. RA does not need to perform pseudonym resolution or maintain CRLs, which overcomes the shortcomings of the traditional VPKIs. We also demonstrate that V-LDAA has significant advantages in computing efficiency and storage consumption compared with the existing lattice-based direct anonymous attestation by adopting an optimized signature scheme based on automorphism stability. Experimental results show that V-LDAA reduces the signature size by 18%. And the speed of blind signing is increased by 30% and blind verification operations are accelerated 3 times compared with the existing LDAA scheme. The main shortcoming of the proposed V-LDAA scheme is the computation and storage costs in the BlindSign protocol. In future work, we will aim to further optimize the proposed scheme to make it more suitable for resource-constrained TPM chips and vehicle platforms.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Key Research and Development Program of China, Joint Research of IoT Security System and Key Technologies Based on Quantum Key (2020YFE0200600).

References

- [1] M. Asghar, R. R. M. Doss, and L. Pan, "A scalable and efficient PKI based authentication protocol for VANETs," in *Proceedings of the 28th International Telecommunication Networks and Applications Conference (ITNAC)*, pp. 1–3, IEEE, Sydney, Australia, November 2018.
- [2] P. Vijayakumar, M. Azees, V. Chang, J. Deborah, and B. Balusamy, "Computationally efficient privacy preserving authentication and key distribution techniques for vehicular ad hoc networks," *Cluster Computing*, vol. 20, no. 3, pp. 2439–2450, 2017.
- [3] D. He, S. Zeadally, B. Xu, and X. Huang, "An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2681–2691, 2015.
- [4] S. F. Tzeng, S. J. Horng, T. Li, X. Wang et al., "Enhancing security and privacy for identity-based batch verification scheme in VANETs," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3235–3248, 2015.
- [5] I. Ullah, A. Wahid, M. A. Shah et al., "VBPC: velocity based pseudonym changing strategy to protect location privacy of vehicles in VANET," in *Proceedings of the 2017 International Conference on Communication Technologies (ComTech)*, pp. 132–137, IEEE, Rawalpindi, Pakistan, April 2017.
- [6] Y. Cui, L. Cao, X. Zhang, and G. Zeng, "Ring signature based on lattice and VANET privacy preservation," *Chinese Journal of Computers*, vol. 42, no. 51, pp. 1–14, 2017.
- [7] J. Liu, Y. Yu, J. Jia et al., "Lattice-based double-authentication-preventing ring signature for security and privacy in vehicular Ad-Hoc networks," *Tsinghua Science and Technology*, vol. 24, no. 5, pp. 575–584, 2019.
- [8] P. Mundhe, V. K. Yadav, S. Verma, and S. Venkatesan, "Efficient lattice-based ring signature for message authentication in VANETs," *IEEE Systems Journal*, vol. 14, no. 4, pp. 5463–5474, 2020.
- [9] R. E. Bansarkhani and A. E. Kaafarani, "Direct anonymous attestation from lattices," *Cryptology ePrint Archive*, Report 2017/1022, 2017.
- [10] N. El Kassem, L. Chen, R. El Bansarkhani et al., "More efficient, provably-secure direct anonymous attestation from lattices," *Future Generation Computer Systems*, vol. 99, pp. 425–458, 2019.
- [11] L. Chen, N. Kassem, A. Lehmann et al., "A framework for efficient lattice-based daa," in *Proceedings of the 1st ACM Workshop on Workshop on Cyber-Security Arms Race*, pp. 23–34, London, UK, November 2019.
- [12] D. Micciancio and C. Peikert, "Trapdoors for lattices: simpler, tighter, faster, smaller," in *Advances in Cryptology—EUROCRYPT 2012*, pp. 700–718, Springer, Berlin, Germany, 2012.
- [13] C. Gentry, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, pp. 197–206, Victoria Canada, May 2008.
- [14] C. Baum, I. Damgård, V. Lyubashevsky, S. Oechsner, and C. Peikert, "More efficient commitments from structured lattice assumptions," in *Lecture Notes in Computer Science*, pp. 368–385, Springer, Cham, Berlin, Germany, 2018.
- [15] V. Lyubashevsky, "Lattice signatures without trapdoors," in *Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 738–755, Springer, Cambridge, UK, April 2012.
- [16] X. Boyen, "Lattice mixing and vanishing trapdoors: a framework for fully secure short signatures and more," in *Proceedings of the International Workshop on Public Key Cryptography*, pp. 499–517, Springer, Paris, France, April 2010.

- [17] S. Katsumata and S. Yamada, "Partitioning via non-linear polynomial functions: more compact IBEs from ideal lattices and bilinear maps," in *Advances in Cryptology—ASIACRYPT 2016*, pp. 682–712, Springer, Berlin, Germany, 2016.
- [18] B. Libert, S. Ling, K. Nguyen, and H. Wang, "Zero-knowledge arguments for lattice-based accumulators: logarithmic-size ring signatures and group signatures without trapdoors," in *Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 1–31, Springer, Vienna, Austria, May 2016.
- [19] R. Del Pino, V. Lyubashevsky, and G. Seiler, "Lattice-based group signatures and zero-knowledge proofs of automorphism stability," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 574–591, Toronto Canada, October 2018.
- [20] Y. Wang, H. Zhong, Y. Xu, and J. Cui, "ECPB: efficient conditional privacy-preserving authentication scheme supporting batch verification for VANETs," *IJ Network Security*, vol. 18, no. 2, pp. 374–382, 2016.
- [21] M. M. Tian, L. S. Huang, and W. Yang, "Efficient lattice-based ring signature scheme," *Chinese Journal of Computers*, vol. 35, no. 4, pp. 712–718, 2012.

Research Article

A LoRa-Based Lightweight Secure Access Enhancement System

Yu Jiang ^{1,2,3}, Hua Fu,^{1,3} Aiqun Hu,^{1,3} and Wen Sun¹

¹School of Cyber Science and Engineering, Southeast University, Nanjing, China

²Key Laboratory of Computer Network Technology of Jiangsu Province, Nanjing, China

³Purple Mountain Laboratories for Network and Communication Security, Nanjing, China

Correspondence should be addressed to Yu Jiang; jiangyu@seu.edu.cn

Received 9 June 2021; Revised 28 July 2021; Accepted 16 August 2021; Published 26 August 2021

Academic Editor: Jingyu Feng

Copyright © 2021 Yu Jiang 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 access control mechanism in LoRa has been proven to have high security risks. In order to improve the secure access ability of LoRa terminals, this paper presents a physical layer-based authentication system for security enhancement. Different from the security access technology of cryptography, a lightweight gateway architecture called LW-LoRaWAN is proposed to realize a data frame-based authentication with radio frequency fingerprint (RFF). A novel RFF feature of Cross Power Spectral Density (CPSD) is used to achieve a fast authentication with one single frame. Theoretical analysis and experimental results show that the proposed system not only reinforces the authentication security of LoRa network but also protects the LoRa terminals against the Sybil attacks. The LW-LoRaWAN provides new security approach from physical layer for LoRa network.

1. Introduction

The Internet of things (IoT) is related to distributed control, dynamic topology, and limited power of terminal nodes. These characteristics of IoT make the security threats different from those of the Internet [1–4]. Due to the openness of wireless communication, the IoT is more vulnerable to eavesdropping, counterfeiting, tampering, and denial-of-service attacks. When nodes are compromised by attackers, they can still access the network with their legitimate identities, making IoT networks exposed to both external and internal threats. Due to the frequent accessing and exiting of nodes and the dynamic changes of network topology, it is impossible to effectively prevent the attack of illegal nodes without a reliable secure access mechanism [5–7].

At present, the network structure of IoT system has not been clearly defined. It is generally considered that the network structure includes the perception layer, the transport layer, and the application layer [8]. There have been mature solutions for security management of the transport layer and application layer, but the research on the identity authentication technology of perception layer is not enough

[9]. Common authentication techniques include authentication based on MAC address, security certificates, instructions, and so on. Unfortunately, all of these methods have certain risks in practical scenarios. MAC address-based authentication can easily be cracked by forging the MAC address, and the black and white list strategy is bypassed [10]. Certificate-based authentication is limited by the lack of resources such as memory or computing ability of IoT terminals [11]. Instruction-based authentication suffers from the instruction leakage and weak instructions, which makes the terminal unable to be effectively protected [12].

The security issues of IoT are different from those of the Internet, because complex security policies cannot be deployed with the limited resources of IoT devices. The problem of accessing security for the IoT has been studied for more than a decade, but it remains challenging to find an effective solution satisfying both lightweight and security.

In order to solve the identity authentication problem of the IoT, it is necessary to propose a complete secure access solution based on the limited resources of IoT devices. In fact, the signal emitted by wireless device carries the unique features on the waveform, which can be deployed as the identity of the terminal to distinguish the counterfeiters

[13–15]. Compared with traditional approaches based on MAC address or authentication instructions, the physical layer features of devices, also known as radio frequency fingerprint (RFF), are difficult to forge and can be applied for identity authentication [16–18].

LoRa is a low-power wireless communication technology, which uses Chirp Spread Spectrum (CSS) to serve remote communication [19]. According to the needs of different scenarios, the physical layer of LoRa is highly configurable, including spread factor (SF), coding rate (CR), bandwidth (BW), optional header, and other parameters. LoRaWAN is an open-source protocol of LoRa, which can realize effective communication and networking between LoRa terminals and gateway.

As the basic technology of Low-Power Wide-Area Network (LPWAN) communication, LoRa is expected to be widely used. Meanwhile, due to the use of unauthorized frequency bands and public protocols, LoRa network is vulnerable to attacks. At present, the analysis of access authentication of LoRa terminals is mostly based on LoRaWAN protocol [20, 21], and the research of access authentication based on the RFF of physical layer is still in the early stage. This paper firstly proposes a physical layer-based authentication system based on RFF features to improve the security of access for LoRa terminals. The main contributions of the proposed authentication solution for LoRa are listed below:

- (1) A lightweight gateway architecture called LW-LoRaWAN is proposed to realize a data frame-based authentication with RFF
- (2) A novel RFF feature of Cross Power Spectral Density (CPSD) is used to achieve a fast authentication within one data frame
- (3) LW-LoRaWAN can protect the LoRa terminals against the Sybil attacks
- (4) The proposed system for security enhancement only needs to upgrade the gateway; no change is needed for terminal devices, which is more feasible than the existing enhancement schemes

The remainder of this paper is organized as follows: In Section 2, the state of the art of the lightweight security enhancement techniques for IoT is presented. The lightweight secure access scheme for LoRa is proposed in Section 3. The novel RFF extraction method for single data frame is presented in Section 4. Experimental results and system analysis are presented in Section 5. Finally, the conclusion is drawn in Section 6.

2. Background and Related Work

Due to the limited resources of IoT devices, the research of the IoT security focuses on the lightweight access technologies, where the lightweight cryptographic algorithm and the lightweight security protocol are the two main research directions. Meanwhile, the RFF-based authentication has been widely studied, which provides a different approach from modern cryptography.

2.1. Lightweight Cryptographic Algorithm. Lightweight cryptographic algorithm [22] is an innovative algorithm for devices with limited computing resources. In recent years, with the increasing security requirements for IoT, the research on lightweight cryptographic algorithm has achieved good results. The International Organization for Standardization (ISO) has developed some standards for algorithms such as lightweight block ciphers and stream ciphers, most of which are symmetric.

Lightweight cipher mainly includes lightweight block cipher, lightweight Hash function, and lightweight stream cipher. Among the symmetric ciphers, the block cipher algorithm [23] has been studied for a long time and has many achievements. It has typical security structures, such as Feistel and Substitution Permutation Network (SPN). After the PRESENT algorithm was published, many lightweight Hash functions have been designed based on PRESENT algorithm, such as C-PRESENT, H-PRESENT, and DM-PRESENT [24]. In recent years, there has been a new trend to design Hash functions by using Sponge structure [25]. The design of stream ciphers is mainly based on the linear and nonlinear feedback shift registers. A variety of lightweight stream cipher algorithms have been proposed such as Espresso, Lizard, Grain-128a, Welch Gong 8 (WG-8), Sprout, Plantlet, and Fruit [26].

2.2. Lightweight Security Protocol. The perception layer of the IoT cannot meet the requirements of computing, storage, and communication overhead of traditional security protocols, so it is necessary to research and develop lightweight protocols. In general, lightweight security protocols, which mainly include lightweight authentication protocol, lightweight key agreement protocol, and lightweight key management protocol, are designed to reduce the amount of computation, information flow, and number of communication rounds by sacrificing certain reliability and even security [27].

The lightweight authentication protocols are mainly used in resource-constrained system to ensure the legitimacy of the identities. It includes one-way and two-way authentication and can be widely used in point-to-point and multihop communications. After the identity authentication is completed, the lightweight key negotiation protocol establishes a session key for the subsequent communication. It can be widely used in access control of RFID, IoT, and other systems [28]. The lightweight key management protocol is used to create, distribute, and maintain the key in the cryptographic mechanism of resource-constrained system and to realize the key management in the secure communication.

2.3. LoRa RFF Technology. As a new wireless communication technology, LoRa RFF begins to receive the attention of researchers. The radar signal has the same modulation scheme as the LoRa signal and the identification of radar equipment is called Specific Emitter Identification (SEI). The U.S. Naval Research Bureau has conducted research on SEI technology for more than decades [29]. The purpose of SEI

research is to extract weak and robust features from radar signals to achieve individual identification of emission sources. Due to the sensitivity of radar research, the information available for inquiry is very limited. In addition, radar is mainly used in the military field, so its production accuracy is better than the commercial LoRa devices. The identification of LoRa devices should be based on its practical scenarios and device characteristics.

In recent years, the research on the communication protocol of LoRa devices [30–32] and the synchronization of LoRa signals [33, 34] are in progress. The identification of LoRa devices based on RFF has also been carried out. In 2017, Eletreby et al. [35] proposed using the time, frequency, and phase offset of the signal to identify LoRa devices and applied them to the access authentication process of LoRaWAN network. Robyns et al. [36] proposed a supervised machine learning method to recognize LoRa devices. It takes the data after signal preprocessing as the recognition object for machine learning. In 2019, Jiang et al. [37] extracted the RFF features of LoRa device based on the differential constellation trace figure (DCTF).

3. Lightweight Secure Access Scheme for LoRa

The current secure access scheme of LoRaWAN protocol adopts modern cryptography technology. This section proposes a physical layer secure access control scheme based on RFF of LoRa terminals and establishes a lightweight access protection architecture, referred to as LW-LoRaWAN.

3.1. Overall Architecture. The proposed LW-LoRaWAN system architecture includes four parts: LoRa terminals, LoRa gateway, RFF database, and remote server. LoRa terminals and remote server continue to use the equipment in LoRaWAN, while LoRa gateway and RFF database are new equipment in LW-LoRaWAN. The system architecture is shown in Figure 1.

Maintaining the original functions of LoRaWAN, LW-LoRaWAN provides the functions listed in the following:

- (1) LoRa modulation and demodulation
- (2) LoRaWAN protocol support
- (3) Bidirectional communication with remote server
- (4) RFF extraction and identification of LoRa terminals
- (5) Establishing the relevance between RFF and data frame of LoRa terminals
- (6) Real-time illegal data blocking and abnormal terminal alarming

Among the above functions, Functions 1–3 are the original functions of LoRaWAN, while Functions 4–6 are the new functions of LW-LoRaWAN to enhance access protection. Function 4 requires the introduction of a new hardware platform, so the Universal Software Radio Peripheral (USRP) is used to receive RF signals from LoRa terminals and extract the RFF. Since the RFF is data-independent and the number of terminals in the LoRa network

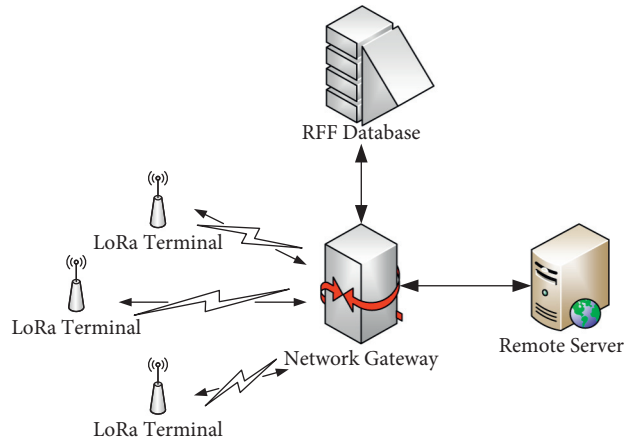


FIGURE 1: LW-LoRaWAN system architecture.

can be very large, Function 5 is deployed to establish the relevance between the RFF and the data frame of LoRa terminals for the implementation of secure policies. Function 6 is in charge of intercepting the data packets from illegal terminals based on the trained RFF database of legitimate terminals. When two or more terminals are found with the same ID but different RFFs, this function will be triggered to alert the remote server for the presence of counterfeit terminal.

3.2. The Design of New Functions in LW-LoRaWAN. The design of Functions 4–6 is described in detail in the following.

3.2.1. RFF Extraction and Identification of LoRa Terminals. The extraction and identification of LoRa RFF can be divided into three stages: signal acquisition and preprocessing, RFF extraction, and recognition and decision of RFF.

The first stage is acquisition and sampling the received LoRa signal. After the signal is collected by USRP, the signal frame is preprocessed including energy normalization and the coarse and fine synchronization.

The purpose of frame synchronization is to accurately locate and extract the signal frame from the received signal segment. Firstly, the coarse synchronization of the frame is conducted with the double sliding window method and the approximate position of the frame head is obtained. The coarse synchronization can meet the requirements of carrier frequency offset calculation, but more accurate synchronization is needed for RFF extraction. Therefore, the cross-correlation calculation is used to find the position of the maximum correlation peak, which corresponds to the frame starting point for the fine synchronization. According to the format of LoRa frame, the preamble part contains at least 6 identical up-chirp symbols. The cross correlation between the coarse-synchronized signal and the standard up-chirp signal is calculated. When the correlation peak is found, the position of the peak is the offset of the coarse-synchronized signal. Then the offset is compensated for the fine synchronization.

In the second stage, the RFFs of LoRa terminals are extracted and the flowchart is shown in Figure 2.

As shown in Figure 2, the RFF features of LoRa terminals are divided into transient features and steady features. The transient features can be found in the rising and falling edge of the signal frame. The steady features include the IQ offset and the carrier frequency offset [35].

In the third stage, the transient and steady features are extracted from each signal frame. The Euclidean distance of feature vectors between the devices is calculated and registered in the legitimate database. According to the size of the distance deviation, the legal and illegal terminals can be identified.

3.2.2. Establishing the Relevance between RFF and Data Frame. This function is the core idea of the proposed lightweight access scheme. LoRaWAN achieves the authentication of the terminals by presharing key, which means that the access control is based on the device ID. The proposed new function demonstrates a novel idea of using data as the object of access control rather than IDs. The data frame contains both the transmitted data and the RFF features of the terminal. When the ID information is included in the payload of the data frame, the binding relationship of “terminal ID-terminal data-RFF” can be established. Then, the access control of data packets based on the RFF can be realized.

The access security policy of the physical layer can achieve the access control of the terminal alone. From the perspective of the overall architecture, the gateway needs to implement the RFF extraction function, while the terminal requires no modification. There are a large number of terminals in the LoRa network, but the data throughput is limited. Therefore, it is possible to implement the RFF binding with the data frame without adding too much resources burden.

3.2.3. Real-Time Illegal Data Blocking and Abnormal Terminal Alarming. The above functions complete the RFF extraction of each packet and determine the attribution of the data packet only from the physical layer. This function performs real-time data processing, forwarding, or discarding according to the validity of the data packets. In addition, this function replaces the LoRaWAN communication function and it needs to be compatible with the data format, protocol, modulation, and demodulation of LoRaWAN.

Different from the existing security scheme, the greatest advantage of the proposed protection scheme is that it uses the uniqueness of RFF to achieve the identification of counterfeit terminals. When two or more different RFFs are found with the same terminal ID, it means that there exist counterfeit terminals in the network. Hence, this function performs real-time data blocking and abnormal terminal alarming for this terminal ID to prevent malicious data from being uploaded to the remote server. The subsequent data with this terminal ID then is blocked until the terminal returns to normal. Compared with some existing access

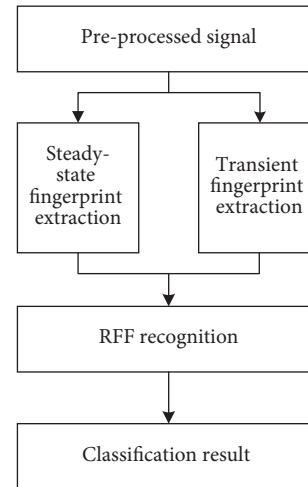


FIGURE 2: The flowchart of RFF extraction.

control strategies based on the terminal behaviors, the proposed function can immediately detect the existence of counterfeit devices and effectively block the malicious data.

3.3. LW-LoRaWAN Workflow. The working procedure of LW-LoRaWAN is simpler than that of LoRaWAN, as shown in the following:

- (1) The initialization phase: The RFF and corresponding ID of the legitimate terminals are stored in the RFF database. The RFF database can be an independent device as shown in Figure 1 or integrated with the gateway. When a new legitimate terminal joins the network, its ID and RFF can be added to the RFF database.
- (2) Normal working phase: The gateway receives the RF signal and extracts the RFFs from LoRa terminals. Meanwhile, the RF signal is demodulated into the link layer packets with the terminal ID according to the LoRa physical layer protocol. The RFF is bonded to the corresponding ID and the terminal is marked as online.
- (3) According to the current packet, the gateway matches the RFFs of the online terminal with that in the RFF database to determine whether the online terminal is legitimate. If the terminal is judged to be illegal, the working process goes to (4). If the terminal is judged to be legal, the gateway forwards the current packet to the remote server.
- (4) The gateway blocks the packets of illegal terminals to prevent them from accessing the gateway. The gateway stores the alarm information of the illegal terminals and sends it to the remote server.
- (5) For the server side, different servers transmit data to the terminals. The gateway receives the packets from the remote server and extracts the terminal ID.
- (6) The gateway queries whether the terminal ID has an alarm message. If not, the terminal ID is judged to be

legal and the working process goes to (7). If the alarm message exists, the terminal is judged to be illegal and the gateway blocks its packets.

- (7) The gateway forwards the legitimate packets of the remote server to the terminals through the physical layer protocol.

The workflow of LW-LoRaWAN is shown in Figure 3.

3.4. Comparative Analysis of LW-LoRaWAN. The proposed lightweight security enhancement scheme based on RFF and the modern cryptographic security enhancement scheme are designed to solve the existing LoRaWAN security issues. From a functional point of view, both of these schemes have promoted the secure access protection for LoRa terminals. Then, the implementation and performance of the two schemes are qualitatively compared and the analysis results are shown in Table 1.

As shown in Table 1, the modern cryptography enhancement scheme maintains the technical specifications of the original gateway and the compatibility of the original communication system due to the software upgrade of the original gateway. However, this implementation requires the software upgrade for all the existing terminals and gateways. Due to the small number of gateways in the LoRa network, the workload of gateway upgrade is limited, but, for the massive and widely deployed LoRa terminals, upgrade is almost impossible. Therefore, the enhancement scheme can only be implemented on the newly deployed devices. In addition, the conflict between modern cryptography and the limited resources of LoRa terminals still exists, which greatly reduces the battery life of LoRa terminals.

The proposed LoRa gateway security enhancement scheme replaces the original LoRaWAN gateway with the USRP and its RF performance depends on the USRP specifications. The advantages of the RFF scheme include the uniqueness of preventing counterfeit attacks; there is no need to upgrade a large number of terminals and the designing freedom for customized functions.

4. RFF Extraction Method for Single Data Frame

The primary requirement of the LoRa gateway security enhancement is not to affect the normal functions of the original network architecture. According to the analysis in the previous section, an access control mechanism for the south side of the gateway is introduced, which requires stable and effective extraction of packet information and the corresponding RFF from each data frame. At present, the LoRa RFF features include the frequency offset [35], the overall data [36], and the DCTF [37]. However, these features generally require the accumulation of a certain number of packets for statistical analysis. In order to solve the problem of extraction efficiency, this section proposes an RFF feature extraction method based on the Cross Power Spectral Density (CPSD) of LoRa signals, which can extract stable and unique RFF information from a single frame.

4.1. LoRa Signal Analysis. Though the RFF features are data-independent, they are generally weak relative to the modulation waveform of the signal. There are multiple identical preambles in RF communication signals, so it is easy to perceive the weak RFF information from the known preambles. Therefore, most of the existing algorithms extract the RFF features from the preambles. However, the preambles only occupy a small proportion of the entire data frame. When the payload behind the preamble is used, more raw data can be utilized for RFF extraction.

Compared with traditional CSS technology, LoRa modulation further improves the deployment of spectrum. LoRa modulation is essentially a circular shift of the standard chirp symbol to obtain the modulated signal and the information transmitted by each symbol is determined by the initial frequency offset. Therefore, all the LoRa symbols can be obtained theoretically by cyclic shift of any symbol, which means that all the symbols in a single data frame can be shifted into the same waveforms for RFF extraction.

The chirp signal is composed of sinusoidal signals and the frequency varies linearly with time. A time-domain waveform of duration T can be expressed as

$$c(t) = \text{rect}\left(\frac{t}{T}\right) \cdot e^{j\varphi(t)}, \quad (1)$$

where $\text{rect}(t/T)$ is a rectangular signal:

$$\text{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & \left|\frac{t}{T}\right| \leq \frac{1}{2} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

In equation (1), $\varphi(t)$ represents the phase of the chirp signal and the equation for calculating the instantaneous frequency $f(t)$ from the signal phase can be expressed as

$$f(t) = \frac{1}{2\pi} \cdot \frac{d\varphi(t)}{dt}. \quad (3)$$

Due to the linear relationship between chirp signal frequency and time, there is

$$f(t) = f_c + \mu \cdot \frac{B}{T} \cdot t = f_c + \mu Kt, \quad (4)$$

where f_c represents the carrier frequency, μ represents the instantaneous frequency changing slope of the chirp signal, B represents the bandwidth, and $K = TB$ represents the frequency modulation slope. $\mu = 1$ means up-chirp and $\mu = -1$ means down-chirp. The IQ signals and instantaneous frequencies of the up-chirp and down-chirp are shown in Figures 4 and 5, respectively, and the signal frequency varies linearly within a bandwidth of 250 kHz.

LoRa modulation encodes the data by cyclic shifting the chirp signal by k bits, where $0 \leq k \leq 2^{SF} - 1$. After k -bit cyclic shifting on equation (4), the result can be expressed as

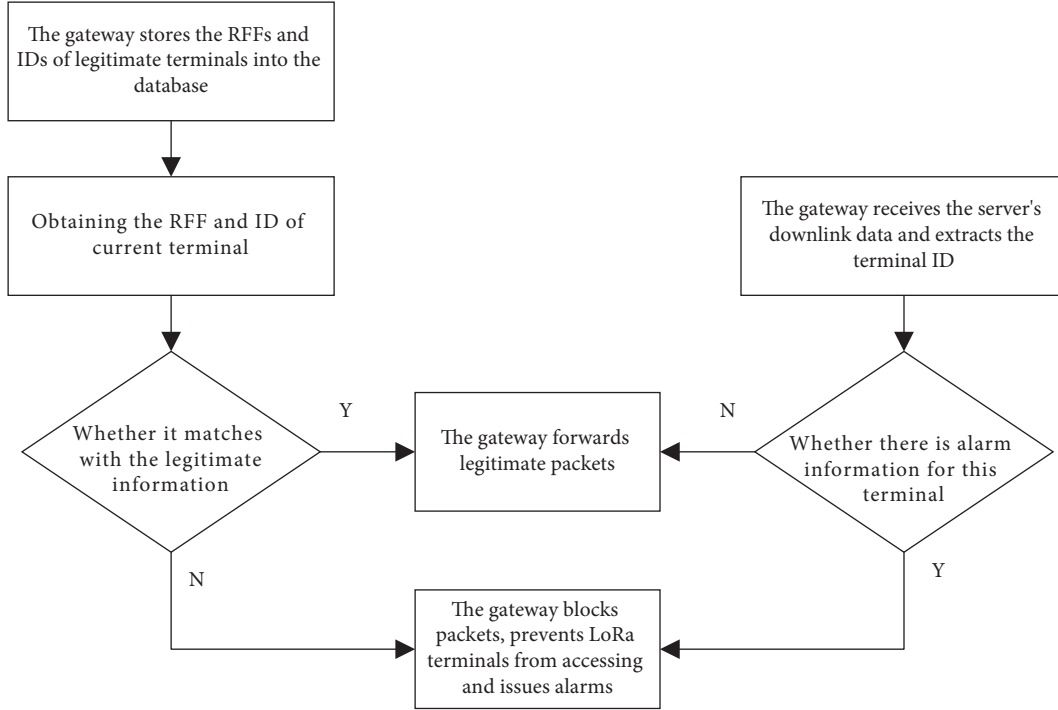


FIGURE 3: The workflow of LW-LoRaWAN.

TABLE 1: Comparison of two LoRa gateway security enhancement schemes.

Schemes	Modern cryptography	RFF technology
Modulation and demodulation	Using LoRaWAN	Self-realization
Communication distance	Far	Normal
Access algorithm freedom	No	High
Gateway upgrading content	Software upgrading	Hardware and software upgrading
Terminal upgrading content	Software upgrading	No
Protection mechanism	Increasing cracking workload	Uniqueness of physical features
Counterfeit attack protection	No	Yes

$$f(t) = \begin{cases} f_c + \mu K \left(t - \frac{k}{B} \right), & -\frac{T}{2} + \frac{k}{B} \leq t \leq \frac{T}{2}, \\ f_c + \mu K \left(t - \frac{k}{B} \right) + B, & -\frac{T}{2} \leq t \leq -\frac{T}{2} + \frac{k}{B}. \end{cases} \quad (5)$$

Taking Figure 4 as the reference, Figure 6 shows the waveform of chirp signal after cyclic shifting of 30 bits.

As shown from Figure 4 to Figure 6, the rule of cyclic shifting for LoRa modulation is obvious, which verifies the feasibility of obtaining the same waveform from the actual LoRa signals. The typical time-frequency diagram of LoRa data frame is shown in Figure 7. In this experiment, the data frame is in an explicit header mode and the data contain 10 up-chirp, 2.25 down-chirp, the explicit header, and the payload. As shown in Figure 7, the up-chirp and down-chirp remain unchanged in each data frame and the subsequent payload varies as the transmitted information changes. As long as each chirp symbol can be synchronized, the cyclic shifting of the payload into the same waveform can be realized.

4.2. LoRa Data Frame Composition. The LoRaWAN protocol mainly defines the technical details and specifications of the LoRa physical layer and MAC layer. The LoRa physical layer frame has two message formats: uplink and downlink. Uplink messages sent by the terminals reach the remote server through the gateway and the downlink is in the opposite direction. Both the uplink and downlink messages include the preamble and PHYPayload. In explicit mode, the message includes the physical layer header (PHDR) and its cyclic redundancy check (PHDR_CRC), which are not included in implicit mode. In terms of frame format, the only difference between uplink and downlink messages is that uplink messages have a cyclic redundancy check (CRC) to protect the integrity of the payload. The data frame format in explicit mode is shown in Figure 8.

The preamble is composed of n up-chirp symbols and 2.25 down-chirp symbols for data synchronization and the value of n is selected from 6 to 65536. The PHDR includes the payload data length, CR, and other values. When the above values are fixed, the implicit mode can be selected to shorten the transmission time. The length of the payload is variable, and its content includes data and MAC layer settings.

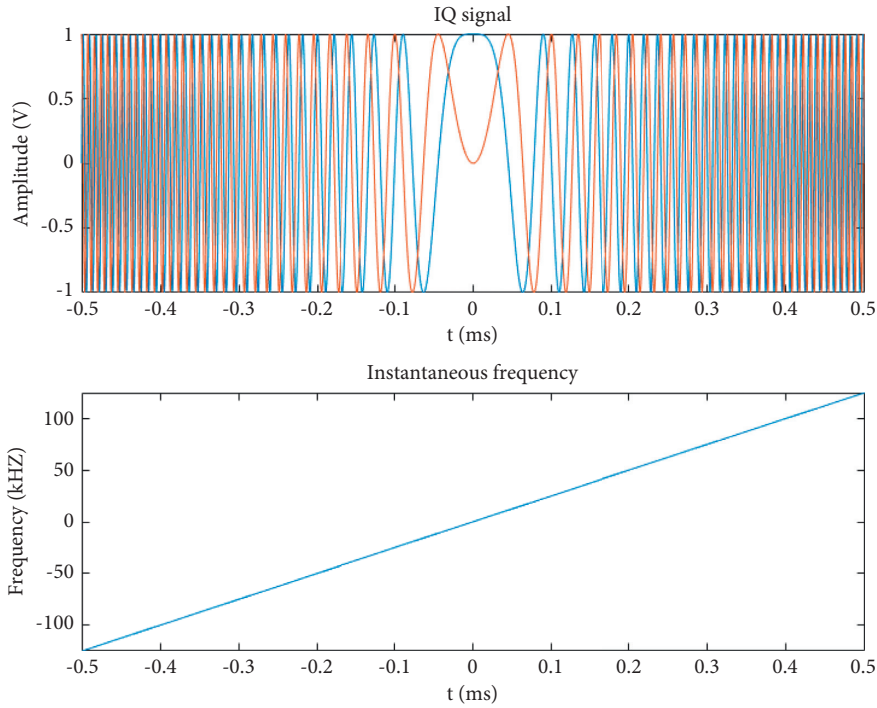


FIGURE 4: Parameters of up-chirp.

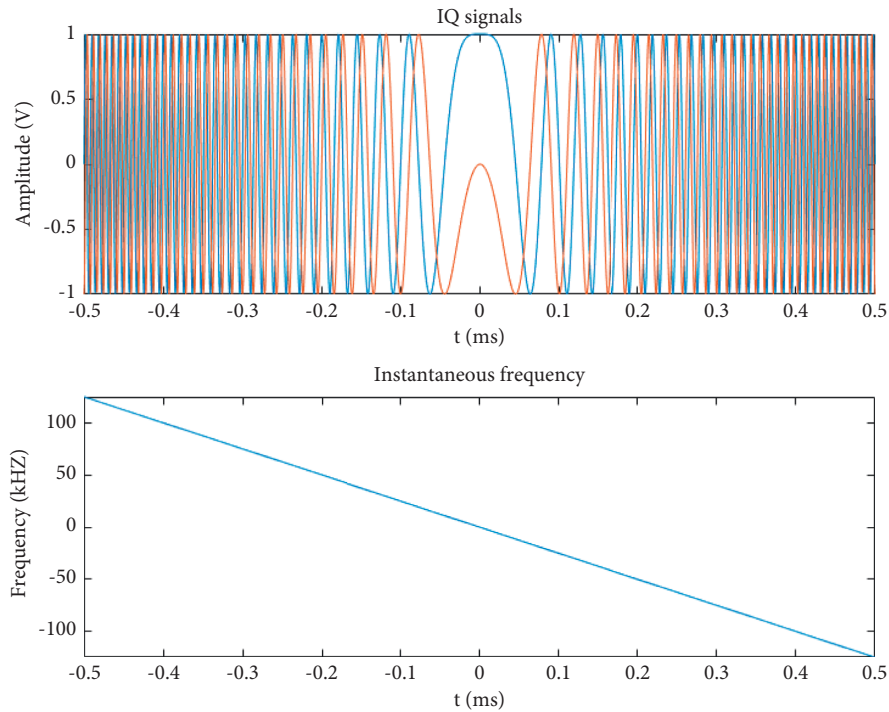


FIGURE 5: Parameters of down-chirp.

4.3. *Cross Power Spectral Density (CPSD) Extraction.* The CPSD can be used to describe the correlation between two random processes at each frequency point. The chirp symbol $x(n)$ in the LoRa signal is cyclically correlated with the up-chirp symbol $y(n)$ and then the CPSD features of the chirp symbol are obtained by the Fourier transform. CPSD reflects the energy

features of the chirp symbols in the amplitude-frequency curve and the cyclic shifting features in the phase-frequency curve. Therefore, through the amplitude-frequency curve, the preamble and payload of the data frame can be effectively analyzed in the same dimension without considering the difference of the initial frequencies.

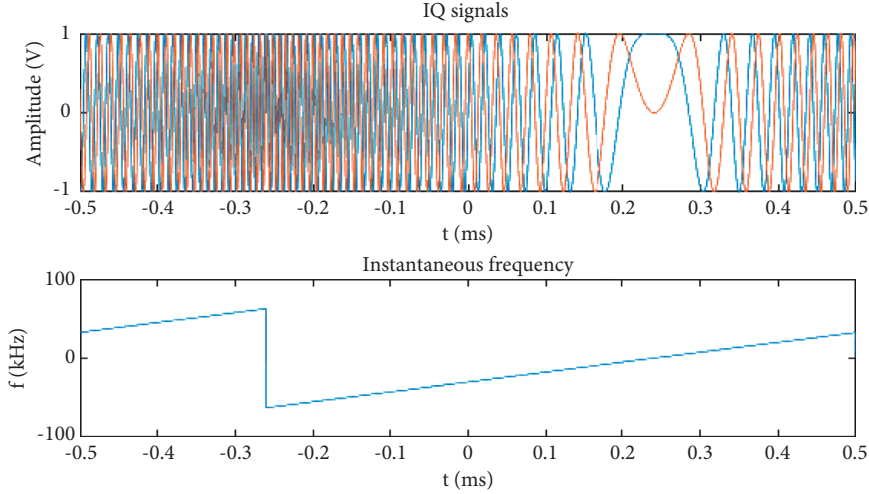


FIGURE 6: Chirp signal waveform after 30-bit shifting.

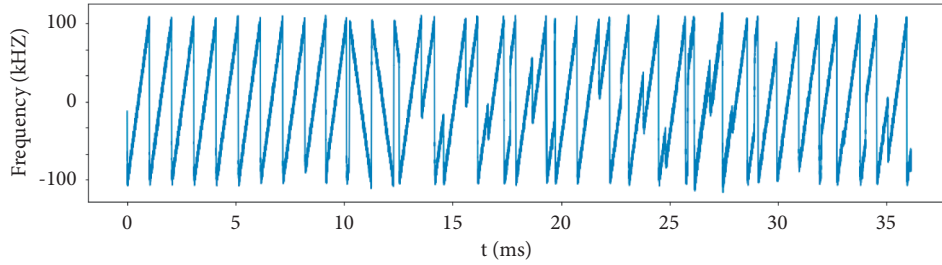


FIGURE 7: Time-frequency diagram of LoRa data frame.

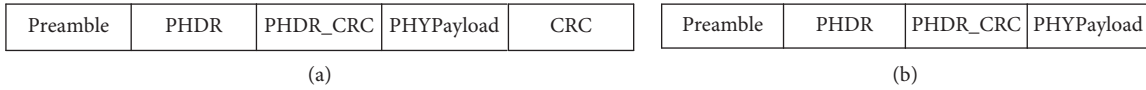


FIGURE 8: LoRa physical layer frame format (explicit mode). (a) Upstream frame format. (b) Downlink frame format.

Cyclic correlation is a kind of correlation operation for the cyclic shifting of the data sequences. Since both $x(n)$ and $y(n)$ are finite-length sequences of equal length, $y(n)$ can be selected for periodic extension. The period is the number of sampling points N of a chirp symbol and then $y(n)$ is shifted to the left by m bits after N points extension $R_N(n)$, which can be expressed as $y((n+m))_N R_N(n)$. The sequences can be shifted in one direction, because, after the periodic extension, shifting one position to the left is equivalent to shifting $N-1$ positions to the right. Finally, taking N values from $(0, N-1)$, the N sequence values are obtained after the cyclic shifting.

The N -point cyclic correlation of $x(n)$ and $y(n)$ can be expressed as

$$r_{xy}(m) = \sum_{n=0}^{N-1} x(n)y((n+m))_N R_N(n), \quad (6)$$

where $m = 0, 1, 2, \dots, N$ and $r_{xy}(m)$ represents the correlation result of $x(n)$ and $y(n)$ after cyclic shifting by m points. The length of r_{xy} is also N points and the N -point CPSD of $x(n)$ and $y(n)$ can be expressed as

$$G_{xy}(k) = \frac{1}{N} \sum_{m=0}^{N-1} r_{xy}(m) e^{-j\frac{2\pi}{N}mk}, \quad 0 \leq k \leq N-1. \quad (7)$$

There are multiple chirp symbols in a LoRa frame. In order to prevent the first and last symbols from possible power instability, the average CPSD of the rest of chirp symbols is taken as the RFF feature of the frame. The specific steps for calculating this feature are as follows:

- (1) Extracting L chirp symbols $X_i, i \in (1, 2, \dots, L)$ in the middle of the data frame.
- (2) Calculating the cyclic correlation r_i between $X_i, i \in (1, 2, \dots, L)$ and the up-chirp symbol $Y(n)$, respectively.

$$r_i(m) = \sum_{n=0}^{N-1} X_i(n)Y((n+m))_N \cdot R_N(n), \quad m \in (0, 1, \dots, N-1), \quad (8)$$

where $i \in (1, 2, \dots, L)$.

- (3) Calculating the CPSD vector G_i from the cyclic correlation vector r_i .

$$G_i(k) = \frac{1}{N} \sum_{m=0}^{N-1} r_i(m) e^{-j\frac{2\pi}{N}mk}, 0 \leq k \leq N-1, \quad (9)$$

where $i \in (1, 2, \dots, L)$.

- (4) Averaging the CPSD vectors to obtain the CPSD features of the frame.

$$\text{CPSD}(n) = \frac{\sum_{i=1}^L G_i(n)}{L}, \quad 0 \leq n \leq N-1. \quad (10)$$

Equation (10) indicates that the CPSD feature of the signal is an N -dimensional vector. When $N=2048$, the CPSD diagram of two terminals is depicted in Figure 9. The curves of the CPSD between different terminals are roughly the same. Then, the expanded view is shown in Figure 10 and there are intuitive differences between the two terminals.

Since the feature dimension is 2048 in Figure 9, the amount of data is large. In order to reduce the computation, it is necessary to reduce the feature dimensions. For example, only the 925th to 1124th dimensions of the CPSD have larger amplitudes and more obvious differences, so this 200-dimensional vector can be used as the CPSD features of the terminal.

5. System Implementation and Testing

5.1. System Structure. The overall structure of the proposed system is shown in Figure 11. After the USRP completes the signal acquisition, downconversion, and analog-to-digital conversion, the RFF System executes the RFF feature extraction of the digital baseband signals and the training and recognition of the physical layer identity of different terminals. Then, the RFF System transmits the judgment results to the Management System for data management and integration, and then the results can be handled and displayed on the remote interface.

The physical hardware diagram of the system is shown in Figure 12. The sending device is a LoRa terminal powered by the USB interface and the receiving device is the NI USRP N210. The USRP and the computer deployed with the RFF System are connected through a gigabit network cable to exchange data. Meanwhile, in order to facilitate the development and demonstration of the program, the Management System and the display interface are deployed on the same computer. In practical applications, the display interface and part of the Management System can be deployed remotely and accessed through the Internet.

During the process, GNU Radio [38] is used to realize the sampling control and signal collection for USRP N210. GNU Radio is an open-source software toolkit for building and deploying software defined radio (SDR) systems. It can process wireless signals and control the parameters such as the sampling frequency, the spectrum range, and the gain. The system uses Python3 to support the signal demodulation and MATLAB to support the RFF feature extraction.

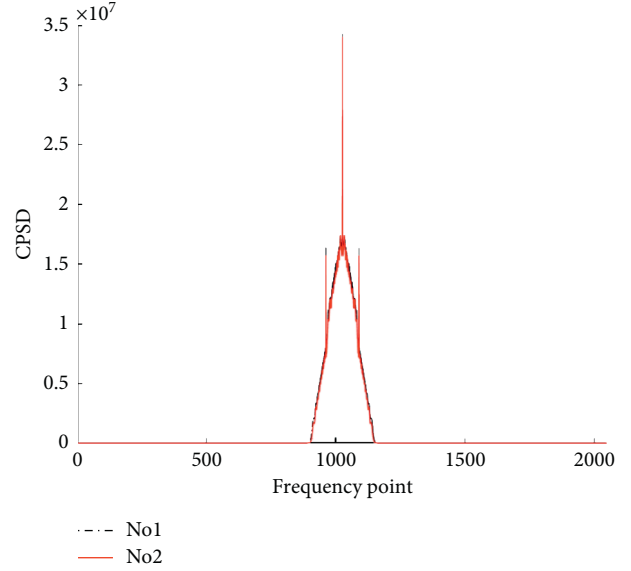


FIGURE 9: The CPSD of two terminals.

The LoRa terminals select the same batch of products from the same manufacturer, as shown in Figure 13. The product consistency makes the RFF features similar and poses a greater challenge for the classification and identification. The LoRa terminals set the carrier frequency to 433 MHz, the SF to 7, the BW to 125 kHz, and the number of up-chirp symbols in the preamble to 10. In the identification phase, the terminal ID is written to the payload to distinguish different terminals.

5.2. System Software Design. According to the system functions and workflow in Section 3, the system software can be divided into seven parts: signal acquisition module, data preprocessing module, RFF extraction module, device registration module, identification module, system management module, and display module.

The downconversion frequency is set at 433 MHz to realize zero IF acquisition, and the sampling frequency is set at 2 MHz. The file receiver module stores the acquired binary data locally, while the header module controls the amount of data collected at once.

GNU Radio Companion [39] is a visual interface supported by GNU Radio to achieve the signal acquisition. The configuration of the parameters is shown in Figure 14. For UHD, USRP Source is the parameter control module to collect signals from USRP. The downconversion frequency is set to 433 MHz to achieve the zero intermediate frequency acquisition and the sampling frequency is set to 2 MHz. The File Sink module stores the acquired binary data locally, while the Head module controls the amount of data collected for one time. The LoRa Receiver module has the same parameters as the LoRa terminal; it demodulates and decodes the LoRa signals. The Message Socket Sink module transmits the demodulated data to the local computer through the Socket and establishes the relevance between the terminal ID in the demodulated data and the baseband data output by the File Sink module.

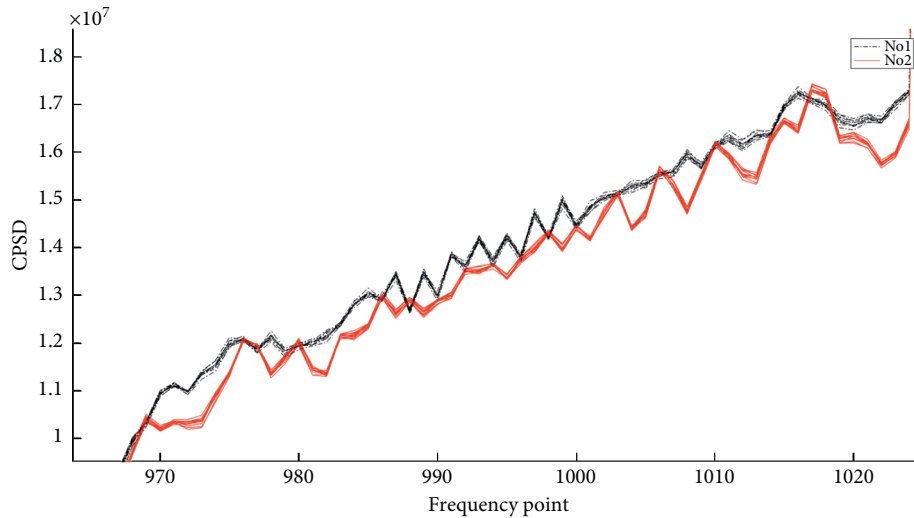


FIGURE 10: The expanded view of CPSD of the two terminals.

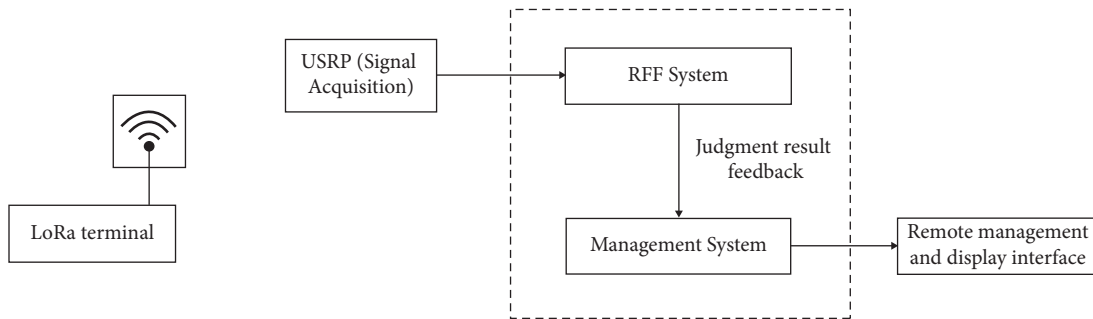


FIGURE 11: System structure diagram.



FIGURE 12: System hardware diagram.

The data preprocessing module performs the signal processing on the baseband signal, including reading valid signals, dividing data segments, normalization, frequency, and phase offset estimation and compensation. The RFF extraction module operates on the preprocessed data and

establishes the relationship among the terminal ID, the timestamp, the signal strength, the signal-to-noise ratio (SNR), the carrier frequency offset, the CPSD, and other values extracted from each signal frame to form the unique RFF. The device registration module, identification module,

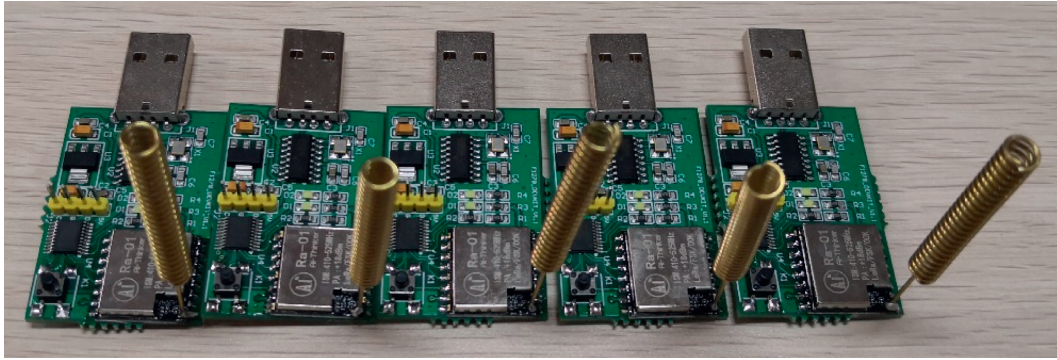


FIGURE 13: LoRa terminals.

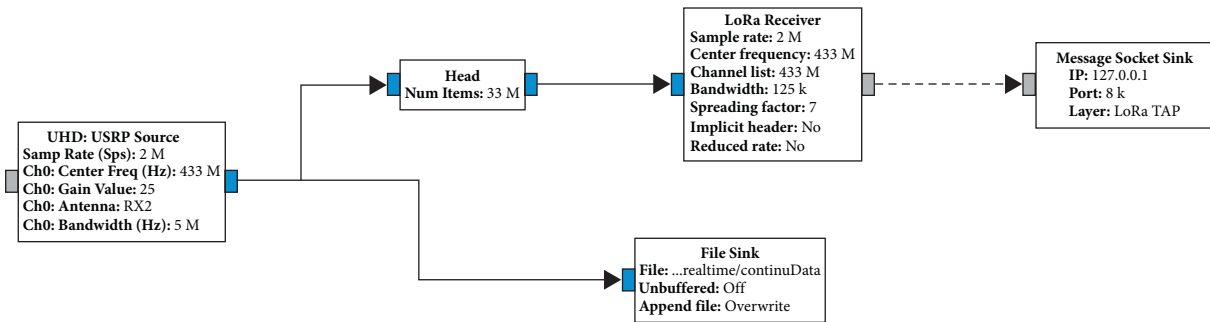


FIGURE 14: Signal acquisition module diagram.

and system management module operate according to the workflow in Figure 3. The display module provides the interface for parameter configuration and information display.

5.3. System Test

5.3.1. Test Scenario. After establishing the recognition system based on LoRa RFF features, the performance tests are carried out to ensure that the system can operate normally in different environments. This section focuses on the analysis of the performance tests.

System performance refers to the classification and recognition accuracy of legitimate terminals and the recognition rate of counterfeit terminals in a certain environment. Three experimental scenarios are tested. In each scenario, 8 LoRa terminals and one USRP are used. 50 sets of single-frame data are firstly collected for each terminal to generate a training model and then 100 sets are used to test the classification and recognition accuracy of the system.

In the first scenario, the LoRa terminals and the USRP communicate in a line-of-sight (LOS) indoor environment; hence, less interference comes from the surrounding and the RFF features are stable. The second scenario is the non-line-of-sight (NLOS) indoor environment and the received signal is greatly affected by multipath channel. The third scenario is a NLOS partition wall environment which leads to long distances, complex channels, and low SNR. By artificially adding white Gaussian noise to change the SNR values, the classification performance with different SNR has been evaluated.

5.3.2. Experiment 1: Comparison for Different RFF Features. The first experiment is carried out in Scenario 1 to verify the recognition efficiency of CPSD features by comparing the recognition accuracy of the carrier frequency offset, the IQ offset, and the CPSD.

Firstly, the performance of the carrier frequency offset features is analyzed, as shown in Figure 15. The linear discriminant analysis (LDA), linear kernel Support Vector Machine (SVM), and Gaussian kernel SVM are used for terminal recognition with multiple SNRs. Even at high SNR, the accuracy of carrier frequency offset features is only about 92%. This feature provides similar results with the three classification algorithms, but unfortunately some terminals with similar frequency offsets are difficult to distinguish.

Secondly, the IQ offset features are tested under multiple SNRs and the results are shown in Figure 16. The test results show that this feature requires high SNR and the recognition accuracy is less than 60% when the SNR is lower than 20 dB. Therefore, the accuracy of this feature is not sufficient, and it can only be used as an auxiliary feature for the device recognition. The analysis of CPSD features is drawn in Experiment 2.

5.3.3. Experiment 2: Comparison of CPSD for Different Scenarios. After dimension reduction for the CPSD features, the recognition accuracies under multiple SNRs for the three scenarios are shown in Figure 17. Compared with the results of Experiment 1 in Scenario 1, the recognition accuracy based on the CPSD feature has been significantly improved and exceeded 99% when SNR is 30 dB. In Scenario 2, when

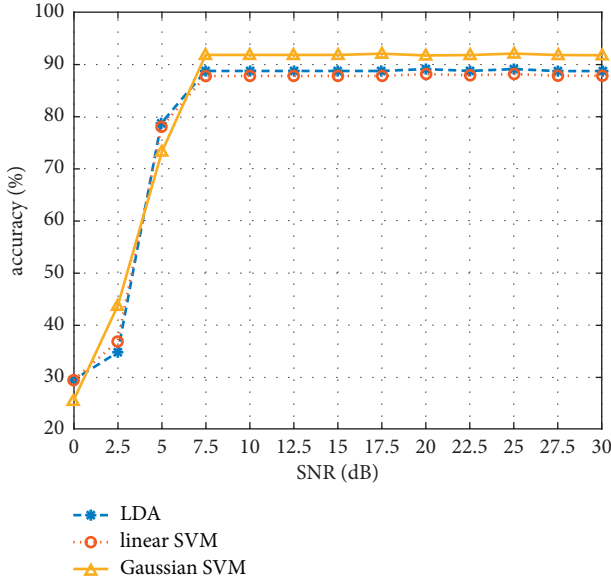


FIGURE 15: Test results of carrier frequency offset features.

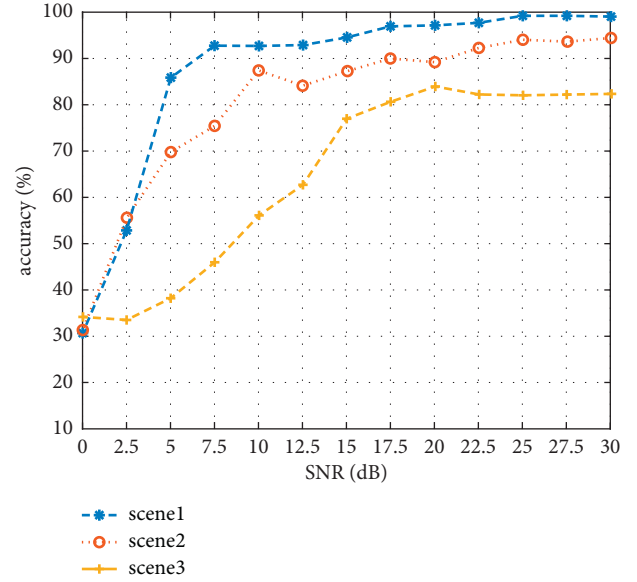


FIGURE 17: Test results of CPSD feature in three scenarios.

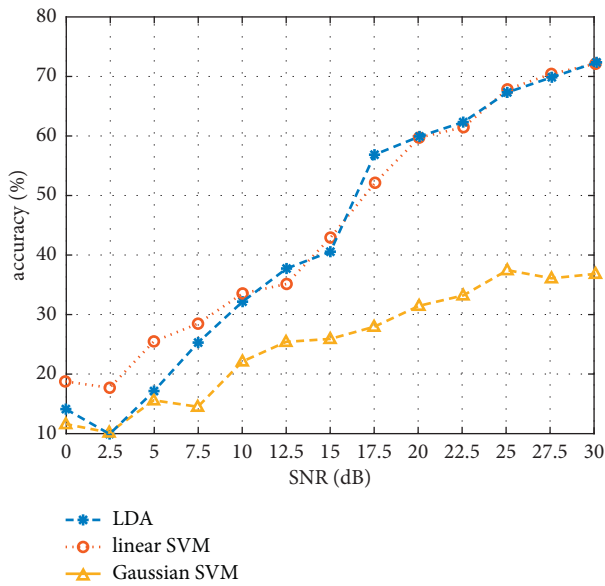


FIGURE 16: Test results of IQ offset feature.

the SNR is greater than 10 dB, the recognition accuracy exceeds 80%. In Scenario 3, under the condition of low SNR, the recognition accuracy is degraded, and the maximum recognition accuracy is about 83%.

The experimental results show that the CPSD feature using the payload information has a better recognition accuracy compared to the other two features. Therefore, the CPSD feature is more feasible in the proposed system.

5.3.4. Experiment 3: Analysis of Counterfeit Attack Protection. This experiment tests the system's ability to recognize counterfeit attacks. Since the ID of the LoRa terminal is written into the payload of the data frame, it can

be tampered by changing the payload to achieve the counterfeit attack for the specified ID.

In this experiment, all terminals are numbered from 1 to 8 in order, and terminal 8 is used as the attacker to impersonate terminals 1 to 7, respectively. In Scenario 1, each terminal is tested for 100 counterfeit attacks and the results are shown in Table 2 for SNR of 30 dB. The detection success rate of counterfeit terminals refers to the percentage of counterfeit attacks recognized by the system. The false alarm rate refers to the probability that a legitimate terminal is identified as an illegal terminal.

The detection success rate of counterfeit terminals indicates the system's ability to recognize counterfeit attacks. The detection success rate reflects the system's ability to resist attacks from counterfeit terminals. The false alarm rate indicates the system's ability to recognize the legitimate terminals. A lower false alarm rate leads to a higher ability to recognize legitimate terminals.

As shown in Table 2, the proposed system has a high success rate of counterfeit attack detection. Meanwhile, the false alarm rate of legitimate terminals remains to be acceptable. Therefore, the proposed system can resist the counterfeit attack, which is unable to be achieved by the existing IoT security protection system.

5.4. Comparative Experiments. The proposed CPSD method is compared with the existing methods [35–37] mentioned in Section 2. The RFF features used in [35] are the time plus frequency offset (TFO). The supervised machine learning methods in [36] include MLP, convolutional neural network, and SVM, where MLP achieves the highest recognition accuracy. The classification method in [37] analyzed the features of the DCTF with the image recognition algorithm. Then, the performances of the four types of methods are evaluated with different fingerprinting experiments.

TABLE 2: Test results of counterfeit attacks.

Counterfeited terminal number	Detection success rate of counterfeit terminal (%)	False alarm rate (%)
1	95	3
2	96	4
3	98	2
4	97	2
5	95	3
6	94	5
7	96	3

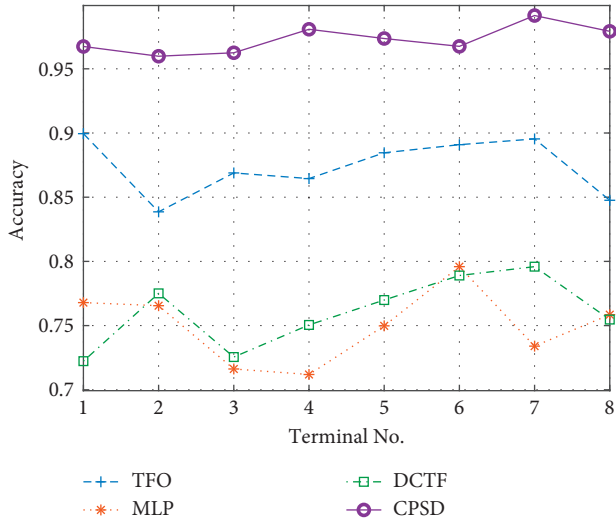


FIGURE 18: Accuracy comparison for 8 terminals.

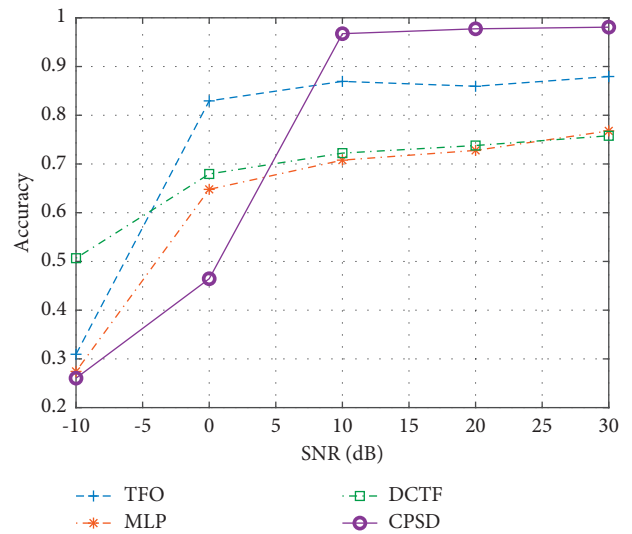


FIGURE 20: Accuracy comparison for different SNR.

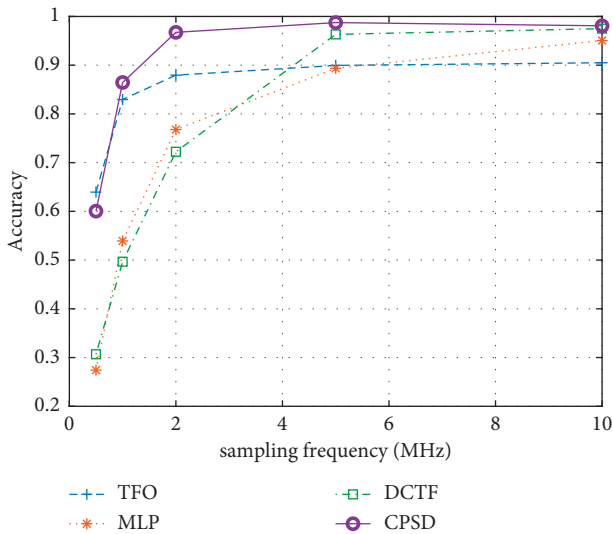


FIGURE 19: Accuracy comparison for different sampling frequency.

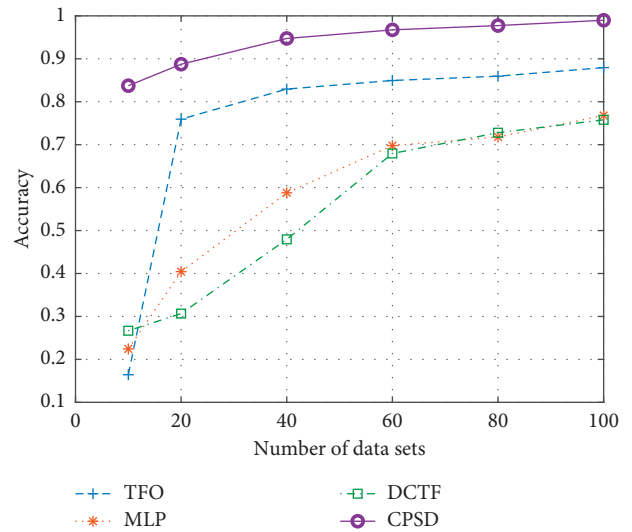


FIGURE 21: Accuracy comparison for different number of data sets.

The first experiment tests the recognition accuracy of 8 terminals in Scenario 1 with a fixed SNR of 30 dB and a sampling frequency of 2 MHz. The comparison results are shown in Figure 18. The MLP and DCTF rely on the signal details at high sampling frequency, so the identification accuracy is poor. The TFO and CPSD are less affected by the sampling frequency, and the accuracy is above 80%.

The second experiment tests the effect of sampling frequency in Scenario 1, where each terminal has a fixed SNR of 30 dB. The average results are shown in Figure 19. The accuracies of TFO and CPSD are significantly higher than those of MLP and DCTF at low sampling frequencies. With the increase of sampling frequency, the accuracy of all the

algorithms is improved. When the sampling frequency exceeds 5 MHz, all the algorithms achieve better accuracy.

The third experiment tests the effect of SNR when the sampling frequency is 2 MHz in Scenario 1, and the average results are shown in Figure 20. The RFF reflects the subtle features of the signal, so the SNR has a great influence on the RFF. With the decline of SNR, the accuracy of all algorithms decreases obviously. Compared with TFO and DCTF, MLP and CPSD are more sensitive to the change of SNR.

The fourth experiment tests the effect of the number of data sets in Scenario 1 with a fixed SNR of 30 dB and a sampling frequency of 2 MHz. The average results are shown in Figure 21. The number of data sets determines the ability of the algorithm to quickly extract stable RFFs. The smaller the amount of data required to extract stable RFFs, the more practical the algorithm is. Unfortunately, most of the current research on RFF algorithms does not consider this problem. Since the CPSD algorithm makes effective use of the data part of the signal, higher accuracy can be obtained in the case of a small amount of data, as shown in Figure 21.

Through the comparative analysis of the above 4 experiments, the following conclusions can be drawn. Compared with the other three algorithms, the proposed CPSD algorithm achieves the best performance under the conditions of low sampling frequency and high SNR and can obtain high stable RFF with the least amount of data.

6. Conclusions and Future Work

Different from the secure access technology of modern cryptography, a lightweight gateway architecture called LW-LoRaWAN is proposed to achieve a data frame-based authentication with RFF. Compared to the two kinds of lightweight access techniques presented in Section 2, the RFF-based access method uses a different security policy. The main advantages and differences are listed as follows:

- (1) LW-LoRaWAN is proposed to achieve a data frame-based authentication by establishing the relevance between RFF and data frame of LoRa terminals, which conforms to the concept of zero trust.
- (2) The current RFF extraction methods for LoRa terminals cannot provide a stable RFF within a small number of packets, so a novel RFF feature of CPSD is proposed to achieve a fast authentication within one data frame.
- (3) Since the RFF is unique and unclonable, LW-LoRaWAN can protect the LoRa terminals against the Sybil attacks.
- (4) The proposed security enhancement system only needs to upgrade the gateway, without any change to the large number of terminals, which is more feasible than the existing enhancement schemes in practical applications.

From the results of our work, we can arrive at a conclusion that the proposed security policy could be a promising approach for LoRa terminal authentication. The theoretical analysis and experimental results show that the

proposed system not only improves the authentication security of LoRa network but also protects the LoRa terminals against the counterfeit attacks. The LW-LoRaWAN provides new ideas from the physical layer for the security of LoRa devices.

In future work, we plan to test the system performance and terminal recognition rate for more communication scenarios. In addition, the RF performance of LW-LoRaWAN depends on the capacity of the USRP, and we will try other RF platforms to improve the gateway performance.

Data Availability

The data supporting this study are available within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported in part by Jiangsu Key R&D Plan BE2019109; the National Natural Science Foundation of China under Grants 61601114, 61602113, 61801115, and 62001106; Natural Science Foundation of Jiangsu Province under Grants BK20160692, BK20200350, and BK20200352; Jiangsu Provincial Key Laboratory of Network and Information (Security no. BM2003201); and the Purple Mountain Laboratories for Network and Communication Security.

References

- [1] M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, and E. K. Markakis, "A survey on the internet of things (IoT) forensics: challenges, approaches, and open issues," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1191–1221, 2020.
- [2] S. H. Haji and S. Y. Ameen, "Attack and anomaly detection in IoT networks using machine learning techniques: a review," *Asian Journal of Research in Computer Science*, vol. 9, no. 2, pp. 30–46, 2021.
- [3] H. Mrabet, S. Belguith, A. Alhomoud, and A. Jemai, "A survey of IoT security based on a layered architecture of sensing and data analysis," *Sensors*, vol. 20, no. 13, p. 3625, 2020.
- [4] W. Iqbal, H. Abbas, M. Daneshmand, B. Rauf, and Y. A. Bangash, "An in-depth analysis of IoT security requirements, challenges, and their countermeasures via software-defined security," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10250–10276, 2020.
- [5] M. M. Ogonji, G. Okeyo, and J. M. Wafula, "A survey on privacy and security of Internet of Things," *Computer Science Review*, vol. 38, p. 100312, 2020.
- [6] K. Lounis and M. Zulkernine, "Attacks and defenses in short-range wireless technologies for IoT," *IEEE Access*, vol. 8, pp. 88892–88932, 2020.
- [7] Y. Jia, F. Zhong, A. Alrawais, B. Gong, and X. Cheng, "FlowGuard: an intelligent edge defense mechanism against IoT DDoS attacks," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9552–9562, 2020.
- [8] I. Yaqoob, E. Ahmed, I. A. T. Hashem et al., "Internet of things architecture: recent advances, taxonomy, requirements, and

- open challenges,” *IEEE wireless communications*, vol. 24, no. 3, pp. 10–16, 2017.
- [9] F. Olivier, G. Carlos, and N. Florent, “New security architecture for IoT network,” *Procedia Computer Science*, vol. 52, pp. 1028–1033, 2015.
- [10] N. Pimple, T. Salunke, U. Pawar et al., “Wireless security—an approach towards secured wi-fi connectivity,” in *Proceedings of 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 872–876, IEEE, Coimbatore, Tamil Nadu, February 2020.
- [11] Y. Wen and T. Liu, “WIFI security certification through device information,” in *Proceedings of 2018 International Conference on Sensor Networks and Signal Processing (SNSP)*, pp. 302–305, IEEE, Xi’an, China, October 2018.
- [12] E. Baray and N. K. Ojha, “WLAN security protocols and WPA3 security approach measurement through aircrack-ng technique,” in *Proceedings of 2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 23–30, IEEE, Erode, India, April 2021.
- [13] B. Danev, D. Zanetti, and S. Capkun, “On physical-layer identification of wireless devices,” *ACM Computing Surveys*, vol. 45, no. 1, pp. 1–29, 2012.
- [14] J. Sakhnini, H. Karimipour, A. Dehghantaha, and R. M. Parizi, “Physical layer attack identification and localization in cyber-physical grid: an ensemble deep learning based approach,” *Physical Communication*, vol. 47, Article ID 101394, 2021.
- [15] L. Bai, L. Zhu, J. Liu et al., “Physical layer authentication in wireless communication networks: a survey,” *Journal of Communications and Information Networks*, vol. 5, no. 3, pp. 237–264, 2020.
- [16] Z. Li, W. Xu, R. Miller et al., “Securing wireless systems via lower layer enforcements,” in *Proceedings of the 5th ACM workshop on Wireless security*, pp. 33–42, Los Angeles California, September 2006.
- [17] R. Zhang, L. Song, Z. Han et al., “Physical Layer Security for Two Way Relay Communications with Friendly Jammers,” in *Proceedings of 2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pp. 1–6, IEEE, Miami, Florida, USA, December 2010.
- [18] J. Chen, R. Zhang, L. Song et al., “Joint relay and jammer selection for secure two-way relay networks,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 310–320, 2011.
- [19] P. Neumann, J. Montavont, and T. Noel, “Indoor deployment of low-power wide area networks (LPWAN): a LoRaWAN case study,” in *Proceedings of 2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 1–8, IEEE, New York, NY, USA, October 2016.
- [20] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, “Understanding the limits of LoRaWAN,” *IEEE Communications Magazine*, vol. 55, no. 9, pp. 34–40, 2017.
- [21] M. Bor, J. E. Vidler, and U. Roedig, “LoRa for the Internet of Things,” in *Proceedings of International Conference on Embedded Wireless Systems and Networks (EWSN) 2016*, pp. 1–7, Graz, Austria, 2016.
- [22] A. Shah and M. Engineer, “A survey of lightweight cryptographic algorithms for iot-based applications,” *Smart innovations in communication and computational sciences*, Springer, Singapore, pp. 283–293, 2019.
- [23] A. Biswas, A. Majumdar, S. Nath et al., “LRBC: a lightweight block cipher design for resource constrained IoT devices,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 1–15, 2020.
- [24] S. Aruna, G. Usha, P. Madhavan, and M. V. R. Kumar, “Lightweight cryptography algorithms for IoT resource-starving devices,” *Role of Edge Analytics in Sustainable Smart City Development: Challenges and Solutions*, pp. 139–169, 2020.
- [25] B. Seok, J. Park, and J. H. Park, “A lightweight hash-based blockchain architecture for industrial IoT,” *Applied Sciences*, vol. 9, no. 18, p. 3740, 2019.
- [26] M. A. Philip, “A survey on lightweight ciphers for IoT devices,” in *Proceedings of 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy)*, pp. 1–4, IEEE, Kollam, India, December 2017.
- [27] X. W. Wu, E. H. Yang, and J. Wang, “Lightweight security protocols for the internet of things,” in *Proceedings of IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–7, IEEE, Montreal, QC, Canada, October 2017.
- [28] A. S. Sani, D. Yuan, P. L. Yeoh et al., “A lightweight security and privacy-enhancing key establishment for internet of things applications,” in *Proceedings of 2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, Kansas City, MO, USA, May 2018.
- [29] K. I. Talbot, P. R. Duley, and M. H. Hyatt, “Specific emitter identification and verification,” *Technology Review*, p. 113, 2003.
- [30] J. Tapparel, O. Afisiadis, P. Mayoraz et al., “An open-source LoRa physical layer prototype on GNU radio,” in *Proceedings of 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, IEEE, Atlanta, Georgia, USA, May 2020.
- [31] M. O. Farooq, “Multi-hop communication protocol for LoRa with software-defined networking extension,” *Internet of Things*, vol. 14, Article ID 100379, 2021.
- [32] J. Souifi, Y. Bouslimani, M. Ghribi et al., “Smart home architecture based on LoRa wireless connectivity and LoRaWAN® networking protocol,” in *Proceedings of 2020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP)*, pp. 95–99, IEEE, El-Oued, Algeria, March 2020.
- [33] C. Bernier, F. Dehmas, and N. Deparis, “Low complexity LoRa frame synchronization for ultra-low power software-defined radios,” *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3140–3152, 2020.
- [34] L. Tessaro, C. Raffaldi, M. Rossi et al., “Lightweight synchronization algorithm with self-calibration for industrial LORA sensor networks,” in *Proceedings of 2018 Workshop on Metrology for Industry 4.0 and IoT*, pp. 259–263, IEEE, Brescia, Italy, April 2018.
- [35] R. Eletreby, D. Zhang, S. Kumar et al., “Empowering low-power wide area networks in urban settings,” in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, pp. 309–321, Los Angeles CA USA, August 2017.
- [36] P. Robyns, E. Marin, W. Lamotte et al., “Physical-layer fingerprinting of LoRa devices using supervised and zero-shot learning,” in *Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 58–63, Boston Massachusetts, July 2017.
- [37] Y. Jiang, L. Peng, A. Hu et al., “Physical layer identification of LoRa devices using constellation trace figure,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–11, 2019.

- [38] T. W. Mathumo, T. G. Swart, and R. W. Focke, "Implementation of a GNU Radio and python FMCW Radar Toolkit," in *Proceedings of IEEE AFRICON*, pp. 585–590, IEEE, Cape Town, September 2017.
- [39] M. Chino, H. Miyashiro, and A. J. Luis, "Implementation of SNR estimation algorithms, using LabVIEW communications and GNU radio companion," in *Proceedings of 2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, pp. 1–4, IEEE, Lima, Peru, August 2018.

Research Article

ICSTrace: A Malicious IP Traceback Model for Attacking Data of the Industrial Control System

Feng Xiao ¹, Enhong Chen,¹ Qiang Xu,² and Xianguo Zhang³

¹Anhui Province Key Laboratory of Big Data Analysis and Application School of Computer Science and Technology, University of Science and Technology of China, Hefei, China

²Electronic Engineering Institute of Hefei, Hefei, China

³School of Cyberspace Security, University of Science and Technology of China, Hefei, China

Correspondence should be addressed to Feng Xiao; xiaof686@mail.ustc.edu.cn

Received 4 June 2021; Accepted 15 July 2021; Published 2 August 2021

Academic Editor: Jingyu Feng

Copyright © 2021 Feng Xiao 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.

Considering that the attacks against the industrial control system are mostly organized and premeditated actions, IP traceback is significant for the security of the industrial control system. Based on the infrastructure of the internet, we have developed a novel malicious IP traceback model, ICSTrace, without deploying any new services. The model extracts the function codes and their parameters from the attack data according to the format of the industrial control protocol and employs a short sequence probability method to transform the function codes and their parameters into a vector, which characterizes the attack pattern of malicious IP addresses. Furthermore, a partial seeded K -means algorithm is proposed for the pattern's clustering, which helps in tracing the attacks back to an organization. ICSTrace is evaluated based on the attack data captured by the large-scale deployed honeypots for the industrial control system, and the results demonstrate that ICSTrace is effective on malicious IP traceback in the industrial control system.

1. Introduction

With the rapid development of the Internet of Things (IoT), more and more Industrial Control Systems (ICS) are connected into the Internet. As the key bond between the virtual signal and the real equipment, an Internet-connected ICS makes the production process be more accurate and agile. But it also narrows the distance between the cyberattacks and the industrial infrastructure. As we know, Stuxnet worm was disclosed to be the first worm attacking the energy infrastructure [1, 2] in 2010. In 2014, the hackers attacked a steel plant in Germany so that the blast furnace cannot be closed properly [3]. On December 23, 2015, the Ukrainian power network suffered a hacker attack, which was the first successful attack to the power grid, resulting in hundreds of thousands of users suffering power blackout for hours [4]. In 2017, the security vendor ESET disclosed an industrial

control network attack weapons named as win32/Industroyer, which implemented malicious attacks on power substation system [5].

ICSs are highly interconnected and interdependent with the critical national infrastructure [6], and thus the attackers have noticed the high returns to attack ICS in recent years. The attackers are diverse in identity. They may be hackers, members of organized criminal groups, or even a hostile country. The worse situation is that ICS has become the new target of terrorists to gain the influence by destroying the real physical world. As traditional ICS is physically isolated from the Internet, most researches just focus on the functional safety of the system rather than the security consideration of the network. There are not any special protective measures, not to mention the attribution mechanism for tracing the attack back [7]. Security researchers are now committed to the intrusion detection technology for ICS. They want to

identify, intercept, and alert the threats, before a severe attack occurs. These intrusion detection technologies can be divided into several categories as follows: state-based [8], behavior-based [9], rule-based [10], characteristic-based [11], model-based [12], and ML-based (machine learning) [13, 14].

Because ICS plays an important role in the critical national infrastructure, the cyberattacks against ICS are mostly organized and premeditated actions. It is significant not only to determine whether there is a threat in ICS but also to trace the attack back. Furthermore, locating the initiators and their motivations before or during an attack is crucial for deterring and cracking down the premeditated and organized attackers.

Attribution is one of the most intractable problems of an emerging field, created by the underlying technical architecture and geography of the Internet [15]. The current dominant IP traceback technologies include packet marking mechanism [16], packet logging mechanism [17], and their hybrid [18, 19]. Packet marking mechanism needs the routers to write a tag (e.g., IP address) into some fields of every packet. The target retrieves all the tags from the received packets and finds out the routing path. Packet marking mechanism includes two categories: probabilistic packet marking (PPM) [16] and deterministic packet marking (DPM) [20]. Packet logging mechanism needs the routers to record all the forwarded packets so as to reveal the routing path. Apparently, this mechanism consumes a lot of storage space. All these IP traceback technologies above need to redesign the Internet or to deploy new services. There is still no applicable IP traceback system to deploy over the network.

The ultimate goal of attribution is identifying an organization or a government, not individuals [15]. Our study identifies an organization by zooming down to a single IP level and then zooming back out to an organization or a unit level without changing the Internet architecture or deploying new services. Instead of tracing back to the source of a packet directly, we just recognize the malicious IP addresses which belong to the same organization.

In this study, we present a malicious IP traceback model (ICSTrace) for industrial control system, and this model makes the following contributions:

- (1) Based on the deep analysis of ICS protocol S7, the function codes and their parameters are extracted from the attack data.
- (2) A feature vector of the function codes and their parameters are designed to represent the attack patterns.
- (3) The slide window method is adopted to reduce the dimension of those multidimensional samples.
- (4) A partial seeded K -means clustering algorithm is proposed based on K -means algorithm.
- (5) ICSTrace is proven to be effective basing on the real attack data captured by the large-scale deployed honeypots for ICS.

Section 2 introduces the research background and our previous work on the attack data collection. Section 3 describes the architecture of our IP traceback model. Sections 4 and 5 introduce the attack pattern extraction method and partial seeded K -means algorithm for clustering, respectively. In Section 6, we evaluate our IP traceback model basing on the real attack data. Section 7 is our related works, and Section 8 is the conclusion.

2. Background

ICS is a business process management and control system which is composed of various automatic control and process control components. It collects and monitors real-time signals to ensure the function of the automatic operation or the process control. Its application fields include program automation, industrial control, intelligent building, power transmission and distribution, smart meter, and car communication. ICS protocol refers to the communication protocol used in ICS. The most well-known ICS protocol includes S7, Modbus, BACnet, and DNP3.

S7 protocol is a Siemens proprietary protocol [21] running on programmable logic controllers (PLCs) of Siemens S7-200, 300, and 400 series. It is suitable for either Ethernet, PROFIBUS, or MPI networks. Because the objects of this study are those industrial control systems which are accessed to the Internet, we only discuss the TCP-based S7 protocol in Ethernet networks. As shown in Figure 1, S7 protocol packets are packed by COTP protocol and then packed by TPKT protocol package for TCP connection.

As shown in Figure 2, the communication procedure of S7 protocol is divided into three stages. The first stage is to establish COTP connection, the second stage is to set up S7 communication, and the third stage is to exchange the request and the response for function code.

The Magic flag of the S7 protocol is fixed to 0×32 , and the following fields are S7 type, data unit ref, parameters length, data length, result info, parameters, and data. In parameters field, the first byte stands for the function code of S7. Table 1 shows the optional function codes of S7. Communication Setup code is used to build a S7 connection; Read code helps the host computer to read data from PLC; Write code helps the host computer to write data to PLC. As for the codes of Request Download, Download Block, Download End, Download Start, Upload, and Upload End, they are designed for downloading or uploading operations of blocks. PLC Control code covers the operations of Hot Run and Cool Run, while PLC Stop is used to turn off the device.

When the function code is 0×00 , it stands for system function which is used to check system settings or status. The details are described by the 4-bit function group code and 1-byte subfunction code in the parameters field. System functions are further divided into 7 groups, as shown in Table 2. Block function is used to read the block, and time function is used to check or set the device clock.

At present, there is not any ICS attacking dataset for security research. Therefore, we developed a high interactive ICS honeypot named as S7commTrace in previous work [22], based on Siemens' S7 protocol.

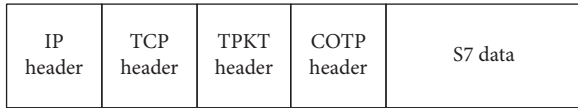


FIGURE 1: Header format of the S7 communication packet.

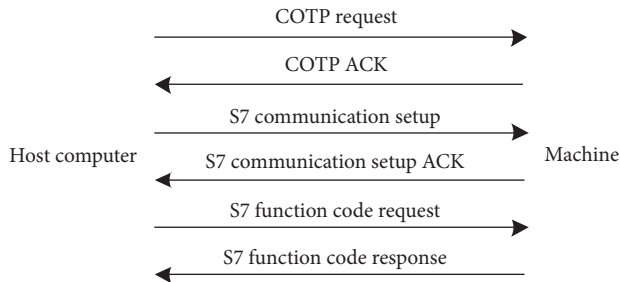


FIGURE 2: Communication procedure of the S7 protocol.

TABLE 1: S7 protocol function code and the corresponding function.

Code	Function
0x00	System functions
0x04	Read
0x05	Write
0x1a	Request download
0x1b	Download block
0x1c	Download end
0x1d	Download start
0x1e	Upload
0x1f	Upload end
0x28	PLC control
0x29	PLC stop
0xf0	Communication setup

TABLE 2: When the function code is 0×00 , it is system function and further divided into 7 groups.

Function group code	Function	Subfunction code	Subfunction
1	Programmer commands	1	Request diag data
		2	VarTab
2	Cyclic data	1	Memory
		1	List blocks
3	Block function	2	List blocks of type
		3	Get block info
		1	Read SZL
4	CPU function	2	Message service
		1	PLC password
5	Security PBC BSEND/BRECV	1	PLC password
		None	None
7	Time function	1	Read clock
		2, 3	Set clock
		4	Read clock (following)

Honey pot is a kind of security resource that is used to attract the attacker for illegal application without any business utility [23]. Honey pot technology is a method to set some hosts, network services, or information as a bait, to induce attackers, so that the behavior of the attacks can be captured and analyzed [24]. Honey pot can be used to better understand the landscape of where these attacks are originating [25].

S7commTrace poses as a real PLC device by simulating the S7 protocol to capture the probing and attacking data. It can be divided into four modules, including TCP Communication module, S7 Protocol Simulation module, Data Storage module and User Template, as shown in Figure 3.

The main function of TCP Communication module is to listen on TCP port 102, submit the received data to the Protocol Simulation module, and reply to the remote peer. S7comm Protocol Simulation module parses the received data according to the protocol format and obtains the valid contents at first. Then, S7comm Protocol Simulation module generates the reply data referring to User Template. At last, the reply data are sent back to TCP Communication module to be packaged. User Template records all the user-defined information such as PLC serial number and manufacturer. The Data Storage module handles the request and the response of data storage.

We deployed S7commTrace honeypots in United States, China, Germany, Russia, Japan, Singapore, and Korea at the same time. Each S7commTrace ran for 272 days on average. At last, we captured 110,501 requests of S7comm protocol, as shown in Table 3. In fact, not all requests are in accordance with S7comm format. Ignoring them, S7commTrace records a total of 46492 valid requests. If we define an uninterrupted TCP communication connection as a session, S7commTrace records 5797 sessions and 4224 valid sessions. Furthermore, a valid IP address indicates that this IP has at least one valid session.

According to the DNS query results, we find that there are 26 IP addresses pointing to Shodan.io, 19 IP addresses pointing to eecs.umich.edu, 16 IP addresses pointing to neu.edu.cn, and 5 IP addresses pointing to plscan.org, as shown in Table 4. This means 573 valid IP addresses belong to four organizations at least.

Shodan.io [26] is the domain suffix of Shodan which is a search engine in cyberspace. In addition to retrieving traditional web services, Shodan has used the ICS protocol directly to crawl the ICS devices on the Internet and visualizes their location and other information. Eecs.umich.edu is the domain suffix of the Department of Electrical and Computer Science (EECS) Department of University of Michigan, which is one of the agencies developing Censys [27, 28]. Censys scans the devices in the Internet and stores the results in its database. It provides not only web and API query interfaces but also raw data to download. Neu.edu.cn is the domain suffix of Northeastern University of China which develops a search engine name as Detecting [29]. Detecting is capable of providing accurate information of ICS devices and their locations. Plscan.org is the domain suffix of Beacon Lab [30] which is committed to the research

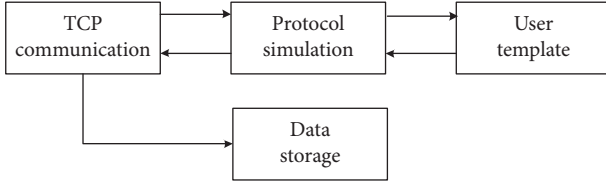


FIGURE 3: Structure of the ICS honeypot (S7commTrace).

TABLE 3: Count of all attack data and valid attack data after 13 honeypots were run for 272 days.

Item	Count
Request	110,501
Valid request	46,492
Session	5797
Valid session	4224
IP address	897
Valid IP address	573

TABLE 4: IP statics by DNS reverse lookup.

Domain	Organization	IP number
Shodan.io	Shodan	26
eecs.umich.edu	Censys	19
neu.edu.cn	Detecting	16
plscan.org	Beacon Lab	5
Others	Unknown	507

and the practice related to ICS security. These four organizations are the well-known security research institutes. They are scanning the devices in the Internet all the time, including the ICS devices. Except for the 66 IP addresses belonging to four well-known organizations, there are still 507 IP addresses which are resolved to be dynamic domain name or none domain name.

3. Structure of the ICSTrace Model

When an attacker launches the attacks, he usually hides the IP address of his own resorting to the anonymous communication networks such as springboard host, VPN, and other measures. As shown in Figure 4, after an ICS suffered an attack from the Internet, the security personnel can only see the last IP address connected to ICS instead of the real IP address of the attacker, not to mention the organization which belongs to.

ICSTrace transforms the features of data from each IP address into a one-dimensional eigenvector. This eigenvector stands for the unique pattern of an attack. Therefore, the problem of attribution turns into a problem of clustering the patterns.

As shown in Figure 5, the input of ICSTrace is a malicious IP and its packets. The output is a cluster containing multiple IP addresses, which indicates an organization. ICSTrace model consists of three stages, including protocol resolution, attack pattern extraction, and partial seeded K -means clustering. The main function of protocol resolution

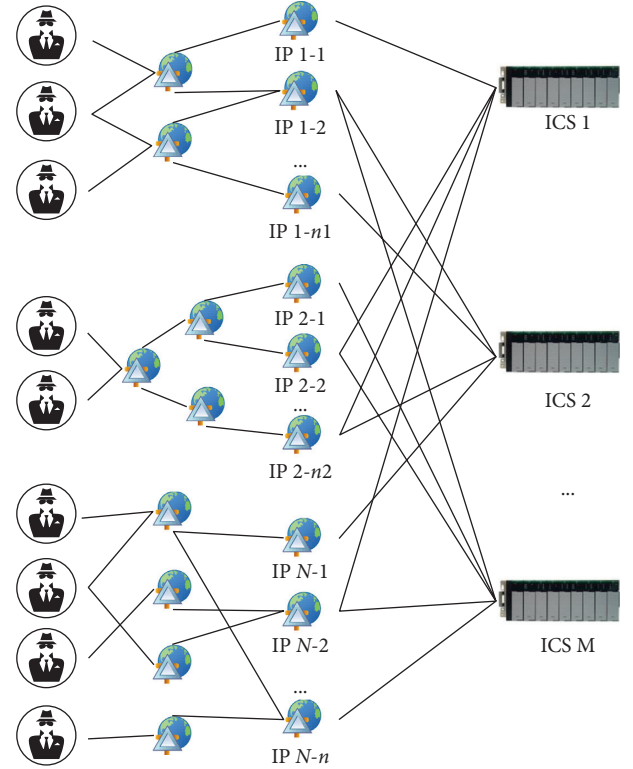


FIGURE 4: Schematic diagram of attacking flow.

is to parse the packets and extract the function codes and their parameters. Attack pattern extraction transforms the function codes and their parameters into one-dimensional vector as the attack pattern of a certain IP address. Partial seeded means is used to cluster the attack patterns so that those IP addresses with the same patterns are aggregated into one cluster. And then, the cluster is labeled as a certain organization according to some auxiliary information (e.g., domain name or geographical location) of the IP addresses in it.

The function codes are used to achieve the effects of operations in most industrial control protocols like S7, modbus, bacnet, and DNP3. Since the inputs of ICSTrace model are function codes and their parameters, the model is applicable for other industrial control protocols besides S7. As for the general Internet protocol, they transmit information not only by function codes and they are much more complex than the industrial control protocols. Therefore, ICSTrace model is not completely applicable for the general Internet protocol.

4. Attack Pattern Extraction

After an attacker has constructed the connection with ICS, he will carry out a series of delicate operations on purpose, which are expressed by the function codes and their parameters [22]. Therefore, the attacking features, which are extracted from the function codes and their parameters of S7comm protocol data, can reveal the intention of the attacker effectively.

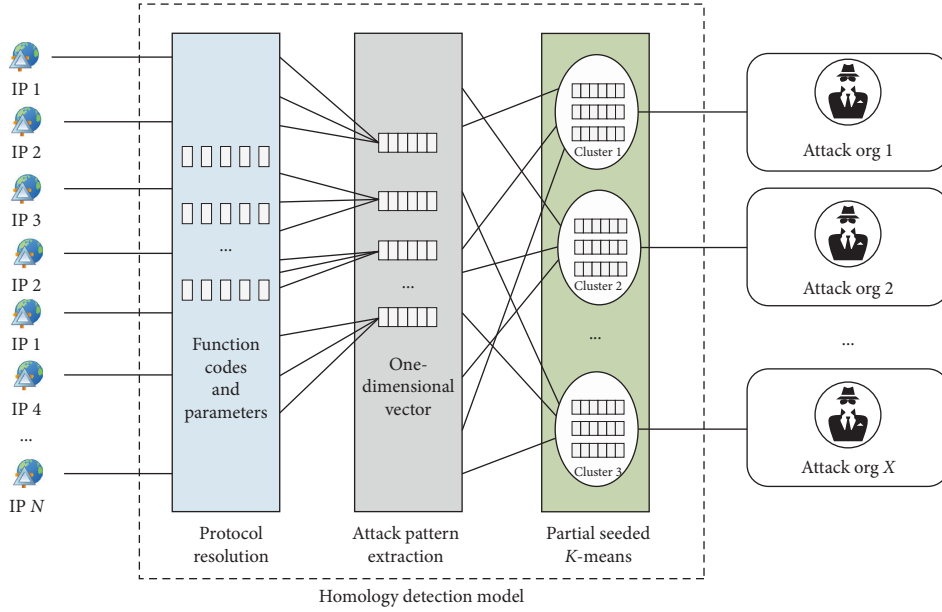


FIGURE 5: Structure of the ICSTrace model.

As shown in Figure 6, one attacker may have several IP addresses to launch attacks. We have defined an uninterrupted TCP communication as a session, and one IP address may attack one or more ICSs for more than one time. Thus, a single-source IP may build several sessions. We call a packet sent by the attacker as a request, and there are several packet interactions, so a session usually contains many requests.

The function codes and their parameters of S7comm protocol are included in these requests, so we extract these from the communication data package, which is sent by the attacker to the receiver, as the feature of the attacker to construct IP traceback model.

4.1. Mean Count of Function Codes and Parameters. Mean count of function codes (MCFC) refers to the average amount of the function codes of each session from the same IP address. Different attackers have different motivations, objectives, and methods while conducting a cyberattack. As a result, quantities of requests and function codes are very different in different sessions:

$$\text{MCFC} = \frac{1}{n} \sum_{i=1}^n (\text{Count_of_function_codes})_{\text{session}_i}, \quad (1)$$

$\text{session}_i \in \text{IP}.$

Mean count of the parameters (MCP) refers to the average amount of the parameters used in the function codes of each session from the same IP address. Some function codes do not need parameters, and some function codes need one or more parameters, so different attackers use different amounts of parameters:

$$\text{MCP} = \frac{1}{n} \sum_{i=1}^n (\text{Count_of_parameters})_{\text{session}_i}, \quad \text{session}_i \in \text{IP}. \quad (2)$$

4.2. Function Code Sequence and Parameter Sequence. Function code sequence (FCS) indicates the change rule of the function codes in all sessions from a single IP address. Different attackers may use the same kind of function codes while launching an attack, but the chronological order is different. As shown in Figure 7, the Function code C_1, C_2, \dots, C_i can be arrayed to form a Markov chain in chronological order.

Array the function codes in the session to form a function code sequence according to the chronological order:

$$F_{\text{session}_i} = (C_1, C_2, \dots, C_i), \quad \text{session}_i \in \text{IP}. \quad (3)$$

For some sessions may belong to the same source IP address, we combine the function codes serials and parameter serials of all sessions from the same IP address into a set of function code sequence:

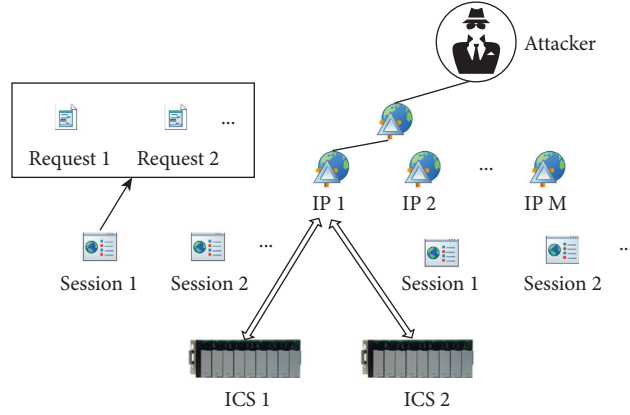


FIGURE 6: Schematic diagram of attack IP session request.

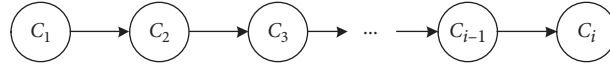


FIGURE 7: Function code sequence.

$$F_n = \begin{pmatrix} F_{\text{session}_1} \\ F_{\text{session}_2} \\ \dots \\ F_{\text{session}_n} \end{pmatrix} = \begin{pmatrix} C_1 & C_2 & \dots & C_{a_1} \\ C_1 & C_2 & \dots & C_{a_2} \\ \dots & \dots & \dots & \dots \\ C_1 & C_2 & \dots & C_{a_n} \end{pmatrix}, \quad \text{session}_i \in \text{IP}, 1 \leq i \leq n. \quad (4)$$

Different amounts of sessions originate from each source IP, and various methods are adopted by the attackers for each time, which results in the different function code sequences in each session. Therefore, F_n of different source IP addresses are two-dimensional matrix vectors with unequal rows and columns.

These FCSs with uncertain amount and unequal length cannot be handled directly, for clustering algorithms like K -means need samples with same dimensions. In this study, we propose a method to convert these sequences with uncertain amount and unequal length into the vectors with the same length; the detailed process is as follows:

Step 1: add the start and the end status to the sequence.

For a sample set of sequence F_n , there are n sequences with unequal length and the length of which are $a_1, a_2, \dots, a_n, a_i \geq 1, i \in [1, n]$, respectively. Add the start and the end status to each sequence in F_n , and then get F'_n . Now the length of each sequence is no less than 3:

$$F'_n = \begin{pmatrix} S & C_1 & C_2 & \dots & C_{a_1} & E \\ S & C_1 & C_2 & \dots & C_{a_2} & E \\ S & \dots & E & & & \\ S & C_1 & C_2 & \dots & C_{a_n} & E \end{pmatrix}. \quad (5)$$

Step 2: obtain the unrepeatable set of short sequences.

Setting the window length equals 3 and the stride equals 1, use the slide window method to process each sequence in F'_n . Then, get a_1, a_2, \dots, a_n short sequences with the same length of 3, $a_i \geq 1, i \in [1, n]$. Then, remove the duplicate sequences and add the short sequences into set $S = (s_1, s_2, \dots, s_m), m \leq \sum_{i=1}^n a_i$.

Step 3: obtain the short sequences set of all sample sets.

Process all of the sequence sample sets according to Steps 1 and 2, and get a short sequence set $S = (s_1, s_2, \dots, s_k)$ without duplication.

Step 4: express the probability vector of the sequences with uncertain amount and unequal length:

$$P_n = \begin{pmatrix} C_1 & C_2 & \dots & C_{b_1} \\ C_1 & C_2 & \dots & C_{b_2} \\ \dots & \dots & \dots & \dots \\ C_1 & C_2 & \dots & C_{b_l} \end{pmatrix}, \quad (6)$$

$$P'_n = \begin{pmatrix} S & C_1 & C_2 & \dots & C_{b_1} & E \\ S & C_1 & C_2 & \dots & C_{b_2} & E \\ S & \dots & E & & & \\ S & C_1 & C_2 & \dots & C_{b_l} & E \end{pmatrix}.$$

For a sequence set P_n corresponding to a certain IP, there are l function code sequences with unequal length and the lengths of them are $b_1, b_2, \dots, b_l, b_i \geq 1, i \in [1, l]$. By adding the start and the end status to each sequence, we get P'_n . Then, process all the function codes sequences with the slide window method to construct a feature vector X_{ip} according to the frequency of these short sequences:

$$\begin{aligned} X_{ip} &= (X_{s_1}, X_{s_2}, \dots, X_{s_k}), \\ \sum_{i=1}^k X_{s_i} &= 1. \end{aligned} \quad (7)$$

The method for FCS feature vector processing is shown in Figure 8. We make an improvement on the short sequence processing method in literature [13]. The improved method has the following advantages: firstly, we transform the FCS with uncertain amount and unequal length from the same IP into feature vectors with the same length, and we retain the information of the function codes and their parameters resorting to the frequency characteristics of the short sequence. Secondly, when the length of the short sequence is set to 3, we can process the sequences with unequal length including the length of 1 or 2, by adding the start and the end status.

Parameters sequence (PS) indicates the change rule of the parameters in all the function codes used by the sessions from the same IP, and it is arrayed by chronological order. Similar to FCS, we use the same method to process PS.

5. Partial Seeded K-Means Algorithm

We have tried machine learning methods for malicious IP traceback. Commonly used machine learning methods include decision tree, SVM, and neural network, but all these methods need supervised training samples. But in the homology test of attacking data, the attack source is unknown, and therefore the sample data has no labels. Unsupervised learning can reveal the inherent nature and law of data by learning the unlabeled training samples. Clustering is the most widely used method in unsupervised learning. Clustering is to divide the data samples into multiple classes or clusters, so that the samples in the same cluster have a higher degree of similarity and the samples in different clusters are more different from one another.

K -means [31] algorithm is one of the most classical clustering methods based on partition. The basic idea is to cluster around K points as centers in space, by classifying other samples which are the closest to them. The values of each cluster center are updated iteratively until the best clustering results are obtained. In application, the clustering effect of K -means algorithm is greatly influenced by the initial center selection method.

Considering that the clustering performance can be improved by using labeled samples to assist the initial center selection, Wagstaff et al. [32] proposed the COP K -means algorithm. By constructing the two constraint sets of Must-List and Cannot-Link, the samples were constrained when

they were added to clusters, but the selection of the initial center point was not constrained. Basu et al. [33] proposed seeded/constrained K -means algorithm. It constrained the choices of initial center through seed, and the constraint was also valid when a sample was added into a cluster. However, in this method, each cluster needs a preexisting seed.

In the IP traceback process, it is possible to know that some IP addresses belong to a certain organization. However, it is very hard to know all the organizations in advance. That means that some clusters do not have preexisting seed. Therefore, we designed a partial seeded K -means algorithm to solve this problem (see Algorithm 1).

Partial seeded K -means algorithm utilizes some sample subsets with known cluster partition (which is partial seed) as seed, to determine the initial center point. Considering there may be a variety of attack modes in an organization, constraints on seed are not applied while adding a sample into the clusters. That means the samples with known cluster partition may be classified into the original cluster or a new cluster during the process of clustering. The purpose of using partially seeded clustering here is not to determine whether the other unknown clusters are correct, but to maintain the known clusters being stable while adjusting the parameters of algorithm.

6. Evaluation

6.1. IP Recall Rate of the Known Organizations. We use the IP addresses of the four known organizations to check how many IP addresses of the same organizations are recalled in the same cluster. The four curves in Figures 9–12 show how the recall rate varies with different K values. Apparently, the IP addresses of Shodan, Censys, and Beacon Labs are all grouped into the same cluster, when the cluster number K is set between 20 and 25. However, the highest recall rate of Detecting's IP addresses is about 40%. That means Detecting's IP addresses are divided into different clusters, and there may be multiple attack modes in the samples of Detecting.

6.2. Similarity between the Predicted Value and the True Value. Given the knowledge of the ground truth class assignments labels_{true} and our clustering algorithm assignments of the same samples labels_{pred}, Adjusted Rand Index (ARI) [34] is a function that measures the similarity of the two assignments, ignoring permutations, and with chance normalization. Mutual Information is a function that measures the agreement of the two assignments, ignoring permutations. Adjusted Mutual Information (AMI) is normalized against chance [35].

We use the 66 IP addresses of the known organizations out of 573 valid IP addresses to compare the similarity between the predicted value and the true value. Figure 13 shows how ARI and AMI scores between the predicted and the true values of the 66 IP addresses vary with different K values. Apparently, the clustering works best when the number of clusters K is set between 20 and 29.

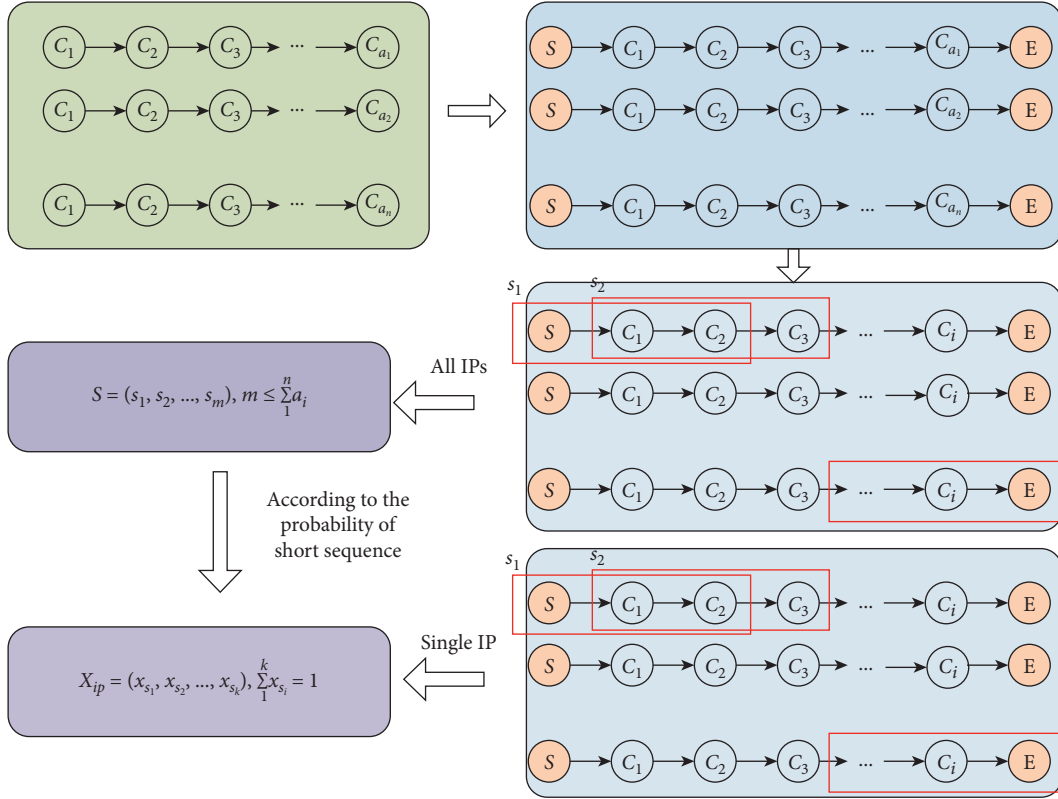


FIGURE 8: Method for FCS feature vector processing.

Input: given a sample set $D = \{x_1, x_2, \dots, x_m\}$, the clustering number k , the known clustering number $l, k \leq l$, the sample subset of known cluster partition $D' = \{x_1, x_2, \dots, x_n\}$, and the sample subset of unknown cluster partition $D - D'$.

- (1) Calculate the mean of the samples in each known cluster $C_i (1 \leq i \leq l)$: $\mu_i = (1/|c_i|) \sum_{x \in c_i} x$.
- (2) Calculate the distance from each sample $x_j (1 \leq j \leq m - n)$ in $D - D'$ to the known mean $\mu_i (1 \leq i \leq l)$, and choose the largest value which equals mean distance added minimum distance as the new initial mean μ_{l+1} and let μ_{l+1} as known mean.
- (3) Repeat Step 2, until $k - l$ samples are chosen as the initial mean vector $\{\mu_{l+1}, \mu_{l+2}, \dots, \mu_{l+k}\}$, make $\mu_i (i \leq l)$ and $\{\mu_{l+1}, \mu_{l+2}, \dots, \mu_{l+k}\}$ to be the initial mean vector with k means.
- (4) Calculate the distance $d_{ij} = \|x_j - \mu_i\|_2$ which is from each sample $x_j (1 \leq j \leq m - n)$ in $D - D'$ to each mean vector $\mu_i (1 \leq i \leq k)$.
- (5) Choose the cluster label for the sample x_j according to nearest initial vector $\lambda_j = \operatorname{argmin}_{i \in \{1, 2, \dots, k\}} d_{ji} (1 \leq j \leq m - n)$, and add x_j into corresponding cluster $C_{\lambda_j} = C_{\lambda_j} \cup \{x_j\}$.
- (6) Calculate new mean vector $\mu'_i = (1/|c_i|) \sum_{x \in c_i} x$, if $\mu'_i \neq \mu_i$ and update μ_i to μ'_i .
- (7) Repeat Steps 4-6, until no mean vector to update.

Output: cluster partition $C = \{C_1, C_2, \dots, C_k\}$.

ALGORITHM 1: Partial seeded K-means.

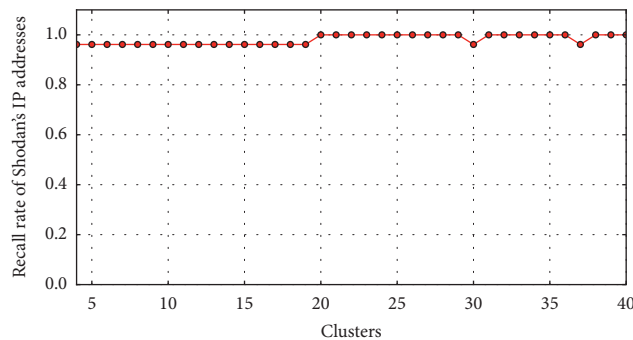


FIGURE 9: The recall rate of Shodan's IP addresses.

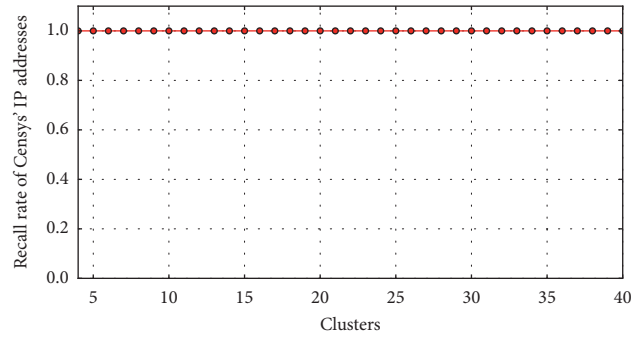


FIGURE 10: The recall rate of Censys' IP addresses.

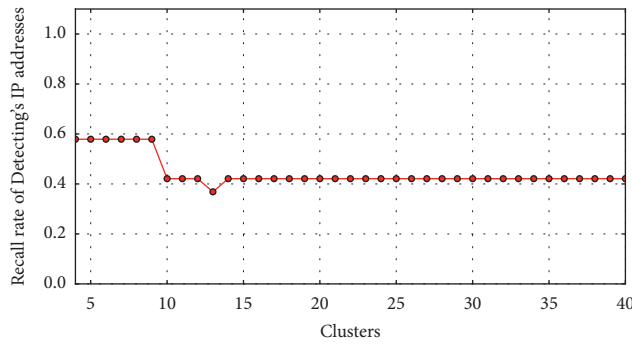


FIGURE 11: The recall rate of Detecting's IP addresses.

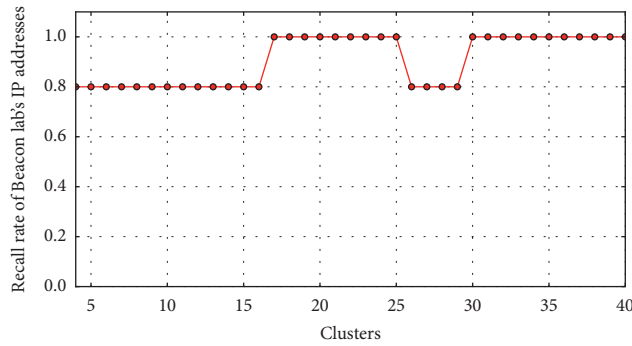


FIGURE 12: The recall rate of Beacon Lab's IP addresses.

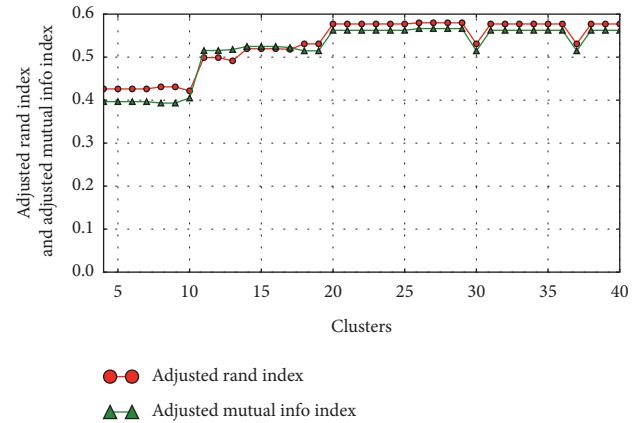


FIGURE 13: ARI and AMI scores between the predicted and the true values of the 66 IP addresses vary with different K values.

6.3. Clustering Performance. In the previous sections, we have evaluated the clustering effect using the samples with known labels. If the ground truth labels are unknown, evaluation must be performed using the model itself. The Silhouette Coefficient [36] is an example of such an evaluation, where a higher Silhouette Coefficient score relates to a model with better defined clusters. Calinski-Harabasz index [37] can be used to evaluate the model too, where a higher Calinski-Harabasz score relates to a model with better defined clusters.

Figures 14 and 15, respectively, show the curves of Silhouette Coefficient score and Calinski-Harabasz score, when the number of clusters K is set differently. Apparently, the clustering works best when K is set to 20.

6.4. Attack Pattern Recognition. Figure 16 shows the total number of clusters, in which those IP addresses of the four known organizations are grouped. No matter what value K is set, the maximum number of clusters is always 6. It indicates that there are only 6 attack patterns at the most in the samples with known organization labels.

The attack pattern of Shodan, Censys, and Beacon Lab is unique, when the cluster number K is set between 20 and 25. But Detecting's attack mode is not unique. All the IP addresses of Detecting belong to three different clusters, except that four IP addresses are labeled as Shodan and two IP addresses are labeled as Censys. The specific distribution of these IP addresses is shown in Figure 17.

6.5. Organization Identification. We set the cluster number K to be 20 for clustering and get 20 clusters at last. That means we find 20 kinds of attack patterns. However, these 20 attack patterns do not indicate that there are 20 organizations because an organization may have multiple attack patterns, and some different organizations may also share a common attack pattern. The DNS query results and the geographical locations of IP Addresses are helpful to identify the organizations. If the IP addresses in a cluster point to the same static domain name or they are very close geographically, we can name this cluster with these labels.

As shown in Table 5, there are 20 clusters with no less than 9 IP addresses in each of them. According to the DNS query results, some IP addresses in clusters 1, 2, 3, and 4 point to a static domain name, and some IP addresses in the clusters 11, 14, and 17 point to a dynamic domain name. There is no domain name for reference in clusters 15, 18, 19 and 20. However, they are located in a particular country or a region, so we can name these clusters with the geographical labels. Furthermore, clusters 3 and 13 are labeled as Detecting, which confirms the existence of multiple attack patterns in a single organization.

7. Related Work

7.1. ICS Intrusion Detection. Khalili and Sami [8] have proposed the SysDetect, which is a systematic approach to Critical State Determination, to solve the problem of determining the critical states in the state-based intrusion

detection. This system built a well-established and iterative data mining algorithm, that is, Apriori. Kwon et al. [9] have proposed a novel behavior-based IDS for IEC 61850 protocol using both statistical analysis of traditional network features and specification-based metrics. Yang et al. [10] have presented a rule-based IDS for IEC 60870-5-104 driven SCADA networks using an in-depth protocol analysis and a Deep Packet Inspection (DPI) method. McParland et al. [11] have proposed the characteristic-based intrusion detection, which is an extension of the specification-based method, by defining a set of good properties and looking for behavior outside those properties. A specification-based intrusion detection model is designed to enhance the protection from both outside attacks and inside mistakes through combining the command sequence with the physical device sensor data. Mo et al. [12] have developed the model-based techniques which are capable of detecting integrity attacks on the sensors of a control system. It is assumed that the attacker wishes to disrupt the operation of a control system in steady state, to which end the attacker hijacks the sensors, observes and records their readings for a certain amount of time, and repeats them afterward to camouflage his attack. The model-based techniques can effectively prevent such attacks. Shang et al. [13] have presented PSO-SVM algorithm which optimizes parameters by advanced Particle Swarm Optimization (PSO) algorithm. The method identifies anomalies of Modbus TCP traffic according to appear frequencies of the mode short sequence of Modbus function code sequence. Zhou et al. [14] have designed a novel multimodel-based anomaly intrusion detection system with embedded intelligence and resilient coordination for the field control system in industrial process automation. In this system, a multimodel anomaly detection method is proposed, and a corresponding intelligent detection algorithm is designed. In addition, in order to overcome the shortcomings of anomaly detection, a classifier based on intelligent hidden Markov model is designed to distinguish the actual attacks and failures.

7.2. IP Traceback. Savage et al. [16] have described a general purpose traceback mechanism based on probabilistic packet marking. Routers probabilistically mark packets with partial path information when they arrive. By combining a modest number of such packets, a victim can reconstruct the entire path. Snoeren et al. [17] have presented a hash-based technique for IP traceback that generates audit trails for traffic within the network and can trace the origin of a single IP packet delivered by the network in the recent past. Belenky and Ansari [20] have proposed a deterministic packet marking algorithm, which only requires the border router to mark the 16-bits Packet ID field and the reserved 1-bit flag in the IP header. Therefore, the victim can obtain the corresponding entry address and the subnet where the attack source is located. This method is simple and efficient compared to Probabilistic Packet Marking algorithm. Belovin et al. [38] have proposed an ICMP Traceback Message. When forwarding packets, routers can, with a low probability, generating a traceback message that is sent along to

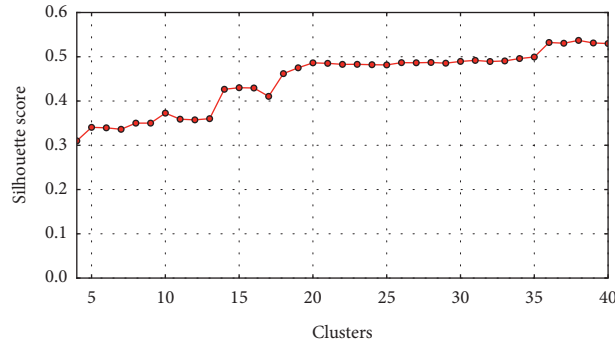


FIGURE 14: Silhouette Coefficient score varies with different K values.

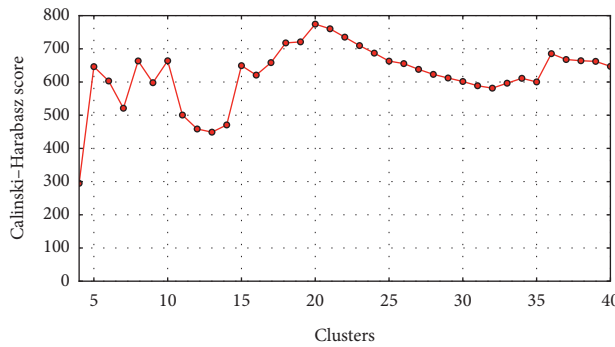


FIGURE 15: Calinski-Harabasz score varies with different K values.

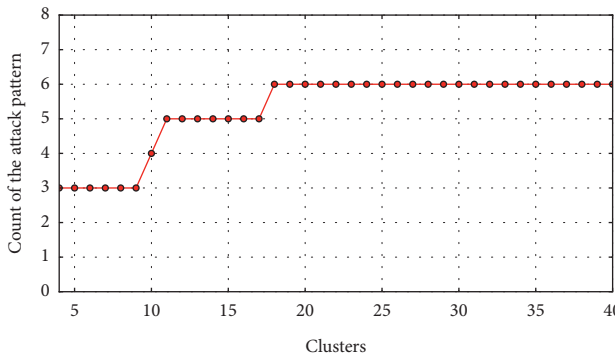


FIGURE 16: The total number of clusters, in which those IP addresses of the four known organizations are grouped.

the destination or back to the source. With enough traceback messages from enough routers along the path, the traffic source and path of forged packets can be determined. Goodrich [39] has presented a new approach to IP traceback based on the probabilistic packet marking paradigm. This approach, which is called randomize-and-link, uses large checksum cords to link message fragments in a way that is highly scalable, for the cords serve both as associative addresses and data integrity verifiers. The main advantage of this approach is that attacker cannot fabricate a message and it has good scalability. Gong and Sarac [18, 19] have presented a novel hybrid IP traceback approach based on both packet logging and packet marking. They maintain the single-packet traceback ability of the hash-based approach

and, at the same time, alleviate the storage overhead and access time requirement for recording packet digests at routers. Their work improves the practicability of single-packet IP traceback by decreasing its overhead. Yang and Yang [40] have proposed a traceback scheme that marks routers' interface numbers and integrates packet logging with a hash table (RIHT) to deal with the logging and marking issues in IP traceback. RIHT has the properties of low storage, high efficiency, zero false positive, and zero false negative rates in attack-path reconstruction. Yu et al. [41] have proposed a marking on demand (MOD) scheme based on the DPM mechanism to dynamically assign marking IDs to DDoS attack related routers to perform the traceback task. They set up a global mark distribution server (MOD server)

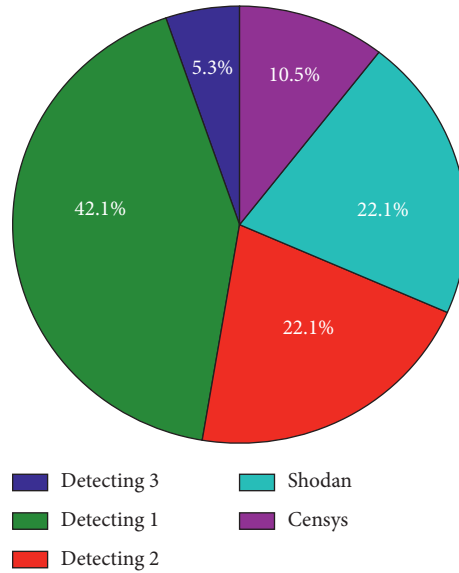


FIGURE 17: Distribution of Detecting's IP addresses.

TABLE 5: Clusters and their labels of organization.

Cluster	IP count	Auxiliary information	Organization
1	93	22 IPs are mapped to the domain name shodan.io	Shodan
2	180	14 IPs are mapped to the domain name eecs.umich.edu	Censys
3	75	8 IPs are mapped to the domain name neu.edu.cn	Detecting
4	43	5 IPs are mapped to the domain name plscan.org	Beacon Lab
11	51	26 IPs are mapped to the dynamic domain name binaryedge.ninja	binaryedge.ninja
13	11	4 IPs are mapped to the domain name neu.edu.cn	Detecting
14	17	6 IPs are mapped to the dynamic domain name amazonaws.com	amazonaws.com
15	20	17 IPs are located in China	China Org
17	35	25 IPs are mapped to the dynamic domain name members.linode.com	linode.com
18	14	11 IPs are located in China	China Org
19	14	12 IPs are located in Europe	Europe Org
20	9	7 IPs are located in China	China Org

and some local DDoS attack detector. When there appear suspicious network flows, the detector requests unique IDs from the MOD server and embeds the assigned unique IDs to mark the suspicious flows. At the same time, the MOD server deposits the IP address of the request router and the assigned marks, which are used to identify the IP addresses of the attack sources, respectively, into its MOD database. Fadel et al. [42] have presented a new hybrid IP traceback framework. This framework is based on both marking and logging techniques. In the marking algorithm, every router is assigned a 12-bit-length ID number; it helps in deploying pushback method to permit legitimate traffic flow smoothly. In the packet logging technique, a logging ratio is managed by changing a value k specified in the traceback system. This framework can save more than 50% of the storage space of routers. Cheng et al. [43] argue that cloud services offer better options for the practical deployment of an IP traceback system. They have presented a novel cloud-based traceback architecture, which possesses several favorable properties encouraging ISPs to deploy traceback services on their networks. This architecture includes a temporal token-based authentication framework, called FACT, for

authenticating traceback service queries. Nur and Tozal [44] exploit the record route feature of the IP protocol and propose a novel probabilistic packet marking scheme to infer forward paths from attacker sites to a victim site and enable the victim to delegate the defense to the upstream Internet Service Providers (ISPs). Compared to the other techniques, this approach requires less many packets to construct the paths from attacker sites toward a victim site.

8. Conclusions

IP traceback for cyberattacks usually needs redesigning the Internet deploying new service. In this study, we have proposed a malicious IP traceback model, that is, ICSTrace, for Industrial Control System without changing the Internet infrastructure or deploying any new services. By analyzing the characteristics of the attack data, we extract the numeric features and the sequence transformation features from the function codes and their parameters. Those features are expressed by a one-dimensional vector, which stands for the unique pattern of an attack. As a result, the problem of IP traceback turns into a problem of clustering those patterns.

We also propose a partial seeded K -means algorithm to cluster the IP addresses with the same pattern into a malicious organization. The effectiveness of ICSTrace is proved by experiments on real attack data. Although ICSTrace cannot recover the whole path of the attack, it is significant in the following aspects:

- (1) Finding out the malicious IP addresses which belong to the same organization
- (2) Revealing the unexposed active IP addresses belonging to the known organizations
- (3) Collecting the anonymous communication networks used by the same organization for launching attacks
- (4) Providing learning samples for subsequent malicious behavior identification by expressing the attack pattern in the form of feature vector

As we know, the concealing methods of attackers are becoming more complex. As a result, it is very difficult to trace the original IP of the attacker directly. The model proposed in our manuscript is helpful for the security experts in an indirect way of tracing the last IPs of attacks which belong to the same attacker. Therefore, we define this model as an IP traceback model.

Data Availability

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

Additional Points

Future Work. In the future, we will improve ICSTrace and apply it to other kinds of ICS protocols, even the traditional internet protocols. At the same time, we will use the attack patterns as the learning samples to design and validate the intrusion detection system based on machine learning to solve the difficult problem of unknown threat detection.

Disclosure

This work has been presented by the authors themselves as arXiv in Cornell University according to <https://arxiv.org/abs/1912.12828>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank Biao Chang, Binglei Wang, and Dazhong Shen for their useful feedback and comments.

References

- [1] T. M. Chen and S. Abu-Nimeh, "Lessons from stuxnet," *Computer*, vol. 44, no. 4, pp. 91–93, 2011.
- [2] D. Kushner, "The real story of stuxnet," *IEEE Spectrum*, vol. 50, no. 3, pp. 48–53, 2013.
- [3] K. Zetter, "A cyberattack has caused confirmed physical damage for the second time ever," 2015, <https://www.wired.com/2015/01/german-steel-mill-hack-destruction>.
- [4] K. Zetter, "Inside the cunning unprecedented hack of Ukraine's power grid," 2016, <http://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/>.
- [5] ESET discovers dangerous malware designed to disrupt industrial control systems. <https://www.eset.com/us/about/newsroom/press-releases/eset-discovers-dangerous-malware-designed-to-disrupt-industrial-control-systems/>.
- [6] K. Stouffer, J. Falco, and K. Scarfone, *Guide to Industrial Control Systems (ICS) Security*, Vol. 800, NIST special publication, Gaithersburg, MD, USA, 2011.
- [7] W. Knowles, D. Prince, D. Hutchison, J. F. P. Disso, and K. Jones, "A survey of cyber security management in industrial control systems," *International Journal of Critical Infrastructure Protection*, vol. 9, pp. 52–80, 2015.
- [8] A. Khalili and A. Sami, "Sysdetect: a systematic approach to critical state determination for industrial intrusion detection systems using Apriori algorithm," *Journal of Process Control*, vol. 32, pp. 154–160, 2015.
- [9] Y. Kwon, H. K. Kim, Y. H. Lim, and J. I. Lim, "A behavior-based intrusion detection technique for smart grid infrastructure," in *Proceedings of the 2015 IEEE Eindhoven PowerTech*, 2015.
- [10] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, and H. Wang, "Rule-based intrusion detection system for SCADA networks," in *Proceedings of the 2nd IET Renewable Power Generation Conference (RPG 2013)*, Beijing, China, 2013.
- [11] C. McParland, S. Peisert, and A. Scaglione, "Monitoring security of networked control systems: it's the physics," *IEEE Security & Privacy*, vol. 12, no. 6, pp. 32–39, 2014.
- [12] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1396–1407, 2014.
- [13] W. L. Shang, S. S. Zhang, and M. Wan, "Modbus/TCP communication anomaly detection based on PSO-SVM," *Applied Mechanics and Materials*, vol. 490–491, pp. 1745–1753, 2014.
- [14] C. Zhou, S. Huang, N. Xiong et al., "Design and analysis of multimodel-based anomaly intrusion detection systems in industrial process automation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 10, pp. 1345–1360, 2015.
- [15] T. Rid and B. Buchanan, "Attributing cyber attacks," *Journal of Strategic Studies*, vol. 38, no. 1–2, pp. 4–37, 2015.
- [16] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Practical network support for IP traceback," *ACM SIGCOMM Computer Communication Review*, vol. 30, no. 4, pp. 295–306, 2000.
- [17] A. C. Snoeren, C. Partridge, L. A. Sanchez et al., "Hash-based IP traceback," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 4, pp. 3–14, 2001.
- [18] C. Gong and K. Sarac, "IP traceback based on packet marking and logging," in *Proceedings of the 2005 IEEE International Conference on Communications*, vol. 2, Seoul, Republic of Korea, 2005.
- [19] C. Gong and K. Sarac, "A more practical approach for single-packet IP traceback using packet logging and marking," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 10, pp. 1310–1324, 2008.
- [20] A. Belenky and N. Ansari, "IP traceback with deterministic packet marking," *IEEE Communications Letters*, vol. 7, no. 4, pp. 162–164, 2003.

- [21] S7 Communication (S7comm), <https://wiki.wireshark.org/S7comm>.
- [22] F. Xiao, E. Chen, and Q. Xu, "S7commTrace: a high interactive honeypot for industrial control system based on s7 protocol," in *Proceedings of the 2017 International Conference on Information and Communications Security*, pp. 368–380, Beijing, China, 2017.
- [23] A. Jicha, M. Patton, and H. Chen, "SCADA honeypots: an in-depth analysis of Conpot," in *Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, pp. 196–198, Tucson, AZ, USA, 2016.
- [24] L. Spitzner, *Honeypots: Tracking Hackers*, Vol. 1, Addison-Wesley Reading, Boston, MA, USA, 2003.
- [25] J.-W. Zhuge, Y. Tang, X.-H. Han, and H.-X. Duan, "Honeypot technology research and application," *Journal of Software*, vol. 24, no. 4, pp. 825–842, 2013.
- [26] Shodan. <https://www.shodan.io/>.
- [27] Censys. <https://censys.io/>.
- [28] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, "A search engine backed by internet-wide scanning," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 542–553, Denver, CO, USA, 2015.
- [29] Ditecting. <http://www.ditecting.com/>.
- [30] ICS Security Workspace. <http://plcscan.org/blog/>.
- [31] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [32] K. Wagstaff, C. Cardie, S. Rogers, and S. Schrödl, "Constrained k-means clustering with background knowledge," in *Proceedings of the 18th International Conference on Machine Learning*, vol. 1, Williamstown, MA, USA, 2001.
- [33] S. Basu, A. Banerjee, and R. Mooney, "Semi-supervised clustering by seeding," in *Proceedings of 19th International Conference on Machine Learning (ICML-2002)*, Sydney, Australia, 2002.
- [34] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985.
- [35] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance," *Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, 2010.
- [36] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.
- [37] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics—Theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
- [38] S. M. Bellovin, M. Leech, and T. Taylor, *ICMP Traceback Messages*, Internet Engineering Task Force (IETF), Fremont, CA, USA, 2003.
- [39] M. T. Goodrich, "Probabilistic packet marking for large-scale IP traceback," *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 15–24, 2008.
- [40] M.-H. Yang and M.-C. Yang, "RIHT: a novel hybrid IP traceback scheme," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 789–797, 2012.
- [41] S. Yu, W. Zhou, S. Guo, and M. Guo, "A feasible IP traceback framework through dynamic deterministic packet marking," *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1418–1427, 2016.
- [42] M. M. Fadel, A. I. El-Desoky, A. Y. Haikel, and L. M. Labib, "A low-storage precise IP traceback technique based on packet marking and logging," *The Computer Journal*, vol. 59, no. 11, pp. 1581–1592, 2016.
- [43] L. Cheng, D. M. Divakaran, A. W. K. Ang, W. Y. Lim, and V. L. L. Thing, "FACT: a framework for authentication in cloud-based IP traceback," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 604–616, 2017.
- [44] A. Y. Nur and M. E. Tozal, "Record route IP traceback: combating dos attacks and the variants," *Computers & Security*, vol. 72, pp. 13–25, 2018.

Research Article

GNFCVulFinder: NDEF Vulnerability Discovering for NFC-Enabled Smart Mobile Devices Based on Fuzzing

Zhiqiang Wang ^{1,2}, Yuheng Lin ¹, Zihan Zhuo ³, Jieming Gu ³ and Tao Yang⁴

¹Beijing Electronic Science and Technology Institute, Cyberspace Security Department, Beijing 100070, China

²State Information Center, Post-Doctoral Scientific Research Workstation, Beijing 100045, China

³National Internet Emergency Center, Beijing 100029, China

⁴Key Lab of Information Network Security, Ministry of Public Security, Shanghai 200031, China

Correspondence should be addressed to Zhiqiang Wang; wangzq@besti.edu.cn and Zihan Zhuo; zzh@cert.org.cn

Received 29 March 2021; Accepted 19 June 2021; Published 28 June 2021

Academic Editor: Jinguang Han

Copyright © 2021 Zhiqiang 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.

Near-field communication (NFC) is a set of communication protocols that enable two electronic devices. Its security and reliability are welcomed by mobile terminal manufactures, banks, telecom operators, and third-party payment platforms. Simultaneously, it has also drawn more and more attention from hackers and attackers, and NFC-enabled devices are facing increasing threats. To improve the security of the NFC technology, the paper studied the technology of discovering security vulnerabilities of NFC Data Exchange Format (NDEF), the most important data transmission protocol. In the paper, we proposed an algorithm, GTCT (General Test Case Construction and Test), based on fuzzing to construct test cases and test the NDEF protocol. GTCT adopts four strategies to construct test cases, manual, generation, mutation, and “reverse analysis,” which can detect logic vulnerabilities that fuzzing cannot find and improve the detection rate. Based on GTCT, we designed an NDEF vulnerability discovering framework and developed a tool named “GNFCVulFinder” (General NFC Vulnerability Finder). By testing 33 NFC system services and applications on Android and Windows Phones, we found eight vulnerabilities, including DoS vulnerabilities of NFC service, logic vulnerabilities about opening Bluetooth/Wi-Fi/torch, design flaws about the black screen, and DoS of NFC applications. Finally, we give some security suggestions for the developer to enhance the security of NFC.

1. Introduction

NFC (near-field communication) is a set of ideas and technologies that enable smart phones and other devices to establish radio communication with each other by touching the devices together or bringing them in proximity to a distance of typically 10 cm or less [1–4]. It evolved from the integration of RFID and interconnection technology, which can realize two-way interactive communication between electronic devices. Devices using NFC technology (such as mobile phones) can exchange data when they are close to each other. NFC has many characteristics like short communication distance, one-to-one device connection, and hardware encryption, and it has excellent security and reliability. With the development of mobile Internet and

mobile payment, more and more smartphones begin to support the NFC function [5]. NFC payment has become a promising means of mobile payment favored by banks, telecom operators, and mobile phone manufacturers.

However, NFC has also drawn more and more attention from researchers and attackers, and the security of NFC is facing increasing threats and challenges. On February 9, 2012, Zvelo Labs found a security vulnerability in Google Wallet [6], a mobile payment system based on NFC. Knowing that the PIN could only be a 4-digit numeric value, they could get the PIN easily by a brute-force attack, but the premise of this attack was that the attacked phone was rooted. The next day, the SmartPhone Champ found another vulnerability; the vulnerability could calculate the PIN in Google Wallet regardless of whether the phone is rooted or

not. The new PIN could be acquired by going into the application menu to clear the Google Wallet app's date and resetting a new PIN by opening it again [7]. In July 2012, Charlie Miller, the chief research consultant of Accuvant Labs, exposed many vulnerabilities in the NFC protocol stack at Blackhat USA [8]. For example, attackers can automatically cause the affected users to open a malicious URL and embed the malware in victims by exploiting these bugs with an NFC tag. At EUsecWest in September 2012, Corey Benninger and Max Sobell from Intrepidus Group were able to quickly and easily reset the number of journeys on the NFC contactless travel card using an NFC-enabled Android phone. This NFC payment flaw existed in the MIFARE Ultralight chips and it could be used to reset certain data on the NFC contactless card. This defect resulted in the provision of free subway tickets, which affected several cities in subway systems, such as Boston, Philadelphia, and Chicago [9]. In August 2013, the Wall of Sheep revealed that the prepared malware in advance could be installed in attacked phones that touched a crafted malicious NFC tag. This vulnerability, which can copy users' short messages in their phones, is discovered in their NFC Security Awareness Project [10]. In July 2015, Rob Miller and Jon Butler of MWR Labs found that remote attackers can attack Samsung NFC-enabled mobile phones to download any or all images on the vulnerable devices without any notifications or user interactions [11]. In the same year, a video [12] published by a group of hackers demonstrated how a thief could use a stolen Apple Watch to make payments using Apple Pay without authenticating the transactions in any way. In May 2016, Martijn Coenen found that sensitive foreground-application information can be obtained by attackers via a crafted background application, which caused NfcService in NFC-enabled phones. In August 2016, security researcher Salvador Mendoza demonstrated a flaw in Samsung Pay at Black Hat [13], in which the transaction tokens could be predicted and be used to authorize fraudulent payments. In 2017, Xinyi Chen et al. found that attackers can use the vulnerabilities of the card payment transactions to compromise the NFC communication message and then transmit the wrong payment information to the communicators [14]. In 2020, S. Akter et al. identified a potential vulnerability in existing contactless payment protocols due to the separation between the card authentication and the transaction authorization phase. They showed how an attacker could compromise the integrity of contactless payments by a malicious MITM smartcard [15].

For all the NFC-related vulnerabilities and security incidents mentioned above, the fundamental and critical things are security vulnerabilities, which are the source of network attacks and defenses. Therefore, the paper mainly studied the technology to discover NFC security vulnerabilities effectively. The main works and contributions of the paper are as follows:

- (i) The test case construction combines manual, generation-based, and mutation-based strategies, which takes advantage of three strategies and overcomes their disadvantages. Also, the known NFC

vulnerability data and abnormal data are adopted to construct test cases. This method solves the single data construction strategy problem and enhances the positive detection rate of vulnerabilities.

- (ii) To achieve the automation of testing and the mobile operating system's independence, we adopted two manners: simulating tag with the NFC reader device and simulating "touch" by operating the NFC process. Besides, we adopted "logcat/Xapsy" and monitoring process to monitor the abnormal operations and other problems.
- (iii) For the NFC Data Exchange Format (NDEF) protocol, we designed and constructed a test case database, which can be used to test all kinds of NFC-related systems and applications, saving time and costs greatly.
- (iv) With the database, we tested lots of NFC system applications and third-party applications on the Android and Windows Phone platforms and found many known and unknown vulnerabilities, including "Wi-Fi on," "Bluetooth on," "black screen," "Flashlight on," and DoS vulnerabilities.

The remainder of this paper is organized as follows. Section 2 surveys related work. Section 3 introduces the proposed methodology. Section 4 illustrates the architecture and implementation of our system used to discover NFC vulnerabilities. Section 5 shows the experimental configurations and environments. Section 6 provides the results in detail. Section 7 evaluates the results by comparing the related works, and Section 8 draws conclusions.

2. Related Work

Currently, many researchers and scholars at home and abroad have done lots of studies on vulnerability discovery and defense security strategies of NFC. The research progress of NFC security will be reviewed as follows.

Ernst Haselsteiner and Klemens Breitfu [16] gave a comprehensive analysis of NFC security, including the various aspects of NFC security threats and related solutions to protect against these threats. This paper provides an overview of the traditional NFC security, including eavesdropping, data corruption, data modification, data insertion, man-in-the-middle attack, and secure channels for NFC, which emphasizes communication security. Collin Mulliner put forward a method to perform vulnerabilities analysis of NFC-enabled mobile phones by fuzzing applications with NFC tags [17]. They analyzed NFC subsystem and components (such as the web browser) and they found several vulnerabilities which can be abused for phishing, NFC-based worms, and DoS attacks. Their method constructs test cases and tests NFC-enabled phones manually, which does not support automatic testing, so the method is time-consuming and manpower-consuming. Gauthier et al. [18] presented an NFC offline payment application and applied it to Nokia 6313/6312. The PKI-based security protocol is used to create a secure channel. Their research is done from the viewpoint

of defense strategies. However, data encryptions will increase the time costs. Antonio et al. [19] evaluated the security capabilities of NFC-powered devices to carry out asymmetric cipher based on the public key encryption algorithm RSA. They proposed a hybrid security scheme that combines asymmetric key encryption with a shared key and asymmetric encryption algorithm. The drawbacks of their scheme are the long size of the key and the time to generate the certificate. Jia and Tong [20] adopted threat modeling methodologies to analyze the threats that may damage the assets from the entries during the mobile payment, which are depicted in the data flow diagram and illustrated by some attack scenarios. They had proposed some migration solutions to these threats, respectively. Their works remain theoretical and need to be verified in the future. Miller [8] proposed a method for fuzzing the NFC protocol stack and found many NFC vulnerabilities. Miller constructed test cases based on the tool Sulley and adopted a single strategy for constructing test cases and also used a single monitor, “logcat,” which may cause a high false-negative rate. Omkar and Hegde [21] analyzed the threats faced by Google Wallet and evaluated its security measures. Their method is based on the existing mobile payment solutions, including the embedded solutions, the SIM-based solutions, and the card-based solution. This paper also analyzed and evaluated the security of NFC payment from a theoretical point of view. Norbert [22] developed a framework called “fuzzing-to-go” using a fuzzing approach for Android NFC APIs and NFC applications, in which test sets were generated by the tool Sulley. The monitor in fuzzing-to-go was also provided by Sulley. This framework’s monitors and data generation strategy are relatively simple and do not output error logs in detail. Besides, the framework is developed to aim for Android OS and specially appointed API version. Thus, its portability is poor. Gummeson et al. designed a passive hardware-based patch called “Engarde” [23] at extremely low power to protect NFC interaction. Engarde can be stuck on the back of an NFC phone and intercept malicious operations and behaviors in the specific blacklisted behaviors. Their solution needs additional hardware, which increases maintenance overhead and hardware cost. Roland [24] designed two scenarios to emulate a secure element of the Android platform. The secure element emulator can be used for debugging and rapid prototyping of secure element applications. Application developers can use it to replace the secure elements for long-term testing, which reduces the development cost and complexity of SE applications.

With the rise of mobile payment, NFC security faces more and more challenges. However, there are many issues and problems in the researches on NFC security. As can be seen from the previous related research works, the research works [16, 20, 21] are purely theoretical and will await further evaluation. The research works [18, 19, 23] propose some NFC security strategies from the point of software or hardware protection. However, the disadvantages are that these strategies will increase time overhead and hardware costs. The research works [17, 22] have some disadvantages, including low automatization, a single strategy for constructing test cases, lack of monitors, and poor extensibility.

Reference [24] develops two scenarios for the open platforms emulating a security element of the Android platform, which is helpful to NFC security.

As can be seen from NFC security research progress, the security of the NFC Data Exchange Format protocol has drawn increased attention from security researchers [8, 17, 22]. Security researchers have done a lot of research work and made some progress. However, there are still many problems with NDEF security that need to be solved. NDEF is a standardized data format maintained by the NFC Forum [25] which can exchange information between any compatible NFC device and another NFC device or tag. The research work in [17] constructs test cases and test targets manually, which are of low automation degree and cost much time and labor power. The research work in [8] constructs test cases with a single strategy, namely, generation-based or mutation-based, and monitors targets with a simple “logcat” that will cause a high false-negative rate. The research work in [22] adopts Sulley’s process monitoring module. The monitoring effect is not good, and there are no detailed logs available to analyze exceptions. In addition, the study in [22] strongly depends on the versions of the Android operating system and APIs, causing poor extensibility.

Aiming at the problems existing in NFC security research, we proposed and designed an NFC application security system called GNFCVulFinder (General NFC Vulnerability Finder) to study the technologies of discovering vulnerabilities, whose strategies of constructing test cases adopt a generation-based manner combined with the mutation-based manner and manual test.

3. Methodology

In this section, the methodology of vulnerabilities discovery will be introduced.

3.1. Outline of the Methodology. Our methodology adopts three technologies to find vulnerabilities: fuzzing test [26–31], manual test, and “reverse analysis.” The combined use of these three methods can make up for each other’s shortcomings and improve efficiency and effectiveness. Fuzzing is a method of finding software vulnerabilities by providing unexpected input to the target system and monitoring the results. It is between a complete manual test and a fully automated test. It is an effective automatic test method, which can improve test efficiency and greatly reduce testing costs. However, it has several shortcomings, including low code coverage and inability to discover logic vulnerabilities. Manual test is a traditional technique for software testing used in our methodology to detect logic vulnerabilities. The complete manual testing is penetration testing. Testers simulate hackers maliciously entering the system and find loopholes. This method completely depends on the testers’ ability. It can make up for fuzzing’s deficiencies. In addition, “reverse analysis,” not the traditional meaning of reverse engineering, is also used in our methodology. Some NDEF messages are defined by third-party developers; they do not strictly follow the NDEF

specifications, so these messages will be obscure. These proprietary protocols lead to some difficulties in constructing test cases. Therefore, we have adopted some techniques, including “write in and read back” and “sniffer,” to analyze NDEF messages using binary editors, which will improve the recognition rate of test case. We call these techniques “reverse analysis.” The procedure of “reverse analysis” will be introduced in Section 4.4.

The methodology is illustrated in Figure 1. We first analyze the vulnerabilities in the tested protocols and then construct test cases based on fuzzing, manual testing, and “reverse analysis.” Afterward, we test targets with test cases and monitor the targets. Finally, we validate the exceptions and output test results. In our method, the key algorithm is to construct test cases, which will be introduced in the next section.

3.2. The Algorithm of Test Case Construction and Test. The construction and testing of test cases are the core points of our methodology, as shown in Figure 2. We call this algorithm GTCT (General Test Case Construction and Test). GTCT is described as follows. The input includes \mathbf{G} , \mathbf{M} , S_1 , and S_2 .

\mathbf{G} is a generator matrix that denotes a flag whether the corresponding protocol field will be filled with the generated data or not. $\mathbf{G} = (g_{ij})_{n \times r}$, $g_{ij} = 0$ or 1 , $1 \leq i \leq n, 1 \leq j \leq r$. $g_{ij} = 1$ means generating test cases with malformed data MD_j at the field f_i , and equaling 0 means the opposite. MD_j is a column vector in \mathbf{MDB} which is a database of malformed data fragments. $\mathbf{MDB} = \{MD_1, MD_2, \dots, MD_r\}$; MD_i denotes a type of malformed data in database \mathbf{MDB} , $1 \leq i \leq r$. f_i is a weak protocol field in a vulnerable field set \mathbf{F} . $\mathbf{F} = \{f_1, f_2, \dots, f_n\}$; f_i denotes a weak field in network protocols, $1 \leq i \leq n$.

\mathbf{M} is a mutation matrix that denotes a flag whether the corresponding protocol field will be mutated randomly or not based on a data sample, $\mathbf{M} = (m_{ij})_{n \times q}$, $m_{ij} = 0$ or 1 , $1 \leq i \leq n, 1 \leq j \leq q$. $m_{ij} = 1$ means mutating f_i field of S_j randomly, and equaling 0 means the opposite.

S_1 and S_2 are the column vectors in \mathbf{SDB} , which are a set of sample data. $\mathbf{SDB} = \{S_1, S_2\} = \{s_1, s_2, \dots, s_q\}$, where q is the number of samples. S_1 denotes sample data from the process of manual analysis and testing process, and S_2 denotes sample data collected from NVD, CVE, and other databases (see Algorithm 1).

Firstly, Initialize() is used to initialize mobile phones and applications. Secondly, the strategies for constructing test cases in the algorithm include three steps: manually construct test cases, generate vulnerable protocol field F_{gen} , and mutate samples' vulnerable fields F_{mut} . The function Generator ($S_{\text{RFC}}, \mathbf{G}, \mathbf{MDB}$) denotes the generation of test cases based on matrices \mathbf{G} , \mathbf{MDB} , and the protocols' fields S_{RFC} based on the RFC specifications. The function Mutator ($\mathbf{SDB}, \mathbf{M}, \text{rand}$) denotes the mutation samples of the \mathbf{SDB} based on \mathbf{M} with a random function. In addition, the “reverse analysis” on the protocols is used to build a sample database \mathbf{SDB} . Thirdly, execute the test cases T_{mau} (test cases constructed by manual strategy), T_{gen} (test cases constructed by multiple-dimension strategy based on the generation),

and T_{mut} (test cases constructed by multiple-dimension strategy based on mutation) with the test function NdefTest(). NdefTest() is a test function; this function sends some data for the mobile phone to read the data for testing. By the way, S_1 is obtained during the previous test, and S_2 is from manual analysis on some vulnerability database, such as NVD. Fourthly, execute the test cases T_{mut} . Finally, output logs and results. The details and advantages of the three strategies will be described in Section 4.3.

4. Design and Implementations of System Architecture

This section will introduce the system architecture design of our vulnerability discovery framework and its implementation process.

4.1. System Architecture. We designed a discovering vulnerability system named GNFCVulFinder for the NDEF protocol, and its architecture is shown in Figure 3. The system architecture includes test case generation, initialization, exception monitor, NFC apps test, exceptions validation, and log output. The architecture contains two entities: the mobile phone and the NFC tag. These two entities do not need to be authenticated during the test. The NFC communication process uses the NDEF protocol, which is used between the NFC mobile phone and the tag and does not require identity verification.

Testcase generation: this module is used to construct test cases about the protocol NDEF, including NFC Forum type, NFC external type, absolute URI type, and MIME type. The strategies for constructing test cases have been introduced in detail in the section titled “Test Case Construction.”

Initialization: since two or more NFC applications will listen to NDEF message events and respond to them simultaneously, mobile phone users have to select an application to handle the events, which will cause a serious disturbance to automated testing. So, we need to initialize the tested mobile phone, namely, installing the tested application and uninstalling other unconcerned applications.

NFC apps test: this module is designed to test NFC services and Android and Windows Phone applications. On Android, emulating tags and “touch” operations are adopted to realize automatic tests. The former is achieved by calling APIs in the open-source library libnfc based on the NFC device ACR 122U, and the latter is achieved by controlling the NFC process with the commands “kill” and “start.” On Windows Phone, we only need to emulate a tag, because Windows Phone can automatically detect the changes of data in tags and automatically read the data.

Exception monitor: this module adopts monitoring abnormal logs and monitoring NFC services and logs the test cases triggering abnormal conditions and the breakpoints.

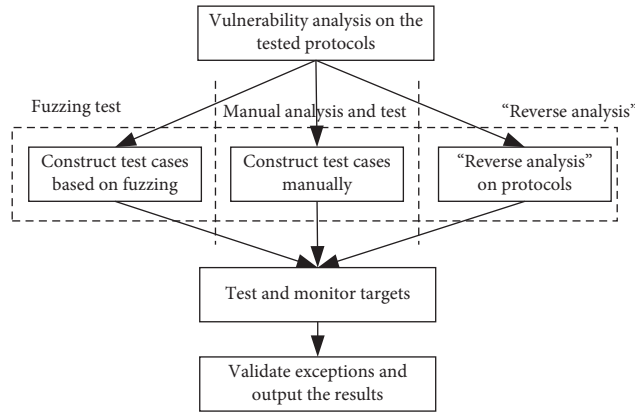


FIGURE 1: The illustration of our methodology.

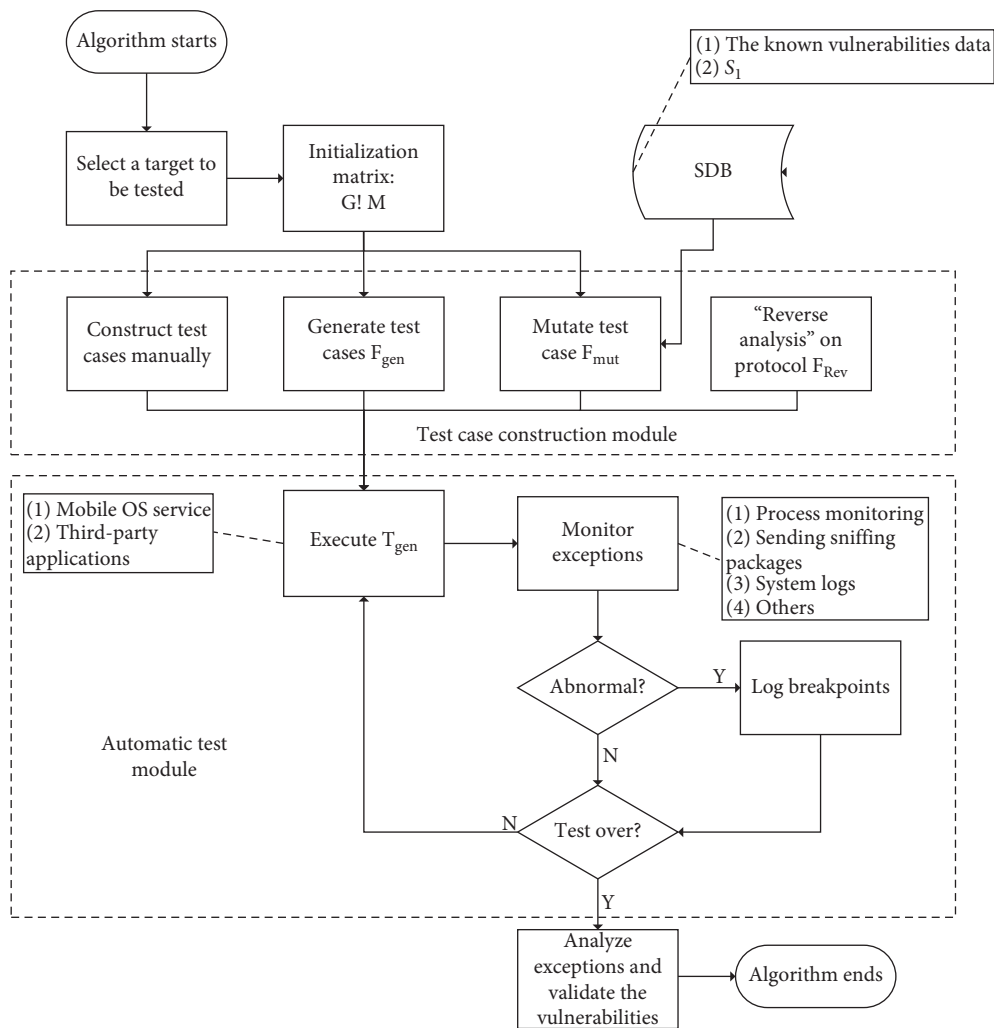
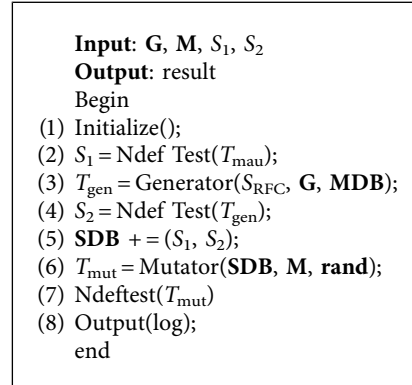


FIGURE 2: The flowchart of Generalized Test Case Construction and Test.

Exception validation: this module validates the exceptions found in the process of a test by resending the test cases and monitoring the targets. The exceptions, if present, will be analyzed manually and verified whether they could be exploited.

Log output: this module will output the results of the test. The system architecture is independent of any mobile platform because the emulated NFC tag does not depend on the mobile phone. Therefore, the architecture is suitable for testing on



ALGORITHM 1: The algorithm of GTCT.

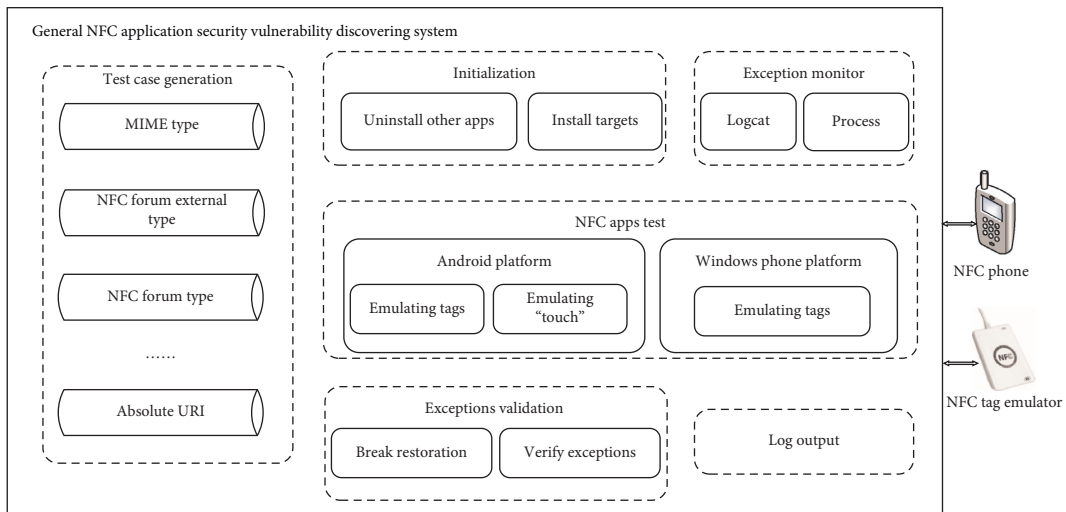


FIGURE 3: The architecture of GNFCVulFinder.

Android, Windows Phone, and other mobile OS platforms.

The following sections will describe the implementation of some important modules in detail, including “test case construction” and “test and monitor.”

4.2. Test Case Construction. NDEF is the NFC Data Exchange Format, which is the data format of the NFC tag agreed by the NFC organization. It is a lightweight and compact binary format with various data types defined by URL, vCard, and NFC. This paper will not detail the detailed definitions of the protocol NDEF, which can be acquired in the NFC Forum. The following will introduce the strategies of test case construction about fuzzing and manual test.

4.3. Strategies

4.3.1. Manual Strategy. Manual strategy constructs test cases by analyzing the NFC forum’s NDEF specifications and NDEF vulnerabilities in the popular vulnerability databases. The strategy can improve the validity of test cases,

reverify known vulnerabilities, and find some logical vulnerabilities that fuzzing cannot detect. By searching the well-known database NVD [32], we analyzed the following several typical vulnerabilities:

CVE-2008-5825: the SmartPoster implementation on the NFC mobile phone Nokia 6131 does not properly display the URI records when the title record contains a certain combination of space(“0×20”), CR(“0x0D”), and (dot, “0x2E”) characters, which allows remote attackers to trick a user into loading an arbitrary URI via a crafted NDEF tag. We can test NFC applications with an NFC tag demonstrated by (a) an http: URI for a malicious website, (b) a tel: URI for a premium-rate telephone number, and (c) an sms: URI that triggers the purchase of a ringtone.

CVE-2008-5826: the NFC phone Nokia 6131 allows attackers to cause a denial of service (device crash) via a large value in the payload length. So, we should test NDEF applications with a large VALUE in the payload length field of an NDEF record or a certain length for a tel or sms: NDEF URI.

CVE-2008-5827: Nokia 6131 automatically installs the software upon completing the download of a JAR file,

making it easier for remote attackers to execute arbitrary code via a crafted URI record in an NDEF tag. We should test it by making the NFC service complete the download of an APK file.

CVE-2015-8041: multi-integer overflows in the NDEF record parser in hostapd before 2.5 and wpa_supplicant before 2.5 allow remote attackers to cause a denial of service (process crash or infinite loop) via a large payload length field value in a WPS or P2P NFC NDEF record, which triggers an out-of-bounds read. So, we can test NFC applications with large payload length values.

CVE-2017-7461: it is a directory traversal vulnerability in the web-based management site on the Intellinet NFC-30ir IP Camera and allows remote attackers to read arbitrary files via a request to a vendor-supplied CGI script that is used to read an HTML text file, but that does not do any URI/path sanitization. So, we can test targets with directory traversal strings.

4.3.2. Multiple-Dimension Strategy Based on the Generation.

This strategy first acquires the NFC Forum's specifications and analyzes the format of NDEF messages. The test cases are constructed by adding one or more fields of NDEF with malformed data and other fields filled with normal data. As shown in Table 1, we designed a malformed database to test fields, including header, length, type, and payload. The database consists of integers, formatted strings, directory traversal data, separators, and other nonalphanumeric characters. An example is illustrated in Figure 4. This message is constructed by generating data for four fields, and other fields are filled with normal data.

4.3.3. Multiple-Dimension Strategy Based on Mutation.

This strategy does not need to analyze NDEF specifications; it only needs to construct test cases by mutating one or more bits of selected samples. Since the previous vulnerability data is likely to trigger an old or new bug [31], the known vulnerability data and test cases that trigger exceptions during the test process are used as mutation samples. We can construct test cases by mutating one or more bits of the "local name field" of the Bluetooth pairing protocol.

4.4. *Reverse Analysis.* During the test process, NDEF messages defined by developers in different NFC applications did not follow the NDEF specifications in the NFC Forum. For example, the messages constructed by an NFC application called "YunNFC" cannot be identified by MIUI OS, and MIUI just opens the site of "YunNFC." As a result, lots of test cases may be denied by tested targets. To solve this problem, "reverse analysis" is adopted to analyze the format of NFC application messages. "Reverse analysis" helps to improve the validity of test cases, and it consists of five steps.

Step 1: construct NDEF message with NFC applications, such as Detail!, TagInfo, and TagWriter. For example, we can construct a Bluetooth pairing message just by inputting the information about Bluetooth.

Step 2: write the constructed messages into an NFC tag.

Step 3: "touch" an NFC mobile phone with the tag. The binary file about the message will be stored on the phone.

Step 4: export the binary file by some tools about phone managers or the command "adb pull."

Step 5: analyze the binary file based on the protocol format of NDEF and the information entered in step 1, with some binary editor such as 010editor.

4.5. *Test and Monitor.* The core issues of automatic tests include simulation of NFC Tags and the "touch" operations. We use NFC reader devices (such as ACS ACR122u and PN512) to simulate NFC Tags. The two devices mentioned above are based on the open-source library libnfc. It is independent of the platforms of mobile terminals to simulate a tag by NFC devices. Therefore, our system is universal to test NFC services and applications. The "touch" is the operation that an NFC device touches an NFC tag from far to near. On the Android platform, "touch" is achieved by the command "kill -s SIGSTOP PID" and the command "kill -s SIGCONT PID" to control the process of the NFC service. The NDEF messages' changes in NFC tags can be detected by the Windows Phone platform, so there is no need to simulate "touch."

Target monitoring is also an important part of an automatic test. For Android, we adopt process monitoring and the typical log command "logcat," which has many types of options and operations and can output detailed logs. For Windows phones, we use a real-time monitor called "Xapsy" to detect exceptions. Xapsy is a dynamic monitor, which is realized by the technologies of program instrumentation and API Hook.

5. Experimental Configurations and Environments

5.1. *Experimental Configurations.* The configurations of our experiments can be divided into hardware configuration and software configuration. The former covers hardware devices needed by GNFCVulFinder and their configuration information. The latter covers NFC system services, NFC applications, and their configuration parameters.

5.1.1. *Hardware Configuration.* GNFCVulFinder runs in a virtual machine with the OS version "Linux Ubuntu 2.6.32-21generic"; NFC phones include Samsung GT-I9300, Mi 5S Plus, OnePlus 3T, and Lumia 920. In addition, NFC devices include ACR 122U and Proxmark3. The detailed information is shown in Table 2.

TABLE 1: The database about malformed data.

Name	Malformed data (hex)
Type	0x02, 0x04, 0x06, 0x05, 0x30, 0x40, 0x41, 0x42, 0x43, 0x44, 0xRR (0xRR is random hex data)
Length	0x0, 0xFF, 0xFFFF, 0xRRRR (0xR-RRR is random hex data)

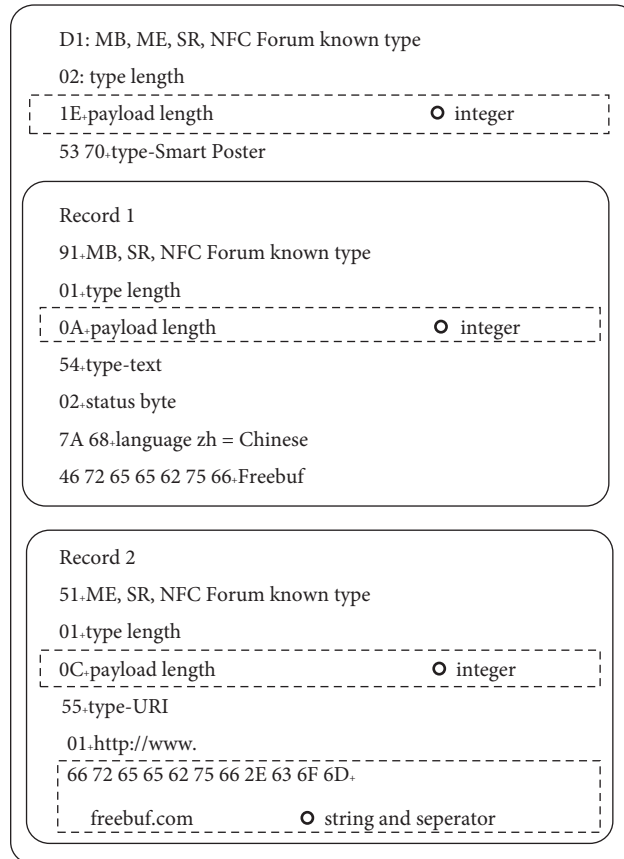


FIGURE 4: A generation example about SmartPoster message.

TABLE 2: Our experiments' hardware configurations.

Hardware name	Configuration
PC	Intel(R) Core(TM) i7, single CPU 2.93 GHz, memory 2 GB, Linux Ubuntu 2.6.32-21generic
Samsung GT-I9300	Samsung Exynos 4412, 4-Core Processor, CPU 1.4 GHz, Mali-400 MP4 GPU, 1 GB RAM, 16 GB ROM
Mi 5S Plus	MIUI 8.2, Android 6.0.1 MXB48T, 4-Core Processor, 6 GB RAM, 128 GB ROM
OnePlus 3T	H2 OS 2.5, Android 6.0, RAM 6 GB, 4-Core Processor, 64 GB ROM
Lumia 920	Windows Phone 8, 2-Core Processor, CPU 1.5 GHz, 1 GB RAM, 32 GB ROM
ACS ACR122U	Frequency 13.56 MHz, compatible with ISO 14443 Types A & B, MIFARE, Felica, and so on
Proxmark3	Frequency 125 KHz, 13.56 MHz, compatible with high-frequency and low-frequency card

5.1.2. *Software Configurations.* The software configurations cover the targets and their information, including NFC system services and third-party applications on Android platform and Windows Phone. The iOS platform does not support reading and writing NDEF messages, so we do not test applications on the iOS platform. By searching the keywords “NFC,” “NDEF,” and “Tag” on Android and Windows Phone’s official and third-party application markets, we currently have collected 42 NFC applications and

NFC services as shown in Table 3. Among these targets, the applications (34–42) are NFC writer applications or need human-computer interaction. These applications have not been tested because they can break the automatic test.

5.2. *Experimental Environments.* The experimental environments are illustrated in Figure 5; (1) in Figure 5 shows the environment of “reverse analysis” and the details are given in

TABLE 3: The targets.

No.	Package name	Name and version	Developer
1	com.android.nfc	Android	Google
2	com.android.nfc	MIUI	MI Corp.
3	com.android.nfc	MK	MoKee
4	com.android.nfc	H2 OS	OnePlus
5	com.android.nfc	Android (Touch Wiz)	Samsung
6	Windows.Networking.Proximity	Windows Phone	Microsoft/Nokia
7	com.nxp.taginfolite	TagInfo 2.00	NXP
8	com.nxp.nfc.tagwriter	TagWriter 2.3	NXP
9	com.fmsh.appsys.nfctag	NFC Tag helper 1.05	Fudan Microelectronics Group
10	hyundai.uni.nfc	NFC Detail 0.5	U&I Research Lab
11	com.sharemore.nfc.transport	ShareMore 3.0	Share More Studio
12	com.antares.nfc	NFC Developer 2.1.1	Thomas Rorvik Skjolberg
13	com.wakdev.wdnfc	NFC Tools 1.7	Wakdev
14	org.ndeftools.boilerplate	NDEF Tool for Android 1.2	Thomas Rorvik Skjolberg
15	com.boco.nfc.activity	elechong 1.0.09	Beijing Yiyang Huizhi
16	com.kddi.nfc.tag_reader	NFC TagReader 2.2.2	KDDI
17	com.microsoft.tag.app.reader	Microsoft tag 5.6.4.90.7866	Microsoft
18	com.mls.handover.wifi	Wifi Handover 1.0.3	STYL Solutions Pte., Ltd
19	at.mroland.android.apps.nfctaginfo	NFC TagInfo 1.11	NFC Research Lab
20	com.touf.mfclassic	MFClassic 2.0	TouF
21	com.widgapp.NFC_ReTAG_FREE	NFC ReTag Free 2.8.2	WidgApp Mobile Solutions
22	com.yunnfc.nfcaction	YunNFC 1.5	Yunfei
23	de.syss.MifareClassicTool	MIFARE Classic Tool 1.3.3	IKARUS Projects
24	se.anyro.nfc_reader	NFC Reader 0.13	Adam Nyback
25	com.nellon.mifareid	MIFARE ID reader 1.1.0	Nellon
26	N/A (Windows Phone)	NFC Commander 2.2.3.5	JasonP
27	N/A (Windows Phone)	NFC Launchit 2.6.0.0	VinApp
28	N/A (Windows Phone)	NFC Writer Reader 1.2.0.0	PCCON
29	N/A (Windows Phone)	NFC Reader 1.2.0.0	DysonChan
30	N/A (Windows Phone)	NFC Tag Writer 2.0.5.9	Mike Francis
31	N/A (Windows Phone)	NFCsms 1.2.0.0	Mopius
32	N/A (Windows Phone)	Nokia NFC Writer 1.0.0.59	Microsoft Mobile
33	jp.co.menox.caffeetime.toiki.nfc.mifare2ndef	MIFARE NDEF TOOL 1.1	N/A
34	com.sony.easyconnect	NFC Easy Connect1.0.02	Sony
35	com.hkphka.mifaredoctor	MIFARE Doctor 2.7	Kings Studio
36	com.tagstand.writer	Tagstand Writer 2.0.14	Egomotion
37	tw.com.method.rfidtool	RFID Tool 0.3	Method
38	com.nfcquickactions.appfree	NFC Actions 2.0.3	Flomio
39	com.anytag.android	AnyTag NFC Launcher 1.3.0	XtraSEC
40	com.skjolberg.nfc.clone	NFC Tag Cloner 1.2.4	Thomas Rorvik Skjolberg
41	nz.intelx.send.lite	Send! 2.0.2	Billy Lam
42	com.Samsung.android.app.gearnfcwriter	NFC Tagwriter 1.4.131022	Samsung

Section 4.4. By the way, Proxmark3 has been used to sniff the messages between NFC phones and NFC tags. (2) in Figure 5 illustrates our test process. Firstly, simulate NFC tags with ACS ACR122u based on libnfc. Secondly, “touch” tags with NFC phone. Finally, monitor the exceptions and output the results.

6. Results

We test the targets in Figure 3 with GNFCVulFinder and find eight new vulnerabilities of mobile OS and applications, as shown in Table 4. We divide the vulnerabilities into five types according to the ways they can be exploited to describe how to use these vulnerabilities to attack. The specific descriptions are as follows:

- (1) NFC service DoS: this vulnerability is triggered by a Bluetooth pairing message with its field “local name length” filled by 0b0000 0000 or 0b1xxx xxxx (x denotes 0 or 1). NFC service will crash, report the exception “java.lang.NegativeArraySizeException,” and stop the service. The results are shown in Figure 6.
- (2) Opening torch: this vulnerability is caused by starting a package “com.android.systemui,” which leads the torch in an NFC phone to be opened automatically. “com.android.systemui” is a core system application in MIUI, which is used to manage battery, configure MIUI OS, and so on. After decompiling the torch apk, we found that the activity of the torch is the only activity kept in the root directory, which will cause the torch to be activated.

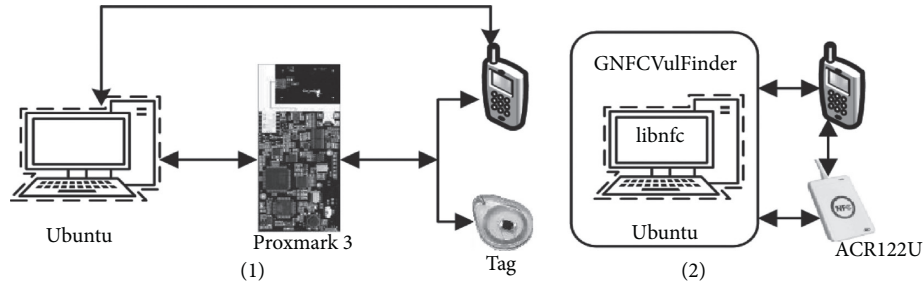


FIGURE 5: Experimental environments.

- (3) Opening Bluetooth and opening Wi-Fi (logic vulnerability): the two vulnerabilities are both logic vulnerabilities. Suppose that Bluetooth and Wi-Fi are closed by default. Bluetooth or Wi-Fi will be closed normally if the pairing or connection fails. However, they are opened, respectively, after pairing or connection failure. The result of Bluetooth opening is shown in Figure 7. No matter which choice is to be selected, Bluetooth will be opened. The result of Wi-Fi opening is shown in Figure 8. When a phone connects to a Wi-Fi SSID “UCAS” successfully or not, the Wi-Fi of the phone will be opened. For the type of logic vulnerability, we give an attack scene as an example; attackers firstly open the Wi-Fi of an NFC phone and then connect NFC phones by NFC or Wi-Fi to get sensitive information combined with other vulnerabilities. Since the messages are normal NDEF packets, we do not need to describe them specially.
- (4) Denial of Service (TagInfo DoS, NFC Tag helper DoS, and NFC Detail DoS): we find lots of DoS vulnerabilities when testing TagInfo, Detail!, and NFC Tag helper. The messages that can cause a Denial of Service of these applications, including smart posters, URLs, and Wi-Fi connections. For the reason that these applications do not handle NDEF messages properly, they will crash when reading these types of messages, as shown in Figure 9. Two types of exceptions are captured including “java.lang.IllegalStateException,” “java.lang.IndexOutOfBoundsException,” “java.lang.ArrayIndexOutOfBoundsException,” and “java.lang.NullPointerException.” These exceptions cannot be exploited by analyzing the logs, and they just lead to a Denial of Service attack.
- (5) Black screen: this vulnerability existing in the NFC Tag helper is a type of design defect. One of the functions of this application is to set the brightness of the Android phone screen. The value scale is between 5 and 100. However, we can break the limit and set the brightness to 0 by writing 0x00 to the NFC tag. The screen of the Android NFC phone will turn black until rebooting the phone or rewriting a new brightness value.

The message that triggers the black screen vulnerability is defined as follows: D1 01 12 54 02 65 6E 10 02 0C 01 04 6E 69 70 63 02 01 1E 03 01 00.

The above test results belong to the Android platform, and no exceptions were found on Windows Phone. The reasons for the results are summarized as follows:

- (i) Many applications need user interaction; namely, when an NFC phone “touches” an NFC tag, the operating system Windows Phone leaves the decision to the users, which interrupts the user interaction testing process. If the user does not make any decision, the messages will be denied by Windows Phone, which results in the low efficiency of the test.
- (ii) The number of NFC applications is small, and most applications just print the binary data of NDEF messages. The operations embedded in the messages have no chance to be executed.

7. Evaluations and Measures

7.1. Evaluations. To illustrate our work’s value and efficiency, we compared it with some related research results in five parts: test case construction, monitoring, portability, automation, and vulnerabilities. The results shown in Table 5 are analyzed as follows:

- (i) Test case construction: Mulliner [17] constructed test cases manually, the efficiency of which is low and has a lot to do with the testers’ experience. Reference [8] and Wiedermann [22] adopt a single strategy: generation strategy (G, for short) (Sulley) or the mutation strategy (M, for short). The shortcomings of single generation are randomness and blindness, resulting in low detection rate. GNFCVulFinder adopts manual generation and mutation to construct malformed data and proposes a method called “reverse analysis” as an auxiliary method for analyzing messages. Our strategy is based on the known vulnerability knowledge of NDEF, which can help find bugs quickly; a manual strategy can detect logical vulnerabilities that fuzzing cannot find. In addition, a multidimensional strategy can detect multidimensional vulnerabilities, which can increase the detection rate.
- (ii) Monitoring: Mulliner [17] monitors exceptions manually; the efficiency is low and the false-negative rate is high. Miller [8] uses logcat to monitor exceptions, and Wiedermann uses Sulley; a single monitor may omit some exceptions.

TABLE 4: Experimental results.

No.	Name	Type	Impact
1	NFC service DoS	Denial of Service	Android 4.1.1, 4.1.2, 4.4
2	Opening torch	Design defect	MIUI-3.6.21, 4.1.17, 4.5.30
3	Opening Bluetooth	Logic vulnerability	MIUI-3.6.21, 4.1.17, 4.5.30
4	Opening Wi-Fi	Logic vulnerability	MIUI-3.6.21, 4.1.17, 4.5.30
5	TagInfo DoS	Denial of Service	TagInfo 2.00
6	NFC Detail DoS	Denial of Service	NFC Detail 0.5
7	Black screen	Design defect	NFC Tag helper1.05
8	NFC Tag helper DoS	Denial of Service	NFC Tag helper1.05

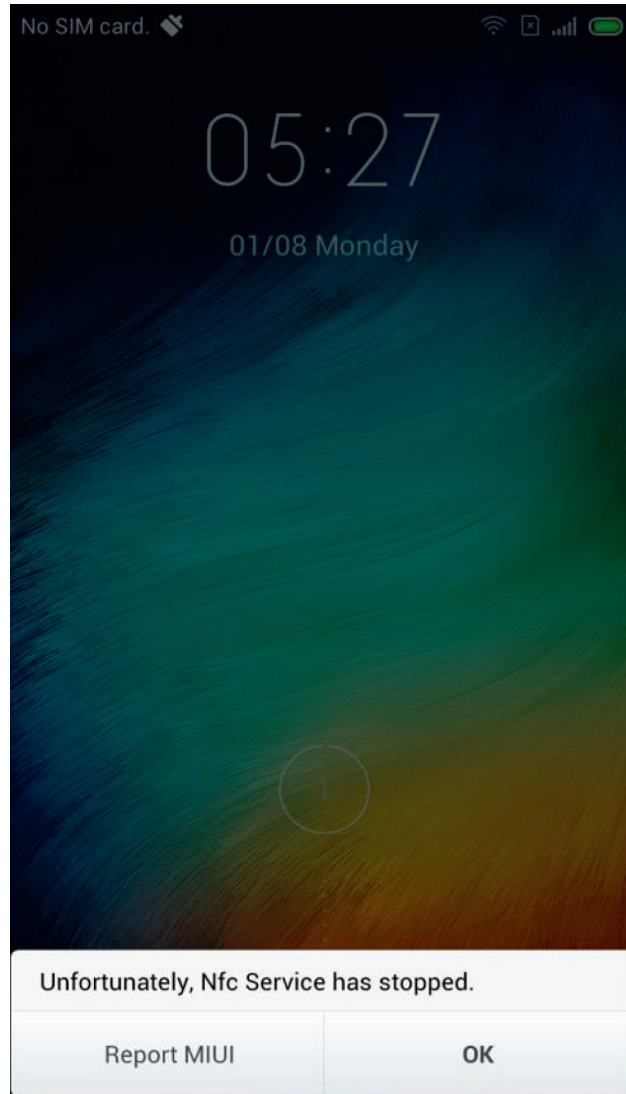


FIGURE 6: NFC service crash.

GNFCVulFinder monitors the targets by logcat, monitoring the NFC process, which increases the detection rate.

(iii) Portability: the portability of GNFCVulFinder and Mulliner’s work is good; others are dependent on Android.

(iv) Automation (auto): except for Mulliner’s work, the automation level of others is high.

(v) Vulnerability: GNFCVulFinder finds eight vulnerabilities and covers all others’ vulnerabilities, including some new vulnerabilities such as opening Wi-Fi/Bluetooth and black screen.

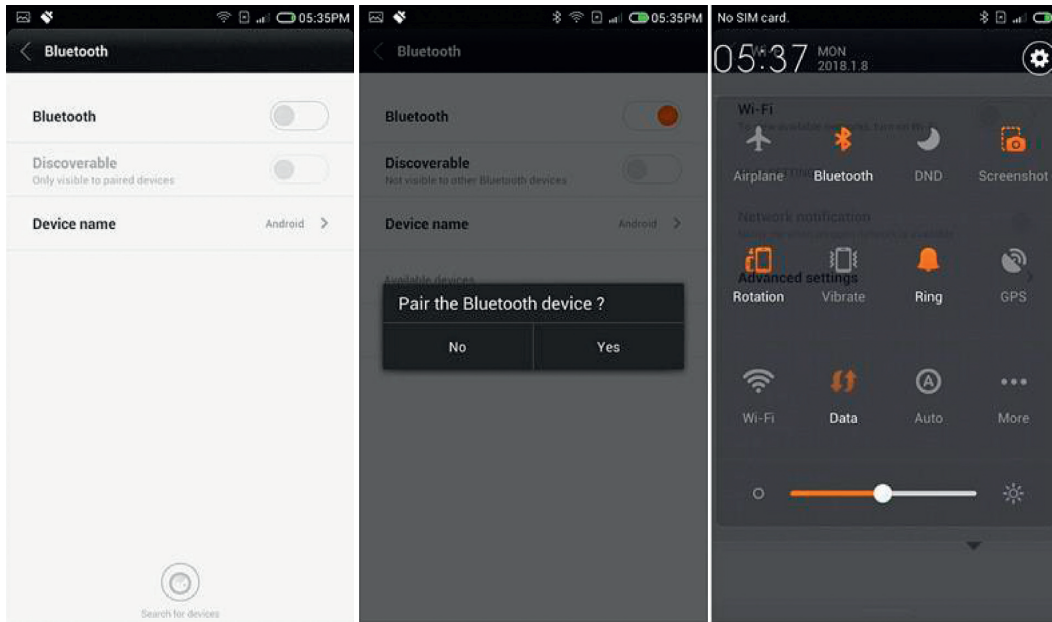


FIGURE 7: The Bluetooth is opened.

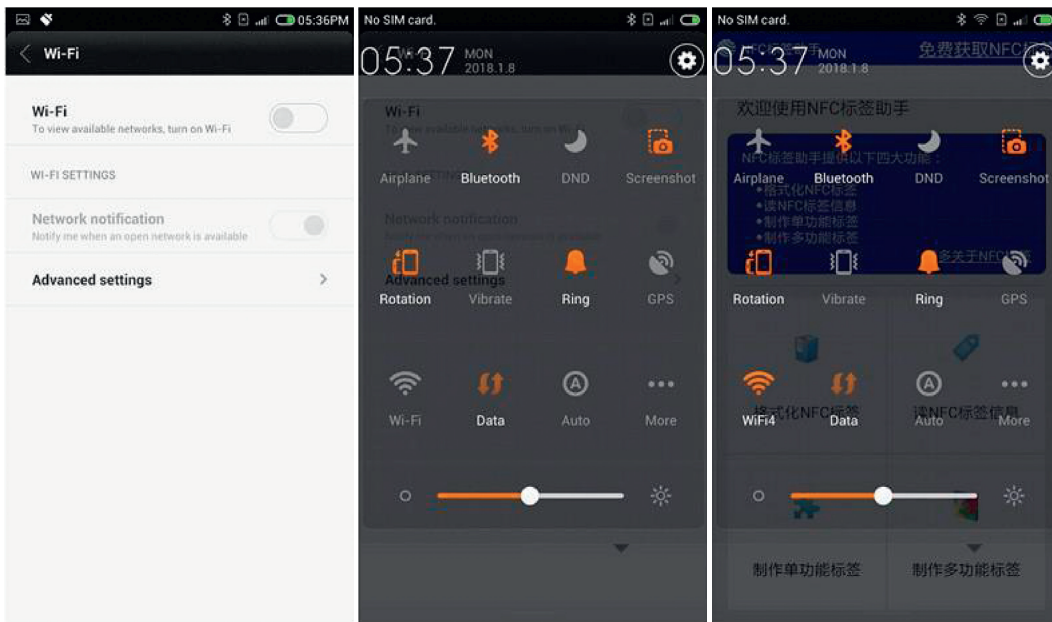


FIGURE 8: The Wi-Fi is opened.

7.2. Measures. To solve the above-mentioned problems and vulnerabilities, we propose the following specific measures and suggestions:

- (1) Detect fields about length in all kinds of messages, and ensure that the value of 0 and the negative value can be resolved correctly.

E/AndroidRuntime (20745): FATAL EXCEPTION: main
E/AndroidRuntime (20745): java.lang.NullPointerException
E/AndroidRuntime (20745): at com.fmsh.appsys.nfctag.nfc.base.1.a (Unknown Source)
E/AndroidRuntime (20745): at com.fmsh.appsys.nfctag.nfc.base.e.a (Unknown Source)
E/AndroidRuntime (20745): at com.fmsh.appsys.nfctag.nfc.EntranceActivity.b (Unknown Source)
E/AndroidRuntime (20745): at com.fmsh.appsys.nfctag.nfc.EntranceActivity.onNewIntent (Unknown Source)
E/AndroidRuntime (32443): FATAL EXCEPTION: main
E/AndroidRuntime (32443): java.lang.IllegalStateException: Couldn't read row 2, col 0 from CursorWindow. Make sure the Cursor is initialized correctly before accessing data from it.
E/AndroidRuntime (32443): at android.database.CursorWindow.nativeGetLong (Native Method)
E/AndroidRuntime (32443): at android.database.CursorWindow.getLong (CursorWindow.java: 511)
E/AndroidRuntime (32443): at android.database.AbstractWindowedCursor.getLong (AbstractWindowedCursor.java: 75)
E/AndroidRuntime (32443): at android.database.AbstractCursor.moveToPosition (AbstractCursor.java: 219)
E/AndroidRuntime (32443): at android.database.CursorWrapper.moveToPosition (CursorWrapper.java: 162)
E/AndroidRuntime (21432): FATAL EXCEPTION: main
E/AndroidRuntime (21432): java.lang.RuntimeException: Unable to start activity ComponentInfo (hyundai.uni.nfc.hyundai.uni.nfc.ReadTagMain_Activity): java.lang.IndexOutOfBoundsException: Invalid index 0, size is 0
E/AndroidRuntime (21432): at android.app.ActivityThread.performLaunchActivity (ActivityThread.java: 2100)
E/AndroidRuntime (21432): at android.app.ActivityThread.handleLaunchActivity (ActivityThread.java: 2125)
E/AndroidRuntime (21432): at android.app.ActivityThread.access\$600 (ActivityThread.java: 140)
E/AndroidRuntime (21432): at android.app.ActivityThread\$H.handleMessage (ActivityThread.java: 1227)
E/AndroidRuntime (21432): at android.os.Handler.dispatchMessage (Handler.java: 99)
E/AndroidRuntime (29559): FATAL EXCEPTION: main
E/AndroidRuntime (29559): java.lang.ArrayIndexOutOfBoundsException: length = 69; index = 69
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.base.a.a (Unknown Source)
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.wifi.e.b (Unknown Source)
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.wifi.f.a (Unknown Source)
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.base.1.a (Unknown Source)
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.base.e.a (Unknown Source)
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.EntranceActivity.b (Unknown Source)
E/AndroidRuntime (29559): at com.fmsh.appsys.nfctag.nfc.EntranceActivity.onNewIntent (Unknown Source)

FIGURE 9: The DoS of NFC Detail.

TABLE 5: Comparison results.

Research work	Construction	Monitoring	Portability	Automation	Vulnerability
Mulliner [17]	Manual	Manual	n/a	n/a	2
Miller [8]	G or M	Logcat	Good	High	2
Wiedermann [22]	G	Sulley	Bad	High	2
GNFCVulFinder	Manual, G, M, and R	Logcat, process monitor	Good	High	8

- (2) Modify the logic of Bluetooth pairing and Wi-Fi connection, and restore their original condition when operations fail.
- (3) When developing NFC applications, follow NDEF specifications strictly.
- (4) Add some means to filter special characters, such as directory traversal characters, to prevent these characters from triggering security vulnerabilities.

8. Conclusions

This paper proposed an algorithm named GTCT to construct test cases and test the protocol NDEF based on fuzzing. The GTCT adopts four strategies to generate test cases: manual, generation, mutation, and “reverse analysis.” The manual strategy is helpful to find logic vulnerabilities that fuzzing cannot detect. In addition, the known vulnerability data is also used to construct test cases, which can improve test efficiency. Based on the algorithm, we designed an NFC vulnerability discovery framework and developed a tool called GNFCVulFinder. By testing lots of NFC system services and applications, we find 8 NDEF vulnerabilities, which can cover all previous vulnerabilities and prove the effectiveness of GNFCVulFinder.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Zhiqiang Wang and Yuheng Lin designed the system architecture and conceived the experiments; Zihan Zhuo and Jieming Gu conducted the experiments; and Tao Yang analyzed the results. All of the authors reviewed the manuscript.

Acknowledgments

This research was financially supported by China Post-doctoral Science Foundation funded project (2019M650606), National Key R&D Program of China (2017YFC1201204), First-Class Discipline Construction Project of Beijing Electronic Science and Technology Institute (3201012), and the Fundamental Research Funds for the Central Universities (328201909).

References

- [1] V. Coskun, B. Ozdenizci, and K. Ok, "A survey on near field communication (NFC) technology," *Wireless Personal Communications*, vol. 71, no. 3, pp. 2259–2294, 2013.
- [2] K. Markantonakis and K. Mayes, *Secure Smart Embedded Devices, Platforms and Applications*, Springer, New York, NY, USA, 2013.
- [3] R. Want, "Near field communication," *IEEE Pervasive Computing*, vol. 10, no. 3, pp. 4–7, 2011.
- [4] K. Curran, A. Millar, and C. Mc Garvey, "Near field communication," *International Journal of Electrical and Computer Engineering*, vol. 2, p. 371, 2012.
- [5] IHS Technology, "NFC-enabled cellphone shipments to soar fourfold in next five years," April 2017, <https://technology.ihs.com/490062/nfc-enabled-cellphone-shipments-to-soar-fourfold-in-next-five-years>.
- [6] J. Rubin, *Google Wallet Security: Pin Exposure Vulnerability*, ZveloBLOG, Greenwood Village, CO, USA, 2012.
- [7] T. Huynh, "Second major security flaw found in Google Wallet," December 2017, <http://thesmartphonechamp.com/second-major-securityflaw-found-in-google-wallet-rooted-or-not-no-one-is-safe-video/>.
- [8] C. Miller, "Exploring the NFC attack surface," in *Proceedings of the Blackhat*, Las Vegas, NV, USA, August 2012.
- [9] C. Benninger, M. Sobell, "NFC for free rides and rooms (on your phone)," in *Proceedings of EUsecWest Conference*, Amsterdam, Netherland, September 2012.
- [10] WallofSheep, "NFC security awareness project," December 2017, <http://www.wallofsheep.com/pages/nfc-security-awareness-project>.
- [11] Z. D. Initiative, "Samsung SBeam image remote information disclosure vulnerability," December 2017, <http://www.zerodayinitiative.com/advisories/ZDI-15-257>.
- [12] GadgetHacks, "Apple watch exploit," December 2017, <https://www.youtube.com/watch?v=2blTo-Ej6mo&feature=youtu.be>.
- [13] S. Mendoza, "Samsung Pay: tokenized numbers, flaws and issues," in *Proceedings of the Blackhat*, Las Vegas, NV, USA, July-August 2016.
- [14] C. Xinyi, C. Kyung and K. Choi, "A secure and efficient key authentication using bilinear pairing for NFC mobile payment service," *Wireless Personal Communications*, vol. 97, 2017.
- [15] S. Akter and S. Chellappan, "Man-in-the-middle attack on contactless payment over NFC communications: design, implementation, experiments and detection," *IEEE Transactions on Dependable and Secure Computing*, p. 1, 2020.
- [16] E. Haselsteiner and K. Breituß, "Security in near field communication (NFC)," in *Proceedings of the Workshop on RFID Security*, Gratkorn, Austria, January 2006.
- [17] C. Mulliner, "Vulnerability analysis and attacks on NFC-enabled mobile phones," in *Proceeding of the 2009 International Conference on Availability, Reliability and Security*, Fukuoka, Japan, March 2009.
- [18] V. Gauthier, K. Damme and Wouters, "Practical experiences with NFC security on mobile phones," in *Proceedings of 2009 International Workshop on RFID Security and Privacy Issue*, Leuven, Belgium, January 2009.
- [19] J. Antonio, Z. Migual, F. A. Alberto, and F. G. A. Skarmeta, "Evaluation of the security capabilities on NFC-powered devices," in *Proceedings of European Workshop on Smart Objects: Systems, Technologies and Applications*, Ciudad, Spain, June 2010.
- [20] F. Jia and X. Tong, "Threat modeling for mobile payments using NFC phones," *Journal of Tsinghua University & Technology*, vol. 52, pp. 1460–1464, 2012.
- [21] O. Ghag and S. Hegde, "A comprehensive study of google wallet as an NFC application," *International Journal of Computing and Applications*, vol. 58, 2012.
- [22] W. Norbert, *Fuzzing-to-go: A Test Framework for Android Devices*, Technische University München, Munich, Germany, 2012.
- [23] J. Gummeson, D. Ganesan, B. Priyantha, D. Thrasher, and P. Zhang, "Engarde: protecting the mobile phone from malicious NFC interactions," in *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 445–458, ACM, Taipei, Taiwan, June-2013.
- [24] M. Roland, "Debugging and rapid prototyping of NFC secure element applications," in *Proceeding of the International Conference on Mobile Computing, Applications, and Services*, pp. 298–313, Springer, Paris, France, November 2013.
- [25] N. Forum, "NFC forum technical specifications," August 2017, Available at: <http://nfc-forum.org/our-work/specifications-and-applicationdocuments/specifications/nfc-forum-technical-specifications/>.
- [26] Z. Wang, Q. Liu and Y. Zhang, "A research on vulnerability discovering for router protocols based on fuzzing," in *Proceedings of 2012 7th International ICST Conference on Communications and Networking in China (CHINACOM)*, IEEE, Kunming, China, August 2012.
- [27] M. Sutton, A. Greene, and P. Amini, *Fuzzing: Brute Force Vulnerability Discovery*, Pearson Education, London, UK, 2007.
- [28] A. Takanen, *Fuzzing: The Past, the Present and the Future*, SSTIC, Rennes, France, 2009.
- [29] M. B. S. Maryam and M. Zolfaghari, "A smart fuzzing method for detecting heap-based vulnerabilities in executable codes," *Security and Communication Networks*, vol. 9, 2016.
- [30] J. Yan, Y. Zhang and Y. Dingning, "Structured grammar-based fuzz testing for programs with highly structured inputs," *Security and Communication Networks*, vol. 6, pp. 1319–1330, 2013.
- [31] Z. Wang, Q. Liu and Y. Zhang, "A framework for discovering router protocols vulnerabilities based on fuzzing," *KSII Transactions on Internet and Information System*, vol. 7, 2013.
- [32] NVD, "National vulnerability database," December 2017, <https://nvd.nist.gov/>.

Research Article

Modeling and Control of Malware Propagation in Wireless IoT Networks

Qing Yan,¹ Lipeng Song ,² Chenlu Zhang,¹ Jing Li,¹ and Shanshan Feng¹

¹School of Data Science and Technology, North University of China, Taiyuan 030051, Shanxi, China

²School of Mechanical, Electrical & Information Engineering, Shan Dong University, Weihai 264209, Shandong, China

Correspondence should be addressed to Lipeng Song; slp880@gmail.com

Received 8 April 2021; Accepted 18 May 2021; Published 15 June 2021

Academic Editor: Jingyu Feng

Copyright © 2021 Qing Yan 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 Internet of Things (IoT) devices densely populate our daily life, but also attract many attackers to attack them. In this paper, we propose a new Heterogeneous Susceptible-Exposed-Infected-Recovered (HSEIR) epidemic model to characterize the effect of heterogeneity of infected wireless IoT devices on malware spreading. Based on the proposed model, we obtain the basic reproduction number, which represents the threshold value of diffusion and governs that the malware is diffusion or not. Also, we derive the malware propagation scale under different cases. These analyses provide theoretical guidance for the application of defense techniques. Numerical simulations validated the correctness and effectiveness of theoretical results. Then, by using Pontryagin's Minimum Principle, optimal control strategy is proposed to seek time-varying cost-effective solutions against malware outbreaks. More numerical results also showed that some control strategies, such as quarantine and vaccination, should be taken at the beginning of the malware outbreak immediately and become less necessary after a certain period. However, the repairing and fixing strategy, for example applying antivirus patches, would be keep on going constantly.

1. Introduction

The smart Internet of Things (IoT), such as intelligent transportation, smart homes, smart grid, and the next industrial revolution (i.e., Industry 4.0), are embedded in billions of wireless devices. Often, wireless devices rely heavily on the wireless connectivity such as WiFi, Bluetooth, and Zigbee [1, 2]. Wireless networks (WNs), as a kind of new information and communication network, connect the physical systems to cyber worlds. In addition, WNs have gained widespread application in IoT including healthcare, public safety, agriculture, and retail due to their low power consumption, flexible networking, unlimited potential, and relevant other features [3–5].

However, the development of IoT brings a series of new challenges to cyber security. On the one hand, the low cost and short time-to-market nature of wireless IoT devices make them, such as sensors, actuators and smart appliances, expose to a high risk of malware infiltration. In addition, most IoT devices are left to operate on consumer

premises without regular maintenance. On the other hand, the majority of wireless IoT devices are installed and controlled by consumers with limited security background. Even consumers may be willing to accept to install certain processes or applications on their devices in exchange for incentives without realizing the fact that this may cause attack [6]. In reality, there are hundreds of malware attacks against IoT that occur every year around the world, and each incident affects production and economy greatly.

Malware (Trojan, virus, and worms) is an intrusive software. It is designed for a variety of criminal and hostile activities such as spying, threatening for monetary benefit, or controlling a large population of devices [7]. The academic research focuses on the spread of malware from the following perspectives: detection technology and mathematical modeling [8–12]. Based on multidimensional hybrid features extraction and analysis, Li and Xu [8] proposed a novel method to detect Android malware. In [9], Du and Liu proposed a packet-based malicious

payload detection and identification algorithm using the deep learning method. However, the above method to detect the malware cannot predict malware spread scale and explore the key factors affecting the spread of malware. In the wireless IoT network, malware hostile activities naturally extend to physical threats and can be launched by one wireless device and spread to another device [13]. Propagation vectors such as services or functions may be used to propagate malware due to different vulnerabilities that emerge across different technologies such as IoT or IPv6 and the lack of experience in technical implementation. As a result, the propagation process is difficult to detect and observe. Based on our understanding of the technology and experience with historic attacks, modeling approaches are used to predict propagation dynamics and to explore influential factors [14–19]. Recently, Chen and Cheng [20] proposed a novel traffic-aware patching scheme to select important intermediate nodes to patch and apply this to the IoT system with limited patching resources and response time constraint. On the basis of the difference of intelligent device's dissemination capacity and discriminant ability, Li and Cui [21] discussed a dynamic malware propagation model to study the malware propagation in industrial Internet of Things. Considering both the heterogeneity and mobility of sensor nodes, Shen and Zhou [22] proposed a heterogeneous and mobile vulnerable-compromised-quarantined-patched-scrapped (VCQPS) model. In [7], Elsayy et al. defined spatial firewalls to control malware spreading in Wireless Network. Gao and Zhuang [23] studied worm propagation with saturated incidence and strategies of both vaccination and quarantine. In [6], Farooq and Zhu proposed an analytical model to study the D2D propagation of malware in wireless IoT networks.

Motivated by the abovementioned research, we propose a Heterogeneous Susceptible-Exposed-Infected-Recovered (HSEIR) model, which characterizes the influence of infected wireless devices' heterogeneity in malware spreading. In our model, a node is a wireless IoT device. And the ability of wireless devices to spread the malware is different due to the factors of its topology links and processing capacity. As far as we know, little work has been done on the heterogeneity of wireless devices. In this work, we make contributions as follows:

- (1) We propose a HSEIR model. To the best of our knowledge, this is the first work for disclosing dynamics of malware propagation in the wireless IoT network.
- (2) We derive differential equations of the HSEIR model, which can reflect the number of wireless devices belonging to different states varying with time.
- (3) We discuss the malware propagation threshold by the proposed model. We also proof the stability of equilibrium, upon which we can judge the malware spread scale.
- (4) We study control strategy in two aspects. One is optimal control to minimize the number of infected

devices (including exposed and infected nodes) and the corresponding cost of the strategies during the process of spreading. And, the other one is to control the malware illness upon the malware propagation threshold.

The remainder of this paper is organized as follows. In Section 2, we establish the HSEIR model in IoT wireless networks. In Section 3, we analyze the dynamics of the proposed model. In Section 4, control strategies by Pontryagin's Minimum Principle and malware propagation threshold are discussed. Section 5 gives the details of numerical simulation of the model. Conclusions and discussion are given in Section 6.

2. The HSEIR Model

In this section, we describe the novel HSEIR model, which considers the heterogeneity of infected wireless devices in malware spreading.

2.1. The States and State Transitions in HSEIR Model. In our model, we assume that the proportion of wireless devices infected by the nodes with weak spreading capabilities is a and the proportion of wireless devices infected by the nodes with strong spreading capabilities is $1 - a$. The total device population N is divided into four different compartments (susceptible (S), exposed (E), infected (I), and recovered (R)). Every device may be in one of such compartments at time tick t .

- (i) S : the wireless devices in this compartment have not been infected by malware, but they are vulnerable to malware M
- (ii) E : the wireless devices in this compartment are exposed to the attacks but do not exhibit due to the latent time requirement
- (iii) I : the wireless devices in this compartment are infected and may infect other devices
- (iv) R : the wireless devices in this compartment are vaccinated and are immune to M

We make the following assumptions:

- (1) Once the wireless devices are vaccinated, they will be permanently immune and cannot be infected by M any more
- (2) The nature death rate is extremely small. In reality, the service life of the wireless device is far more than the time from malware appearance to the end of the attack

At time $t = 0$, all nodes are in the susceptible compartment. Once the malware M intrudes into the system, a node may move from one state to another, as shown in Figure 1, which illustrates the state transition diagram of a node. Table 1 shows the parameters involved in this paper. The nodes changing their states upon the following rules:

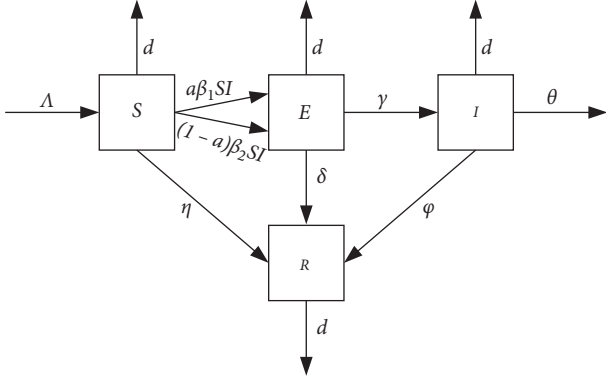


FIGURE 1: The state transition in the HSEIR model.

- (i) Due to the partial efficiency of the vaccine, there is only η fraction of the vaccinated susceptible nodes that move to R state per unit time
- (ii) The remaining susceptible devices move to the exposed state, and the new exposed devices at time t are given by the expression $a\beta_1 I(t)S(t) + (1-a)\beta_2 I(t)S(t)$, where β_1 and β_2 stand for the transmission coefficient of devices with weak spreading capability and strong capability and a and $1-a$ represent the fraction of susceptible devices targeted by devices with weak spreading capability and strong capability, respectively
- (iii) The exposed devices transit into I with γ when the malware begins actively, where γ is the mean latent period
- (iv) Using some sufficient defense mechanisms, a portion of the exposed, infectious devices can recover at rates δ and φ , respectively:

$$\begin{cases} \frac{dS}{dt} = \Lambda - [a\beta_1 I + (1-a)\beta_2 I]S - \eta S - dS, \\ \frac{dE}{dt} = [a\beta_1 I + (1-a)\beta_2 I]S - \gamma E - \delta E - dE, \\ \frac{dI}{dt} = \gamma E - \varphi I - dI - \theta I, \\ \frac{dR}{dt} = \eta S + \delta E + \varphi I - dR. \end{cases} \quad (1)$$

3. Model Analysis

In this section, we firstly calculate the malware-free, malware-existence equilibria and the basic reproduction number of the HSEIR model. Then, we proof the local and global stability of each equilibria.

3.1. Equilibria and Basic Reproduction Number. Summing the right hand of (1), we have

$$\frac{d(S + E + I + R)}{dt} = \frac{dN}{dt} = \Lambda - dN - \theta I \leq \Lambda - dN, \quad (2)$$

and after a simple computation, we have

$$N(t) \leq \frac{\Lambda}{d} + \left(N(0) - \frac{\Lambda}{d}\right)e^{-dt}. \quad (3)$$

Then, we have

$$\limsup_{t \rightarrow \infty} N(t) \rightarrow \frac{\Lambda}{d}. \quad (4)$$

One can verify that the positive cone \mathbb{R}_+^4

$$\Omega = \left\{ (S, E, I, R) \in \mathbb{R}_+^4 \mid 0 \leq S + E + I + R \leq \frac{\Lambda}{d} \right\}. \quad (5)$$

Denote that the region Ω is the positively invariant of system (1).

Next, we calculate the basic reproduction number R_0 by the method of van den Driessche and Watmough [24]. R_0 is a threshold value of the epidemiological model, which indicates the number of wireless IoT devices infected by an infectious device during its average period of illness at the beginning of the disease, when all are susceptible. It is easy to obtain that system (1) always has a malware-free equilibrium $P_0 = (S_0, E_0, I_0, R_0) = ((\Lambda/d + \eta), 0, 0, (\Lambda\eta/d + \eta))$, and the associated next generation matrices are given by R_0 :

$$F = \begin{pmatrix} 0 & \frac{\Lambda}{d + \eta} [a\beta_1 + (1-a)\beta_2] \\ 0 & 0 \end{pmatrix}, \quad (6)$$

$$V = \begin{pmatrix} \gamma + \delta + d & 0 \\ -\gamma & \varphi + d \end{pmatrix}.$$

Then, the basic reproduction number R_0 of the system is as:

$$R_0 = \rho(FV^{-1}) = \frac{\Lambda\gamma(a\beta_1 + (1-a)\beta_2)}{(d + \eta)(\gamma + \delta + d)(d + \varphi + \theta)}. \quad (7)$$

It can be seen that system (1) has a malware-existence equilibrium $P^* = (S^*, E^*, I^*, R^*)$ in Ω if $R_0 > 1$, which satisfies that

$$\begin{cases} S^* = \frac{\Lambda}{(a\beta_1 + \beta_2 - a\beta_2)I^* + (d + \eta)}, \\ E^* = \frac{d + \varphi + \theta}{\gamma} I^*, \\ I^* = \frac{(R_0 - 1)(d + \eta)}{(a\beta_1 + \beta_2 - a\beta_2)}, \\ R^* = \frac{\delta E^* + \varphi I^* + \eta S^*}{d}. \end{cases} \quad (8)$$

TABLE 1: Nations and parameter values in the model.

Nation	Explanation
$N(t)$	Total number of wireless device at time t
$S(t)$	Number of susceptible devices at time t
$E(t)$	Number of exposed devices at time t
$I(t)$	Number of infected devices at time t
$R(t)$	Number of secured devices at time t
a	The proportion of wireless devices infected by weak spreading capabilities nodes
$1 - a$	The proportion of wireless devices infected by strong spreading capabilities nodes
β_1	Infection rate of devices with weak spreading capabilities
β_2	Infection rate of devices with strong spreading capabilities
Λ	The recruitment rate
d	The natural death rate
γ	State transition rate from E to I
δ	State transition rate from E to R
φ	State transition rate from I to R
η	State transition rate from S to R
θ	Malware-related death rate

3.2. Stability of Equilibrium

$$\begin{aligned}
|\lambda \mathbf{E} - J(P_0)| &= \begin{vmatrix} \lambda + d + \eta & 0 & \frac{\Lambda(a\beta_2 - \beta_2 - a\beta_1)}{\eta + d} & 0 \\ 0 & \lambda + \gamma + \delta + d & -\frac{\Lambda(a\beta_1 + \beta_2 - a\beta_2)}{\eta + d} & 0 \\ 0 & -\gamma & \lambda + d + \varphi + \theta & 0 \\ -\eta & -\delta & -\varphi & \lambda - d \end{vmatrix} \\
&= (\lambda + d)(\lambda + \eta + d) \left[(\lambda + d + \varphi + \theta)(\lambda + \gamma + \delta + d) - \frac{\Lambda\gamma}{\eta + d} (a\beta_1 + \beta_2 - a\beta_2) \right] \\
&= 0.
\end{aligned} \tag{10}$$

It is clear that (10) has two negative roots $\lambda_1 = -d$ and $\lambda_2 = -(d + \eta)$, and other roots of (10) are determined by the following equation:

$$(\lambda + d + \varphi + \theta)(\lambda + \gamma + \delta + d) - \frac{\Lambda\gamma(a\beta_1 + \beta_2 - a\beta_2)}{\eta + d} = 0. \tag{11}$$

If $R_0 < 1$, all roots of (11) have negative real parts, so all roots of (10) have negative real parts. Therefore, the malware-free equilibrium P_0 is locally asymptotically stable by the Hurwitz criterion. If $R_0 > 1$, the root of (11) has both positive and negative real parts, so the malware-free equilibrium P_0 is unstable. \square

Lemma 1. *The malware-free equilibrium P_0 of system (1) is locally asymptotically stable if $R_0 < 1$, and unstable if $R_0 > 1$.*

Proof. The Jacobian matrix of system (1) at P_0 is

$$J(P_0) = \begin{pmatrix} -(\eta + d) & 0 & \frac{\Lambda(a\beta_2 - \beta_2 - a\beta_1)}{\eta + d} & 0 \\ 0 & -(\gamma + \delta + d) & \frac{\Lambda(a\beta_1 + \beta_2 - a\beta_2)}{\eta + d} & 0 \\ 0 & \gamma & -(d + \varphi + \theta) & 0 \\ \eta & \delta & \varphi & -d \end{pmatrix}, \tag{9}$$

and its characteristic equation is

Theorem 1. *The malware-free equilibrium P_0 of (1) is globally asymptotically stable if $R_0 \leq 1$, and unstable if $R_0 > 1$.*

Proof. Consider the Lyapunov function as

$$L = \gamma E + (\gamma + \delta + d)I. \tag{12}$$

Then,

$$\begin{aligned}
L' &= \gamma' + (\gamma + \delta + d)I', \\
&= [\gamma S(a\beta_1 + \beta_2 - a\beta_2) - (\gamma + \delta + d)(\varphi + d + \theta)]I, \\
&\leq [\gamma S_0(a\beta_1 + \beta_2 - a\beta_2) - (\gamma + \delta + d)(\varphi + d + \theta)]I, \\
&= (\gamma + \delta + d)(\varphi + d + \theta)(R_0 - 1)I.
\end{aligned} \tag{13}$$

Thus, $L' \leq 0$ when $R_0 \leq 1$. The equality is holding if and only if $R_0 = 1$, $E = 0$, and $S = S_0$ or $E = 0$ and $S = S_0$, or $R_0 = 1$ and $S = S_0$. If $R_0 = 1$, $E = 0$, and $S = S_0$, then the only compact invariant subset in the set $L' \leq 0$ is the singleton P_0 ; if $E = 0$ and $S = S_0$, the only compact invariant subset in the set $L' \leq 0$ is also the singleton P_0 ; if $R_0 = 1$ and $S = S_0$, the only compact invariant subset in the set $L' \leq 0$ is also the singleton P_0 ; therefore, the largest invariant subset in the set $\{(dL/dt) = 0\}$ also is the singleton P_0 . If $R_0 > 1$, we have $(dL/dt) > 0$. Finally, taking into account $L(E, I) = \gamma E + (\gamma + \delta + d)I$ and LaSalle invariance principle [25], the result follows. This means that the malware will disappear with time varying if the basic reproduction number is less than one. \square

Theorem 2. *The malware-existence equilibrium P^* of (1) is globally asymptotically stable if $R_0 > 1$, and unstable if $R_0 < 1$.*

Proof. For system (1), we consider the following Lyapunov function:

$$V = \gamma \left(S - S^* - S^* \ln \frac{S}{S^*} \right) + \gamma \left(E - E^* - E^* \ln \frac{E}{E^*} \right) + (\gamma + \delta + d) \left(I - I^* - I^* \ln \frac{I}{I^*} \right). \quad (14)$$

Then, the derivative of V along solutions of system (1) is

$$V' = \gamma S' \left(1 - \frac{S^*}{S} \right) + \gamma E' \left(1 - \frac{E^*}{E} \right) + (\gamma + \delta + d) I' \left(1 - \frac{I^*}{I} \right). \quad (15)$$

$$V' \leq \gamma S^* I^* [a\beta_1 + (1-a)\beta_2] \left[\left(1 - \frac{S^*}{S} + \ln \frac{S^*}{S} \right) + \left(1 - \frac{I^* E}{IE^*} + \ln \frac{I^* E}{IE^*} \right) + \left(1 - \frac{SE^* I}{S^* I^* E} + \ln \frac{SE^* I}{S^* I^* E} \right) \right]. \quad (17)$$

For the function $f(x) = 1 - x - \ln x$, we know that if $x > 0$, $f(x) \leq 0$, and $x = 1$ leads to $f(x) = 0$. Therefore, we can obtain $V' \leq 0$, and the equality is holding if and only if $S = S^*$, $E = E^*$, $I = I^*$, and $R = R^*$. It means that the largest invariant subset, where $V' = 0$, is P^* . By LaSalle's Invariance Principle [25], P^* is globally asymptotically stable when $R_0 > 1$. This means that the malware will be outbreak if the basic reproduction number is more than one. \square

4. Control Strategy

In this section, we investigate the control strategy from two aspects. Firstly, an optimal control model has been proposed by Pontryagin's Minimum Principle. Secondly, we give some control strategies to prevent the malware outbreak from the explicit expression of the malware spreading threshold value.

By direct calculations, we have that

$$\begin{aligned} V' &= [a\beta_1 + (1-a)\beta_2] \gamma S^* I^* \left(1 - \frac{S^*}{S} \right) \left[1 - \frac{SI}{S^* I^*} \right] \\ &\quad + [a\beta_1 + (1-a)\beta_2] \gamma S^* I^* \left(1 - \frac{E^*}{E} \right) \left[\frac{SI}{S^* I^*} - \frac{E^*}{E} \right] \\ &\quad + (\gamma + \delta + d) \gamma E^* \left(1 - \frac{I^*}{I} \right) \left(\frac{E}{E^*} - \frac{I}{I^*} \right) \\ &\quad - \gamma (d + \eta) \frac{(S^* - S)^2}{S}, \\ &\leq [a\beta_1 + (1-a)\beta_2] \gamma S^* I^* \left(1 - \frac{E^*}{E} \right) \left[\frac{SI}{S^* I^*} - \frac{E^*}{E} \right] \\ &\quad + (\gamma + \delta + d) \gamma E^* \left(1 - \frac{I^*}{I} \right) \left(\frac{E}{E^*} - \frac{I}{I^*} \right) \\ &\quad (a\beta_1 + (1-a)\beta_2) S^* I^* = (\gamma + \delta + d) E^*, \end{aligned} \quad (16)$$

then

4.1. Optimal Control Strategy Formulation. We aim to minimize the number of infected devices (including exposed and infected nodes) and the corresponding cost of the strategies during the process of spreading. Four control functions $u_1(t)$, $u_2(t)$, $u_3(t)$, and $u_4(t)$, where $0 \leq u_i(t) \leq 1$. In particular, $u_i(t) = 0$ means no control strategy and $u_i(t) = 1$ means the maximal use of control strategy. The meanings of $u_i(t)$ are shown as follows:

- (1) $u_1(t)$ is used to represent the quarantine strategy that aims to reduce the contact between wireless devices with weak spreading capabilities and susceptible devices at time t
- (2) $u_2(t)$ is used to represent the quarantine strategy that aims to reduce the contact between wireless devices with strong spreading capabilities and susceptible devices at time t

- (3) $u_3(t)$ is used to represent the vaccination strategy that can improve the immunocompetence of susceptible devices at time t
- (4) $u_4(t)$ is used to represent the repairing and fixing strategy that can increase the recovery rate of infected wireless devices at time t

The transmission dynamics of the optimal control model is formulated as

$$\begin{cases} \frac{dS}{dt} = \Lambda - (1 - u_1(t))a\beta_1SI - (1 - u_2(t))(1 - a)\beta_2SI - (d + u_3(t))S, \\ \frac{dE}{dt} = (1 - u_1(t))a\beta_1SI + (1 - u_2(t))(1 - a)\beta_2SI - (\gamma + \delta + d)S, \\ \frac{dI}{dt} = \gamma E - (\theta + d + u_4(t))I, \\ \frac{dR}{dt} = \delta E + u_4(t)I + u_3(t)S - dR. \end{cases} \quad (18)$$

The main purpose is to minimize the number of infected devices at a minimum cost. And, as a consequence, we consider the objective functional:

$$J(u_1, u_2, u_3, u_4) = \int_0^T \left[E + I + I + \frac{K}{2}u_1^2 + \frac{W}{2}u_2^2 + \frac{P}{2}u_3^2 + \frac{Q}{2}u_4^2 \right] dt, \quad (19)$$

where the parameters $K \geq 0$, $W \geq 0$, $P \geq 0$, and $Q \geq 0$ are the weight constants for the control strategies. $(K/2)u_1^2(t)$, $(W/2)u_2^2(t)$, $(P/2)u_3^2(t)$, and $(Q/2)u_4^2(t)$ describe the cost associated with quarantine strategy, quarantine strategy, vaccination, and repairing and fixing strategies, respectively. Our aim is to seek the optimal control functions $\{u_1^*, u_2^*, u_3^*, u_4^*\}$, such that

$$J(u_i^*) = \min\{J(u_i): u_i \in U, i = 1, 2, 3, 4\}, \quad (20)$$

where U is the control function set defined as $U = \{u_i: u_i(t) \text{ is Lebesgue measure, } 0 \leq u_i(t) \leq 1, t \in [0, T]\}$.

Next, we discuss the existence of an optimal control functions by Fleming and Rishel [26].

- (1) The set of controls and the corresponding state variables are nonempty
- (2) The admissible control set is convex and closed
- (3) The right hand side of the optimal control system is bounded by a linear function in the state and control variables

(4) The integrand of the objective function, $E + I + (K/2)u_1^2 + (W/2)u_2^2 + (P/2)u_3^2 + (Q/2)u_4^2$, is convex

(5) There exist constants $c_1 > 0$, $c_2 > 0$, and $l > 1$ such that the integrand of the objective cost functional is convex and satisfied $J(E, I, u_i, t) \geq c_1 (\sum_{i=1}^4 |u_i|^2)^{l/2} + c_2$

Theorem 3. An optimal control pair $(u_1^*, u_2^*, u_3^*, \text{ and } u_4^*)$ subject to system (18) exists if the following conditions hold:

Proof. By the results of [27], it is easy to check that the set of controls and corresponding state variables exist. By the definition, the control set is bounded and convex. Since optimal control system (18) is bilinear in u_i , the right hand side of it satisfies condition 3 by using the boundedness of the solutions. Additionally, the integrand of objective function (19) is convex on the control set U . $E + I + (K/2)u_1^2 + (W/2)u_2^2 + (P/2)u_3^2 + (Q/2)u_4^2 \geq c_1 (|u_1|^2 + |u_2|^2 + |u_3|^2 + |u_4|^2)^{l/2} + c_2$ because the state variables are bounded, considering $l=2$ and c_1 and c_2 are smaller enough. \square

In order to find the optimal control solution, we should describe the Lagrangian and Hamiltonian function of control system (18). Let $\mathbf{x} = (S, E, I, R)$, and the Lagrangian of the control system is

$$L(\mathbf{x}, U) = E + I + \frac{K}{2}u_1^2 + \frac{W}{2}u_2^2 + \frac{P}{2}u_3^2 + \frac{Q}{2}u_4^2. \quad (21)$$

Then, we define the Hamiltonian function H as

$$\begin{aligned} H(\mathbf{x}, U, \lambda) = & E + I + \frac{K}{2}u_1^2 + \frac{W}{2}u_2^2 + \frac{P}{2}u_3^2 + \frac{Q}{2}u_4^2 \\ & + \lambda_1\{\Lambda - (1 - u_1(t))a\beta_1SI - (1 - u_2(t))(1 - a)\beta_2SI - (d + u_3(t))S\} \\ & + \lambda_2\{(1 - u_1(t))a\beta_1SI + (1 - u_2(t))(1 - a)\beta_2SI - (\gamma + \delta + d)S\} \\ & + \lambda_3[\gamma E - (\theta + d + u_4(t))I] \\ & + \lambda_4[\delta E + u_4(t)I + u_3(t)S - dR]. \end{aligned} \quad (22)$$

Next, we show the following theorem.

Theorem 4. Given an optimal control pair $(u_1^*, u_2^*, u_3^*, u_4^*)$ and a solution $(S^\circ, E^\circ, I^\circ, R^\circ)$ of corresponding control

system (18), there exists an adjoint variable $\lambda_i, i = 1, \dots, 4$, satisfying

$$\begin{aligned} \frac{d\lambda_1}{dt} &= (\lambda_1 - \lambda_2)[(1 - u_1)a\beta_1I + (1 - u_2)(1 - a)\beta_2] + (d + u_3)\lambda_1 - u_3\lambda_4, \\ \frac{d\lambda_2}{dt} &= -1 + \lambda_2(d + \gamma + \delta) - \lambda_3\gamma - \lambda_4\delta, \\ \frac{d\lambda_3}{dt} &= -1 - u_4\lambda_4 + (\lambda_1 - \lambda_2)S[(1 - u_1)a\beta_1 + (1 - u_2)(1 - a)\beta_2] + (\theta + d + u_4)\lambda_3, \\ \frac{d\lambda_4}{dt} &= d\lambda_4, \end{aligned} \quad (23)$$

with the transversality condition

$$\lambda_1(T) = \lambda_2(T) = \lambda_3(T) = \lambda_4(T) = 0. \quad (24)$$

Furthermore, by the necessary condition, we have

$$u_i^* = \max[0, \min(\tilde{u}_i, 1)], \quad i = 1, 2, 3, \text{ and } 4, \quad (25)$$

where

$$\begin{aligned} \tilde{u}_1 &= \frac{(\lambda_2 - \lambda_1)a\beta_1I^\circ S^\circ}{K}, \\ \tilde{u}_2 &= \frac{(\lambda_2 - \lambda_1)(1 - a)\beta_2I^\circ S^\circ}{W}, \\ \tilde{u}_3 &= \frac{(\lambda_1 - \lambda_4)S^\circ}{P}, \\ \tilde{u}_4 &= \frac{(\lambda_3 - \lambda_4)I^\circ}{Q}. \end{aligned} \quad (26)$$

$$\begin{aligned} \frac{\partial H}{\partial S} &= (\lambda_2 - \lambda_1)[(1 - u_1)a\beta_1I + (1 - u_2)(1 - a)\beta_2] \\ &\quad - (d + u_3)\lambda_1 + u_3\lambda_4, \\ \frac{\partial H}{\partial E} &= 1 - \lambda_2(d + \gamma + \delta) + \lambda_3\gamma + \lambda_4\delta, \\ \frac{\partial H}{\partial I} &= 1 + u_4\lambda_4 + (\lambda_2 - \lambda_1)S[(1 - u_1)a\beta_1 + (1 - u_2)(1 - a)\beta_2] \\ &\quad - (\theta + d + u_4)\lambda_3, \\ \frac{\partial H}{\partial R} &= -d\lambda_4. \end{aligned} \quad (27)$$

According to Pontryagin's Maximum Principle [28], we obtain

$$\frac{d\lambda}{dt} = -\frac{\partial H(\mathbf{x}, U, \lambda)}{\partial \lambda}, \quad (28)$$

Proof. Calculating the partial derivatives of the five states of Hamiltonian function (22), respectively, we have

where $\mathbf{x} = (S, E, I, R)$. Hence, the adjoint variable $\lambda_i, i = 1, \dots, 4$, satisfies (23).

Consider control system (18) with $\lambda_i(T) = 0$. By using the optimal necessary condition, we have

$$\frac{\partial H}{\partial u_i^*} = 0. \quad (29)$$

On the interior of the control set space, we can obtain optimal control pair solution (25). \square

4.2. Control Strategy Based on the Basic Reproductive Number. There is an important epidemiological threshold R_0 in the epidemic model. As we discussed in Section 3, the threshold value plays a critical role to control the malware outbreak. Moreover, we obtain the local and global stability of the worm-free equilibrium when $R_0 \leq 1$. Consequently, it is crucial to reduce the value of R_0 below 1, as a result of that to design the efficient security countermeasures to prevent malware outbreak.

From equation (19), we obtain $(\partial R_0 / \partial \beta_1) > 0$, $(\partial R_0 / \partial \beta_2) > 0$, $(\partial R_0 / \partial \Lambda) > 0$, $(\partial R_0 / \partial \gamma) > 0$, $(\partial R_0 / \partial a) < 0$, $(\partial R_0 / \partial \eta) < 0$, $(\partial R_0 / \partial \delta) < 0$, $(\partial R_0 / \partial \varphi) < 0$, $(\partial R_0 / \partial \theta) < 0$, and $(\partial R_0 / \partial d) < 0$, which mean that R_0 increases as the parameters Λ , γ , β_1 , and β_2 increase or the parameters a , d , δ , θ , φ , and η decrease. Figure 2 further describes the trend of R_0 over time with different parameters δ and a . However, the value of d and θ is usually constant because of the properties of wireless devices. Consequently, we control the following coefficients to make $R_0 \leq 1$:

- (i) To decrease the transmission rate of infectious devices with weak and strong spreading capabilities by increasing the security background of the consumer
- (ii) To ensure configuration integrity and wipe out potential malicious software by taking into account an efficient defense mechanism

5. Experimental Validation for the HSEIR Model

In this section, we simulate the HSEIR model via MATLAB R2018a. We illustrate the theoretical results for system (1) by numerical simulations. Besides, by Forward-Backward Sweep Method, we obtain the optimal control strategies and show their effectiveness via comparing the trajectories of infected devices with optimal control and without optimal control.

5.1. Stability of Worm-Free and Worm-Existence Equilibrium. In the real world, there are not much data about malware attacks against the IoT wireless network. Following [15, 23], we attempt to choose appropriate parameter values from the Slammer worm. Assume model parameters $\Lambda = 0.75$, $d = 0.001$, $\theta = 0.005$, $\beta_1 = 0.00001$, $\beta_2 = 0.0001$, $\gamma = 0.004$, $\eta = 0.0006$, $\delta = 0.003$, $a = 0.6$, and $\varphi = 0.005$. The initial numbers of susceptible, exposed, infected, and recovered devices are $S(0) = 350$, $E(0) = 80$, $I(0) = 20$, and $R(0) = 300$. By simple computing, we obtain that $R_0 = 0.9801 < 1$, the worm-free equilibrium $P_0 = (468.75, 0, 0, 282.15)$. Thus, by

Theorem 1, the malware-free equilibrium is globally asymptotically stable, which means that when the time t goes to infinity, the exposed and infected nodes will approach to 0, while the susceptible and recovered nodes will be 468.75 and 282.15. Figure 3 shows that the stable behavior of malware-free equilibrium when $R_0 = 0.9801 < 1$. From Figure 3, we can see that there only exist susceptible nodes and recovered nodes, in accordance with the conclusion in Theorem 1.

Secondly, we set the infected rate $\beta_1 = 0.0001$ and $\beta_2 = 0.001$, and the other parameters are the same as above. By equations (7) and (8), we obtain that $R_0 = 9.801 > 1$, and the malware-existence equilibrium $E^* = (30.61, 47.83, 84.18, 434.31)$. Besides, by Theorem 2, the malware-existence equilibrium is globally asymptotically stable, which means that when the time t goes to infinity, the number of $S(t)$, $E(t)$, $I(t)$, and $R(t)$ is 47.83, 84.18, 30.61, and 434.31, respectively. Figure 4 shows that the stable behavior of malware-existence equilibrium when $R_0 = 9.801 > 1$. From Figure 4, we can see that infected nodes (including exposed and infectious), as well as susceptible and recovered nodes, persist at the endemic level.

5.2. Sensitivity Analysis of Infected Rate of Heterogeneity Devices. In this section, we do some sensitivity analysis of parameters a , β_1 , and β_2 to observe the malware spreading scale. Firstly, we set $a = 0.2, 0.4, 0.6$, and 0.8 , while keeping other parameters the same as those in 5.1. In Figure 5, we can see that when the proportion of wireless devices infected by weak spreading capabilities nodes decreases, that is, the proportion of wireless devices infected by strong spreading capabilities nodes increases, and the speed of malware spreading goes fast, which also causes an increase in the malware spreading scale.

Secondly, we assume 0.0001 as the interval for the infection rate of devices with weak spreading capabilities to compare the malware spreading scale, as shown in Figures 6 and 7. We can see that when $a = 0.2$, parameter β_1 has little effect on the scale and speed of malware transmission, but when $a = 0.8$, the larger the parameter β_1 , the faster the malware spreading and the larger of the number of infected nodes simultaneously. In Figures 8 and 9, let 0.01 be the interval for the infection rate of devices with strong spreading capabilities. The malware spreading speed as well as the spreading scale are increasing with β_2 increasing when $a = 0.2$ and $a = 0.8$. Figure 5 shows that when the proportion of wireless devices that infected by strong spreading capability nodes increases, the number of infected wireless devices also increases. When the proportion of wireless devices infected by strong spreading capabilities nodes remains unchanged, the infected rate β_2 is the main factor in malware spreading. Thus, we must control the proportion of wireless devices with strong spreading capabilities and spread the patch to them as soon as possible.

5.3. Simulation of Optimal Control Strategies. In [29], Lenhart and Workman combined the Runge-Kutta fourth-

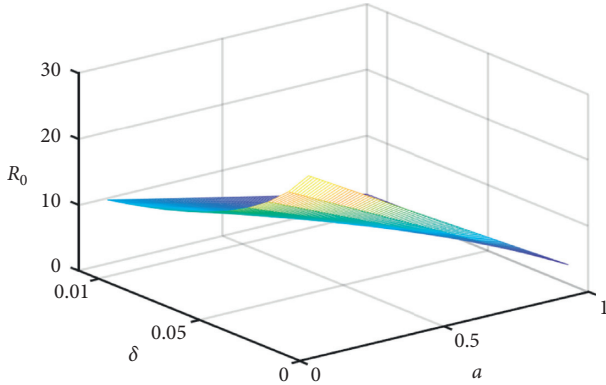


FIGURE 2: Analysis of R_0 .

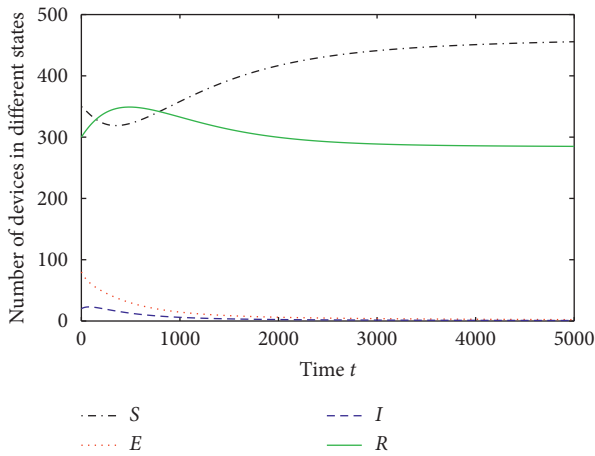


FIGURE 3: Dynamics of the malware-free equilibrium.

order schemes and Forward-Backward Sweep Method to get the optimal solution. Based on their method, we conduct some numerical simulations to illustrate the effectiveness of the optimal control theoretic approach by using a MATLAB code. Considering the limitation of technology and cost, we set $u_{1_{\max}} = 0.8$, $u_{2_{\max}} = 0.8$, $u_{3_{\max}} = 0.7$, $u_{4_{\max}} = 0.9$, $K = 0.3$, $W = 0.6$, $P = 0.4$, and $Q = 0.5$, and the initial numbers and the other parameters are taken as the same as those in V.A. As shown in Figure 10, we give the optimal control strategies. We observe that $u_1(t)$ and $u_2(t)$, namely, quarantine strategies to infected wireless devices with weak spreading capabilities and strong spreading capabilities, could be reduced 3 seconds later from the beginning of the malware outbreak, which saves much of quarantine costs. This conclusion is consistent with the use of wireless IoT devices that disconnected from the wireless network for a long time are not allowed. Equally, after 8 seconds, $u_3(t)$, vaccination strategy, could be canceled gradually. Different from the abovementioned control strategies, $u_4(t)$, repairing and fixing strategy, for example, applying antivirus patches, would be keep going. In Figures 11 and 12, we illustrate the trend of the number of exposed devices and infected devices over time with control and without control, respectively. It is clear that the exposed devices, as well as the infected devices, with control are much smaller than those without control,

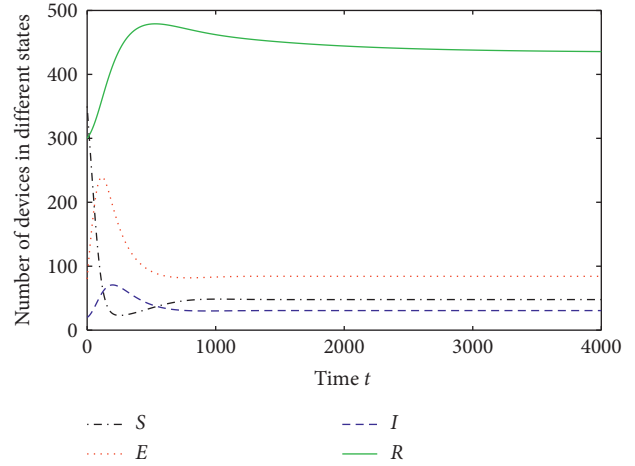


FIGURE 4: Dynamics of the malware-existence equilibrium.

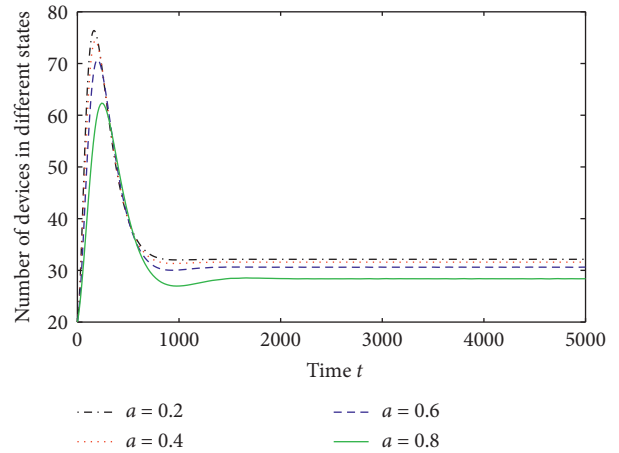
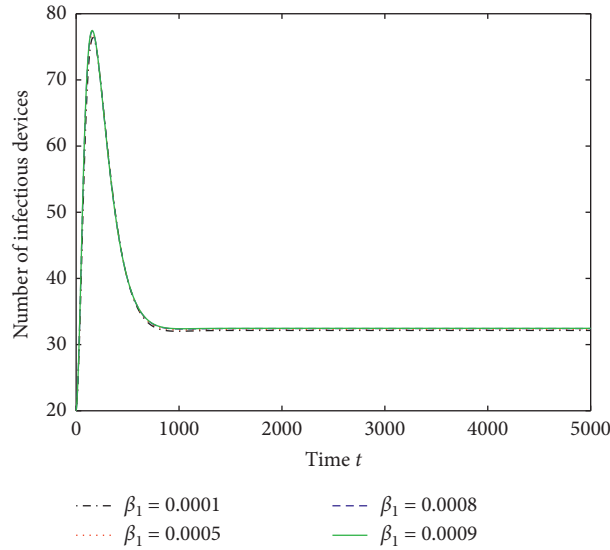
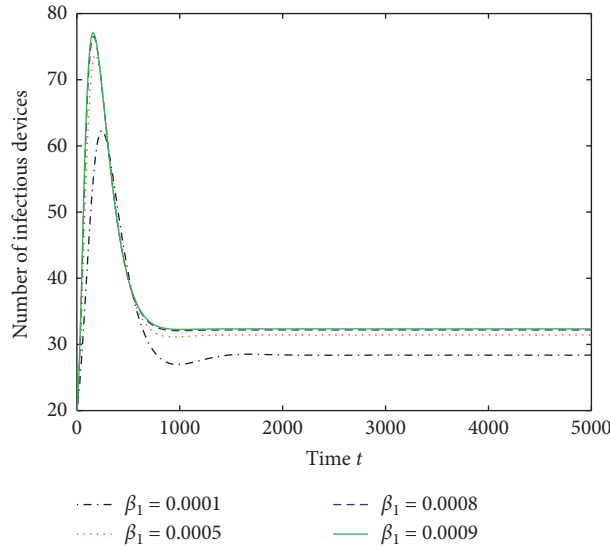


FIGURE 5: The state transition in the HSEIR model.

and we suggest that early control strategies play a significant role on reducing the number of infected devices remarkably, which are also cost-effective optimal strategies.

6. Conclusions and Discussion

In this paper, we have proposed a new HSEIR model to investigate the malware propagation in wireless IoT networks, while considering the heterogeneity of infected wireless IoT devices. According to the ability of wireless devices on malware spreading, devices are divided into two different level groups in a fuzzy way. Based on the proposed model, we obtained the basic reproduction number R_0 , which represents the malware spreading threshold. Moreover, we analyzed the final size of malware propagation under special cases. Numerical simulations vividly illustrate the main results of stability analysis for system (1). Our simulations also show that when the proportion of wireless devices infected by strong spreading capabilities nodes increases, the malware spreading scale will also increase. In addition, the proportion of susceptible wireless devices infected by weak spreading capabilities or strong spreading capabilities devices also affects the malware propagation size.

FIGURE 6: Parameter β_1 at $a = 0.2$.FIGURE 7: Parameter β_1 at $a = 0.8$.

These results will provide some useful insights on preventing the global spread of malware.

Furthermore, aiming to minimize the costs of control strategy as well as minimize the infected mobile devices, we obtained the optimal control solution by the Forward-Backward Sweep Method. At the beginning of the malware outbreak, the strategies of quarantine and vaccination can effectively control the malware propagation. As for repairing and fixing strategy, it is the essential method to control malware spreading and reduce the death rate due to the malware attack. More numerical results show the effectiveness of the optimal control strategy. Also, the analysis of

R_0 allows us to give the efficient malware-epidemic control strategies to prevent the malware propagation through IoT wireless networks, including decreasing the transmission rate of infectious devices with weak and strong spreading capabilities by increasing the security background of the consumer and to ensure configuration integrity and wipe out potential malicious software by taking into account an efficient defense mechanism.

Although we have investigated the issue of the heterogeneity of wireless IoT devices in malware spreading, there are still some problems in this paper to be further solved.

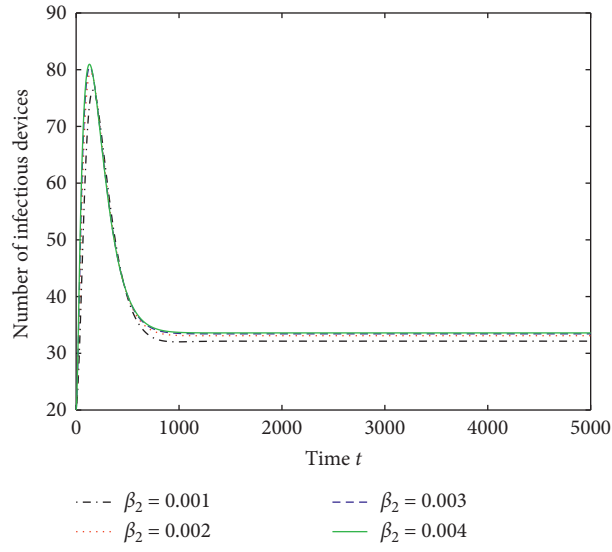


FIGURE 8: Parameter β_2 at $a = 0.2$.

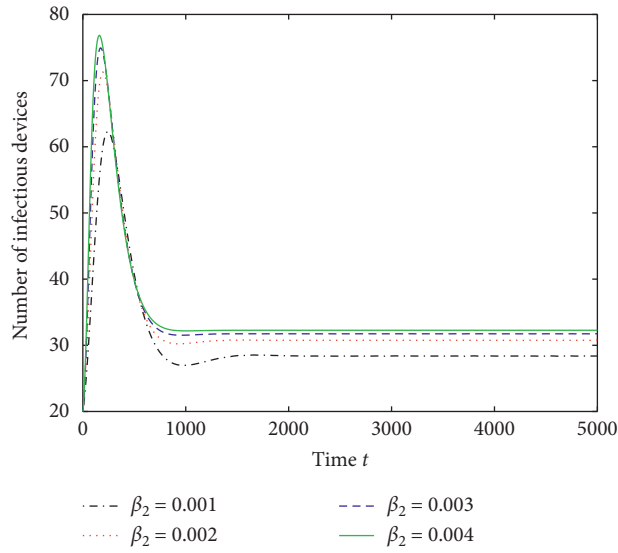


FIGURE 9: Parameter β_2 at $a = 0.8$.

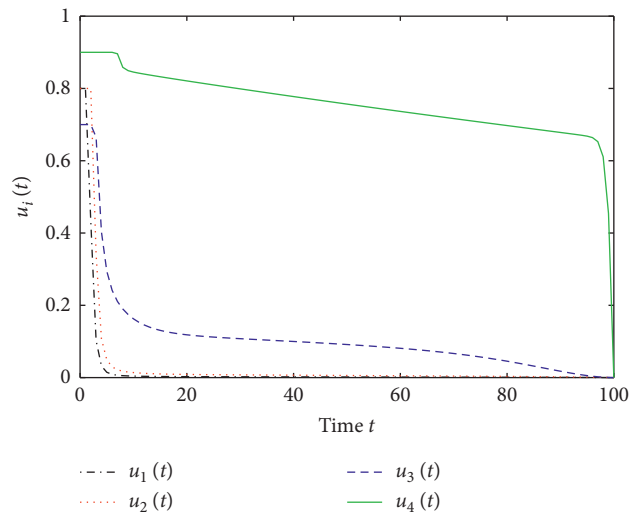


FIGURE 10: The graph trajectories of four optimal control strategies.

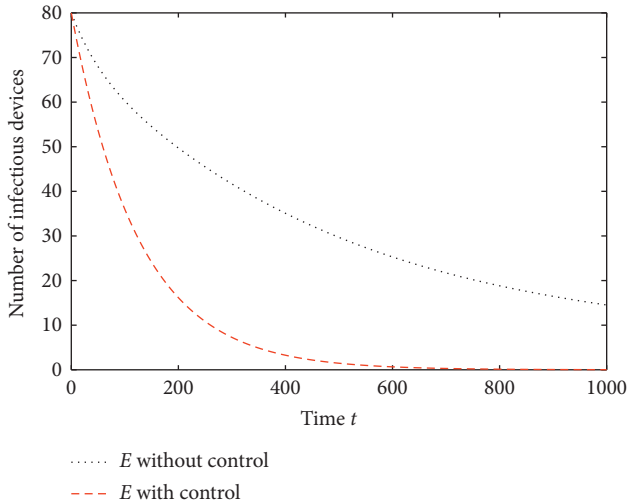


FIGURE 11: A comparison of $E(t)$ with optimal control and without optimal control.

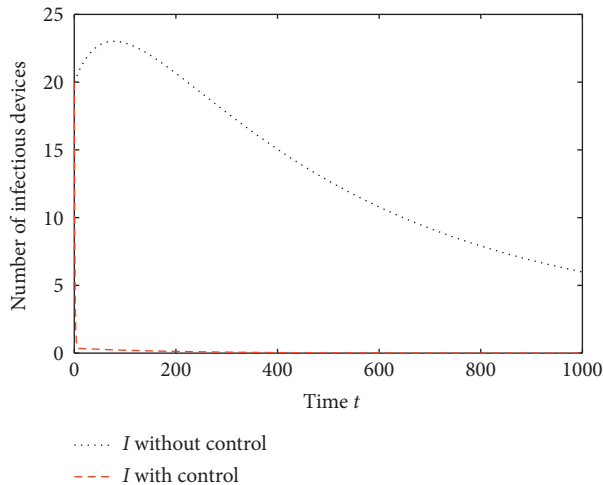


FIGURE 12: A comparison of $I(t)$ with optimal control and without optimal control.

- (i) How to determine the transmission capability of the infected wireless devices more accurately is still a question. In this paper, we only divided them into two different level groups in a fuzzy way.
- (ii) It is significant to study the case where there is a limited supply of control strategies at each instant of time.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 61772478.

References

- [1] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security-a survey," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1802–1831, 2017.
- [2] B. Kuang, A. Fu, L. Zhou et al., "DO-RA: data-oriented runtime attestation for IoT devices," *Computers and Security*, vol. 97, 2020.
- [3] Y. Hou and J. Wang, "Investigation of wireless sensor network of the Internet of Things," in *Proceedings of the International Conference on Intelligent and Interactive Systems and Applications*, pp. 21–29, London, UK, 2019.
- [4] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of things: architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, 2017.
- [5] J. Wu, W. Jiang, Y. Mei, Y. Zhao, and T. Wang, "A survey on the progress of testing techniques and methods for wireless sensor networks," *IEEE Access*, vol. 7, pp. 4302–4316, 2018.
- [6] M. J. Farooq and Q. Zhu, "Modeling, analysis, and mitigation of dynamic botnet formation in wireless IoT networks," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 9, pp. 2412–2426, 2019.
- [7] H. Elsawy, M. A. Kishk, and M. S. Alouini, "Spatial firewalls: quarantining malware epidemics in large scale massive wireless networks," *IEEE Communications Magazine*, vol. 58, no. 9, pp. 32–38, 2020.
- [8] Y. Li, G. Xu, H. Xian, L. Rao, and J. Shi, "Novel android malware detection method based on multi-dimensional hybrid features extraction and analysis," *Intelligent Automation and Soft Computing*, vol. 25, no. 3, pp. 635–644, 2019.
- [9] C. Du, S. Liu, L. Si, Y. Guo, and T. Jin, "Using object detection network for malware detection and identification in network traffic packets," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1785–1796, 2020.
- [10] D.-W. Kim, G.-Y. Shin, and M.-M. Han, "Analysis of feature importance and interpretation for malware classification," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 1891–1904, 2020.
- [11] J. Liu, Y. Zeng, J. Shi, Y. Yang, R. Wang, and L. He, "Mal-Detect: a structure of encrypted malware traffic detection," *Computers, Materials and Continua*, vol. 60, no. 2, pp. 721–739, 2019.
- [12] M. S. Arif1, A. Raza, M. Rafiq, M. Bibi, and J. N. Abbasi, "Numerical simulations for stochastic computer virus propagation model," *Computers, Materials and Continua*, vol. 62, no. 1, pp. 61–77, 2020.
- [13] M. Knysz, X. Hu, Y. Zeng, and K. G. Shin, *Open WiFi Networks: Lethal Weapons for Botnets?*, Proceedings IEEE Infocom, Orlando, FL, USA, 2012.
- [14] L. X. Yang, P. D. Li, X. F. Yang, Y. Xiang, and Y. Y. Tang, "Simultaneous benefit maximization of conflicting opinions: modeling and analysis," *IEEE Systems Journal*, vol. 14, no. 2, pp. 1623–1634, 2020.

- [15] F. W. Wang, Y. K. Zhang, C. G. Wang, J. F. Ma, and S. J. Moon, "Stability analysis of a SEIQV epidemic model for rapid spreading worms," *Computers and Security*, vol. 29, no. 4, pp. 410–418, 2010.
- [16] J. N. C. Goncalves, H. S. Rodrigues, and M. T. T. Monteiro, "Optimal control measures for a susceptible-carrier-infectious-recovered-susceptible malware propagation model," *Optimal Control Applications and Methods*, vol. 40, no. 4, pp. 691–702, 2019.
- [17] L. P. Song and X. Q. Ding, "Hopf bifurcation of an epidemic model with delay," *PLoS One*, vol. 11, no. 6, 2019.
- [18] D. Acarali, M. Rajarajan, N. Komninos et al., "Modelling the spread of Botnet malware in IoT-based wireless sensor networks," *Security and Communication Networks*, vol. 2019, Article ID 3745619, 2019.
- [19] S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: a survey," *IEEE Communication Surveys and Tutorials*, vol. 16, no. 2, pp. 952–1041, 2014.
- [20] S. M. Cheng, P. Y. Chen, C. C. Lin, and H. C. Hsiao, "Traffic-aware patching for cyber security in mobile IoT," *IEEE Communications Magazine*, vol. 55, no. 7, 2017.
- [21] L. Li, J. Cui, R. Zhang et al., "Dynamics of complex networks: malware propagation modeling and analysis in Industrial Internet of Things," *IEEE Access*, vol. 8, pp. 64184–64192, 2020.
- [22] S. G. Shen, H. P. Zhou, S. Feng, J. H. Liu, H. Zhang, and Q. Y. Cao, "An epidemiology-based model for disclosing dynamics of malware propagation in heterogeneous and mobile WSNs," *IEEE Access*, vol. 8, pp. 43876–43887, 2020.
- [23] Q. W. Gao and J. Zhuang, "Stability analysis and control strategies for worm attack in mobile networks via a VEIQS propagation model," *Applied Mathematics and Computation*, vol. 368, Article ID 124584, 2020.
- [24] P. van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, pp. 29–48, 2002.
- [25] J. P. Lasalle, *The Stability of Dynamical Systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1976.
- [26] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin, Germany, 1975.
- [27] D. L. Lukes, *Differential Equations: Classical to Controlled*, Elsevier, Amsterdam, Netherlands, 1982.
- [28] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze et al., *Mathematical Theory of Optimal Processes*, Gordon and Breach Science Publishers, New York, NY, USA, 1986.
- [29] S. Lenhart and J. T. Workman, *Optimal Control Applied to Biological Models*, Chapman and Hall/CRC Press, Boca Raton, FL, USA, 2007.

Research Article

An Adaptive Industrial Control Equipment Safety Fault Diagnosis Method in Industrial Internet of Things

Hanrui Zhang ¹, Qianmu Li ^{2,3}, Shunmei Meng,¹ Zhuoran Xu,¹ and Chaoxian Lv¹

¹School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China

²School of Cyber Science and Engineering, Nanjing University of Science and Technology, Nanjing, China

³Intelligent Manufacturing Department, Wuyi University, Jiangmen, China

Correspondence should be addressed to Qianmu Li; liqianmu@126.com

Received 10 February 2021; Revised 8 April 2021; Accepted 27 May 2021; Published 14 June 2021

Academic Editor: jingyu feng

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

With the rapid development of intelligent manufacturing and Industrial Internet of Things, many industrial control systems have high requirements for the security of the system itself. Failures of industrial control equipment will cause abnormal operation of industrial control equipment and waste of resources. It is very meaningful to detect and identify potential equipment abnormalities and failures in time and implement effective fault tolerance strategies. In the Industrial Internet of Things environment, the instructions and parameters of industrial control equipment often change due to changes in actual requirements. However, it is impractical to customize the learning method for each parameter value. Aiming at the problem, this paper proposes a fault diagnosis model based on ensemble learning and proposes a method of updating voting weights based on dynamic programming to assist decision-making. This method is based on Bagging strategy and combined with dynamic programming voting weight adjustment method to complete fault type prediction. Finally, this paper uses different loads as dynamic conditions; the diagnostic capability of the Bagging-based fault diagnosis integrated model in a dynamically changing industrial control system environment is verified by experiments. The fault diagnosis model of industrial control equipment based on ensemble learning effectively improves the adaptive ability of the model and makes the fault diagnosis framework truly intelligent. The voting weight adjustment method based on dynamic programming further improves the reliability of voting.

1. Introduction

Traditional industrial control system safety fault diagnosis mainly focuses on mechanical failure of industrial control equipment. However, with the rapid development of intelligent manufacturing and Industrial Internet of Things, the links between industrial control equipment have become closer, and the operating environment has become more complicated. This also increases the probability of equipment failure to a certain extent, leading to further increase in equipment management and control costs. In the Industrial Internet of Things environment, the instructions and parameters of industrial control equipment often change due to changes in actual needs. For example, since more industrial tasks need to be completed in a short time, the current, voltage, speed, load, and other parameters of industrial control equipment must be adjusted so

that the operation can be successfully completed. As the actual environmental parameters are changing, the safety fault signal returned by the sensor will also change with the change of operating parameters. These dynamic factors have brought huge challenges to equipment safety fault diagnosis in the industrial control environment. In the environment of Industrial Internet of Things, the parameters are constantly changing, so it is not practical to customize a learning method for each parameter value.

How to intelligently detect the changed safety fault sequence according to the dynamic changes of actual parameters is the focus of this paper.

The main contents of this paper are as follows:

- (i) A set of safety fault diagnosis models based on ensemble learning is proposed. The model has high

accuracy for safety fault diagnosis under the condition that the equipment parameters change.

- (ii) A method for adjusting voting weight based on dynamic programming is proposed. The weights can be evaluated based on the current performance of each learner. The weights are then updated after the evaluation is completed.

2. Related Work

In recent years, a large amount of research has been conducted on intelligent fault diagnosis and prediction in the Industrial Internet of Things. Fault diagnosis and prediction is one of the most important functions in complex and safety-critical engineering systems, especially fault diagnosis, which has been the subject of in-depth research in the past forty years. This capability allows detecting and isolating early developing failures and predicting failure propagation, which can allow preventive maintenance or even as a countermeasure to the possibility of catastrophic events due to failures. Chen et al. [1] proposed a distributed fast fault diagnosis approach for multimachine power systems based on deterministic learning (DL) theory. In the study of Mousavi et al. [2], an efficient strategy for fault detection and isolation (FDI) of an Industrial Gas Turbine is introduced based on ensemble learning methods. Soleimani et al. [3] studied the impact of early stages insulation deteriorations on the temperature inside the transformer using a finite-element electromagnetic-thermofluid method and proposed an online sensor-based decision-making predictive fault diagnosis approach based on the observations. Kumar et al. [4] built up on the ideas of inference-based ambiguity management in the setting of decentralized control and developed a framework for inference-based decentralized diagnosis. Guan et al. [5] proposed a control and protection framework based on multiagent system (MAS), in which situation awareness of zone agents plays an important role.

Ensemble learning is a very popular research direction in the field of machine learning (Dietterich, 2002) [6]. Its core idea is to build multiple models and then merge their decision results to get better results than a single model. Common ensemble learning methods include Bagging (Breiman, 1996) [7], Boosting (Freund, 1996) [8], Random Forest (RF) (Breiman, 2001) [9], XGboost (Chen, 2016) [10], and GBDT (Friedman, 2001) [11]. There are many similarities between integrated learning and information fusion. Information fusion is mainly divided into three categories: data layer fusion, feature layer fusion, and decision layer fusion. Integrated learning includes the content of decision fusion. Integrated learning has two important steps: one is the construction of the base model and the other is the selection of the fusion method. A good fusion method can effectively improve the effect of the integration.

Next, we summarized the base model selection method and decision fusion method.

Ensemble learning can be viewed as a multimodel system. There are two main factors that affect the integration effect: one is the performance of a single model; the other is the diversity between models. The greater the difference

between the models, the more obvious the effect after integration. For improving model diversity, there are mainly the following methods:

- (1) Use different training data sets to train the same type of model [12, 13]. Bagging and random forest methods both generate different base models in this way. In this way, different basic models are generated. The above two methods generate different training data sets by resampling the original training data set. Then, the same base model is trained separately. For example, the Random Forest uses different training sets to train the decision tree model. Bagging can choose any base model, such as a decision tree model or a neural network model [14, 15].
- (2) The model is trained through different feature subsets, and the Random Forest is building a decision tree at each node, and a different feature subset is randomly selected [16]. This further enhances the difference between decision trees and can achieve better integration performance [17, 18].
- (3) Completely different types of models are used as the base model [19, 20].
- (4) Different model parameters are used to construct the base model [21].

Most of the existing ensemble learning methods construct the base model through the above 4 methods.

Decision fusion methods can mainly be classified into two categories: one is the fusion method based on label output and the other is the fusion method based on probability output. The label-based methods mainly include voting method (Lam, 1997) [22], Borda counting method (Emerson, 2013) [23], and the behavioral knowledge space method (BKS) (Huang, 1993) [24]. The probability-based methods mainly include the decision template method (Kuncheva, 2001) [25], the Dempster-Shafer (DS) evidence fusion method (Sentz, 2002) [26], and Bayesian fusion method (Stathaki, 2011) [27].

3. Safety Fault Detection Model of Industrial Control Equipment Based on Integrated Learning

In order to be able to adaptively face the changes of the industrial control system environment, this paper proposes a safety fault diagnosis model for industrial control equipment based on integrated learning. The model is a composite model composed of multiple individual safety fault classifiers. Individual safety fault classifiers vote on safety fault detection results, and safety fault combination classifiers perform safety fault diagnosis based on the voting results of each classifier. Figure 1 shows the specific structure of the safety fault classifier based on the composite model. Compared with individual safety fault classifiers [28, 29], safety fault classifiers based on composite models tend to have more accurate results.

For the safety fault detection of industrial control equipment, let the set of safety fault types be

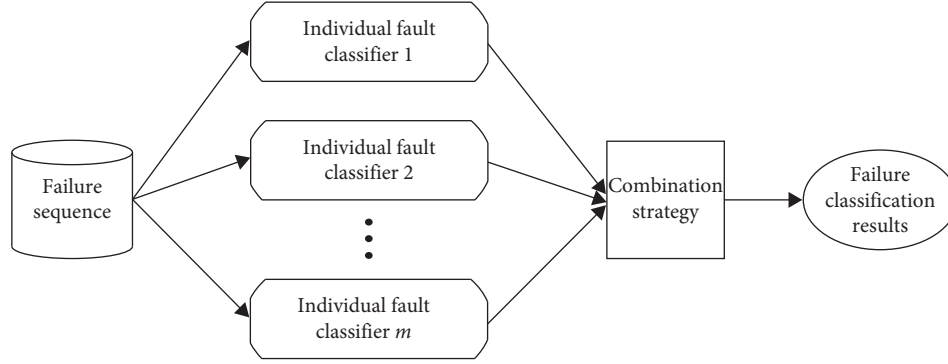


FIGURE 1: Structure of the safety fault detection model based on ensemble learning.

$y = \{c_1, c_2, \dots, c_n\}$, the true function of safety fault classification be f , and the error rate of the individual safety fault classifier be ε . Then, for each individual safety fault classifier $h_i, i \in (1, 2, \dots, m)$, there is

$$P(h_i(x) \neq f(x)) = \varepsilon. \quad (1)$$

Here, m is the number of individual safety fault classifiers and is an odd number, $c_i, i \in \{1, 2, \dots, n\}$ is the safety fault type, and x is the input safety fault sequence. When more than one-half of the individual safety fault classifiers can correctly classify the safety fault sequence, the safety fault integrated classifier can correctly classify the following:

$$h(x) = \text{sign}\left(\sum_{i=1}^m h_i(x)\right). \quad (2)$$

Assume that the error rate of each safety fault classifier is independent of each other. Then, according to the Hoeffding inequality, it can be known that the detection error rate of the safety fault combination classifier is as follows:

$$P(h(x) \neq f(x)) = \sum_{k=0}^{m/2} \binom{m}{k} (1-\varepsilon)^k \varepsilon^{m-k} \leq \exp\left(-\frac{1}{2}m(1-2\varepsilon)^2\right). \quad (3)$$

It can be known from formula (3) that when the number of individual safety fault classifiers increases, the error rate of the safety fault combination classifier will become lower and lower.

According to whether there is a strong correlation between the individual learners, the current integrated learning methods can be divided into two categories. In the first category, the individual learners have a strong dependency relationship. In this case, serial generation must be used to integrate learning. The main representative of this method is Boosting. On the other hand, the dependency relationship between individual learners is low, and it does not have strong correlation. In this case, use a parallelization method that can be generated simultaneously. The main representatives of this method are Bagging and Random Forest.

Since the safety fault diagnosis in the industrial control environment has high real-time requirements, if the serial inheritance method is used, the result selection of each

round of the training set is related to the learning results of the previous rounds. This process has a large time overhead, and it is difficult to meet the real-time requirements of industrial control systems. Therefore, this paper adopts a parallelized integrated learning model to adaptively diagnose industrial control equipment safety faults.

3.1. Integrated Model of Safety Fault Diagnosis Based on Bagging. In the safety fault diagnosis of industrial control systems, if a safety fault diagnosis integrated model with strong generalization ability is needed, the individual safety fault learners in the integrated model should be as independent as possible from each other [30-32]. However, due to the impact of the industrial control environment, it is almost impossible to achieve independence from each other in actual situations. How to make a relatively large difference between safety faulty individual learners has become the key to research.

As a representative of parallelized integrated learning, the Bagging algorithm has broad application prospects in the industry. Assume that the samples of the safety fault data set in the current industrial control environment are all safety fault sequences under the parameter θ_t , where $i \in \{1, 2, \dots, n\}$. The safety fault data under the parameter θ_t need to be detected, where $t \notin \{1, 2, \dots, n\}$.

The basic process of the fault diagnosis method based on the Bagging algorithm is as follows.

First, the training set composed of device state data which are divided and sampled, and several different sub-training sets are obtained. Then, the safe failure data set expansion method based on periodic overlap sampling is used to expand the safe failure data set. Next, through deep model convolution training, we train multiple individual fault learners from these subtraining sets. Then, we input the failure training under the changing parameters. Through voting decisions, we can get the fault category output.

The basic flow of the safety fault diagnosis method based on the Bagging algorithm is shown in Figure 2.

First, the safety fault data set augmentation method based on periodic overlapping sampling is used to expand the safety fault data set. It is assumed that the enhanced data set contains m safety fault samples. A random sample is selected from the safety faulty data set and added to the

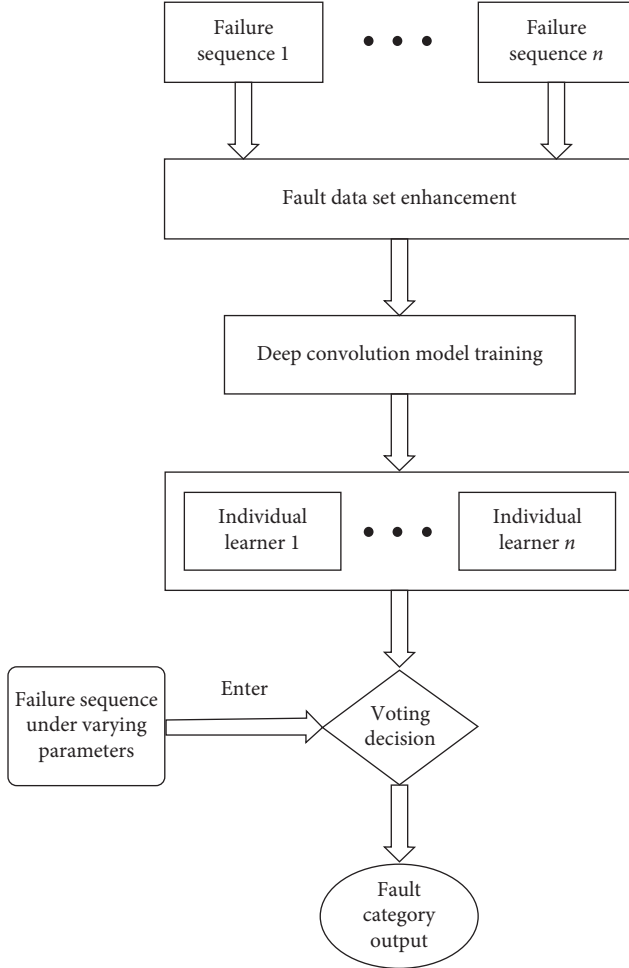


FIGURE 2: Bagging-based safety fault diagnosis process.

sampling set. After the selection is completed, the safety faulty data set is returned. In this way, the sample may be selected next time. After repeating this process m times, a sampling set containing m samples can be obtained. In this way, some samples in the initial failure data set will appear multiple times in the sampling set, while some will not appear in the sampling set. With this sampling method, the probability that a sample will never be selected is

$$f(m) = \left(1 - \frac{1}{m}\right)^m, \quad (4)$$

$$\lim_{m \rightarrow \infty} \left(1 - \frac{1}{m}\right)^m \mapsto \frac{1}{e} \approx 0.368.$$

It can be seen from Figure 3 that when the number of samples m of the safety faulty data set is sufficiently large, the probability that a sample is not always sampled tends to be stable. It can be known from formula (4) that with the increase in the safety faulty data set, about 63.2% of the samples appear in the sampling set. By repeating the above process k times, respectively, we can get k sample sets containing m samples.

Using the convolutional neural network-based safety fault diagnosis method as an individual learning algorithm,

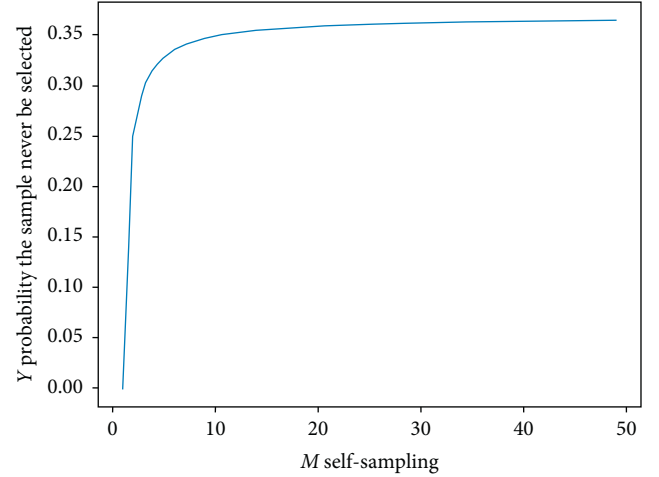


FIGURE 3: Self-sampling probability map.

the sample set under each parameter is trained to obtain an individual learner. Then, these individual safety fault learners are combined to predict the safety fault category through a voting strategy. If there are multiple categories with the highest number of votes after voting, a failure type is randomly selected as the algorithm output.

The complexity analysis of the safety fault detection method based on Bagging is performed below. Assume that the computational time complexity of each individual safety fault learner is $O(m)$. Because the Bagging method supports parallel computing, it is assumed that the time for context switching between individual safety fault learners is t_0 and the time chosen for voting is t_1 . Therefore, the time complexity of the Bagging algorithm is $k(O(m) + t_0n + t_1)$. In fact, compared with the calculation time of individual learners, the time for context switching and the time for voting selection in Bagging parallel computing are negligible. Therefore, in terms of time complexity, the Bagging-based safety fault diagnosis integration method has only a constant level gap compared with individual learners and has higher performance.

3.2. Combination Strategy of Safety Fault Diagnosis Based on Voting Method. Common basic integration strategies include voting, averaging, and learning. These methods are relatively simple but very powerful. For classification problems such as industrial control equipment safety fault diagnosis, the combination strategy based on the voting method has been widely used and has achieved good results. The voting method adopts the principle that the minority is subject to the majority. The individual safety fault classifier needs to select a predictive label from $y = \{c_1, c_2, \dots, c_n\}$. Further, the voting method can be subdivided into absolute majority voting method, relative majority voting method, and so on.

For the absolute majority voting method, for a certain safety fault category c_i , if it gets more than half of the individual safety fault classifiers, then this category is used as the output category; otherwise, prediction is rejected. The formula is as follows:

$$h(x) = \begin{cases} c_j, & \text{count}(h_i(x) = c_j) > 0.5\text{count}(*), \\ \text{reject}, & \text{otherwise.} \end{cases} \quad (5)$$

Here, count represents the counting function and * represents all prediction results.

$$h(x) = c_{\underset{j}{\text{argmax}}(\text{count}(h_i(x)=c_j))}, \quad i \in \{1, 2, \dots, m\}, j \in \{1, 2, \dots, n\}. \quad (6)$$

4. Method for Adjusting Voting Weight Based on Dynamic Programming

Each of the existing k sampling sets contains m safety fault samples. Because these sample sets are collected from the initial failure data set and the initial data set contains n failure samples under different parameters, the sample type is $\theta_i, i \in \{1, 2, \dots, n\}$. How to determine the initial weight of the voting weight adjustment method based on dynamic programming and the weight state transition equation is the key to whether the Bagging-based ensemble learning method can adaptively predict the safety fault sequence under unknown parameters.

In theory, each individual classifier should have the same initial weight. In theory, each individual classifier should have the same initial weight because they are all extracted from the initial failure data set. Due to the randomness of the extraction, the proportion of parameters corresponding to the safety fault sequences in the m sampling sets cannot be exactly the same. For the fault sequence to be detected under parameter θ_i , if the distance between θ_i and θ_t is closer, theoretically, the fluctuation and amplitude of the fault sequence θ_t can be considered to be more similar to those of the fault sequence to be detected. In other words, under the θ_i parameter, the individual learner should have better generalization ability, compared with the individual learner under the parameters which are far from θ_t .

Based on the analysis, the initial voting weight determination process of the individual learner is as follows.

First, according to the distance between $\{\theta_1, \theta_2, \dots, \theta_n\}$ and θ_t , set the parameter distance influence coefficient $\{\delta_1, \delta_2, \dots, \delta_n\}$. The closer the parameter θ_i is, the larger the value of the influence coefficient δ_i is, where $i \in \{1, 2, \dots, n\}$.

Second, calculate the proportion of data under each parameter θ_i in each sampling set as $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, $\sum_{i=1}^n \lambda_i = 1$.

Third, the initialization weight parameters are solved, and the solution formula is as follows:

For the relative majority voting method, the strategy will choose the type with the most votes as the final prediction result. If there are multiple safety fault types with the same number of votes and the highest votes, then randomly select one of these safety fault types as the final output, which is

$$\text{weight}_j = \sum_{i=1}^n \lambda_i \delta_i, \quad j \in \{1, 2, \dots, k\}. \quad (7)$$

Here, k is the number of individual failure learners. At this point, the determination of the initial parameters of the voting weight is completed. Next, the state transition equation of the voting weight parameters is determined.

The weighted state of each individual safety fault learner vote is transferred according to the previous safety fault detection results. If the previous detection is successful, then the safety faulty individual learner is considered to have a better effect under the parameter θ_i , and the weight of the safety faulty individual learner should be appropriately increased. If the previous safety fault detection fails to predict success, the voting weight of the safety faulty individual learner should be appropriately reduced.

The proportion of weight increase or decrease is also dynamically changed. The proportion of change is determined by the scale of the test data and the accuracy of the safety fault detection. When the data size is small, the proportion of weight adjustment is relatively small regardless of the success or failure of the individual safety fault learner detection. Because the statistical characteristics are not obvious when the data are small, there may be some chance. Conversely, when the data size is large, the overall performance of individual learners can be well reflected. If the safety faulty individual learner has shown good results in the previous detection, if it successfully detects the current test sample, the adjustment proportion of weight increase will also increase; if the current sample fails, the weight reduction is relatively small. If the safety fault learner did not perform well in the previous detection and then if it successfully predicts the current safety fault sample, then the proportion of weight improvement is small. If the current failure test sample prediction fails, the reduction in voting weight is also relatively large. The specific voting weight transfer equation is as follows:

$$W_p = \begin{cases} W_{p-1} + \text{success} \left(p, \frac{\sum_{i=1}^p (h_i^j(x) = f(x))}{p} \right) \\ W_{p-1} - \text{failure} \left(p, \frac{\sum_{i=1}^p (h_i^j(x) \neq f(x))}{p} \right) \end{cases}, \quad j \in \{1, 2, \dots, k\}. \quad (8)$$

Here, W_p is the weight to be calculated, W_{p-1} is the last voting weight, the success function represents a weighted incremental function of the number of test failure samples and detection accuracy when the detection is correct, the failure function represents a weight reduction function between the number of test failure samples and the detection error rate when detecting errors, p represents the number of failure samples currently tested, $h_i^j(x)$ represents the detection result of the j -th individual safety fault learner on the i -th test set, x is the failure data sample space, and $f(x)$ represents the true function of the safety fault category.

In the integrated safety fault diagnosis model based on Bagging, a combination of absolute majority voting and relative majority voting is used to vote. When the absolute majority voting method can directly determine the failure type, the type selected by voting is output; when the absolute majority voting method cannot determine the failure type, the relative majority voting method is used for selection.

5. Experimental Results and Analysis

This paper uses the CWRU bearing fault data set to verify the feasibility of the proposed fault diagnosis method in the experiment. The CWRU bearing failure data set comes from the Laboratory of Case Western Reserve University (<http://www.grouplens.org/node/73>). In this experiment, the data set is divided into normal data, 12K drive end bearing fault data, 48K bearing drive end fault data, and 12K bearing fan end fault data according to different frequencies. For each type of data, fault samples are collected under four different loads (0 hp, 1 hp, 2 hp, and 3 hp). This experiment uses the 12K drive end bearing fault data to carry out the experiment. The fault locations are divided into inner ring faults, outer ring faults, and rolling element faults. The fault diameters of different parts are divided into four types: 0.007 inch, 0.014 inch, 0.021 inch, and 0.028 inch. For outer ring failures, there are no samples with a failure diameter of 0.028 inches. Therefore, with normal data, there are a total of 12 fault states. In order to verify the diagnostic capability of the integrated fault diagnosis model based on Bagging in the dynamic industrial control system environment, this experiment will use different loads as dynamic conditions to verify the diagnostic capability of the model. The specific experimental strategy is as follows.

The safety fault data set used in the experiment contains safety fault samples with load sizes of 0 hp, 1 hp, 2 hp, and 3 hp. This experiment uses safety fault data under three loads for integrated training and uses the trained model to classify the data set under another load. The safety fault data sets are defined under 0 hp, 1 hp, 2 hp, and 3 hp as A, B, C, and D data sets. There are four groups of experiments for verification. The specific experimental groups are shown in Table 1.

Here, there are 18,000 samples of each type of failure for each set of data. In the following, the first group of experiments is taken as an example for analysis. Since the

Bagging-based safety fault diagnosis integrated model uses a voting mechanism for classification, the number of learners to be learned is set to an odd number to facilitate voting. In this experiment, the number of learners to be learned was set to 9. The samples of the three types of data sets A, B, and C are mixed first, and the mixed data set is called S. S is divided into 9 parts, and make sure that each data set does not contain duplicate samples. The sample distribution of each sampling set after the equalization is shown in Table 2.

For each sampled data set, the training set is selected by the self-sampling method and the uncollected data set is used as the validation set. The self-collection of the data set corresponding to each safety fault diagnosis model after self-sampling is shown in Table 3:

As shown in the table, the self-sampled sample data distribution of each individual learner to be learned is obtained. The traditional safety fault diagnosis model based on convolutional neural network learns the training set to obtain each safety faulty individual learner. Then, the existing model is used for ensemble learning. The experimental data are shown in Table 4.

Using the safety fault data sets A, B, and C to diagnose the data of the predicted safety fault data set D, the average diagnostic accuracy of the individual learner is 91.23% and the average accuracy of the integrated learner is 94.93%; with the safety fault data sets A, B, and D, the data of the prediction failure data set C are diagnosed. The average diagnosis accuracy of the individual learner is 91.53%, and the average accuracy of the integrated learner is 95.79%. The failure data sets A, C, and D are used to predict the failure data set B. *Data Diagnosis*. The average diagnosis accuracy rate of individual learners is 91.95%, and the average accuracy rate of integrated learners is 95.86%. The safety fault data sets B, C, and D are used to diagnose the data of predicted safety fault data set A. The average diagnostic accuracy of the individual learner is 91.48%. The diagnostic accuracy is 91.48%, and the average accuracy of the integrated learner is 95.41%. The experimental results show that the safety fault diagnosis method based on ensemble learning can achieve higher diagnostic accuracy under the dynamic changes of equipment parameters.

From the above experimental results, we can clearly see the advantages of the industrial control equipment safety fault detection algorithm based on integrated learning:

- (i) We deal with the dynamic factors of the industrial control system and propose a fault diagnosis model of industrial control equipment based on ensemble learning. This effectively improves the adaptive ability of the model and makes the fault diagnosis framework truly intelligent.
- (ii) We use the voting weight adjustment method based on dynamic programming to further improve the reliability of voting.

It has been verified that the safety fault detection model of industrial control equipment based on integrated learning

TABLE 1: Validation experiment data set based on the Bagging integrated safety fault diagnosis model.

Number	Experimental data set	Validation data set
1	A, B, C	D
2	A, B, D	C
3	A, C, D	B
4	B, C, D	A

TABLE 2: Sampling situation summary table.

Number of the sample set	Number of class A failure samples	Number of class B failure samples	Number of class C failure samples
1	7982	7982	8036
2	7939	8060	8001
3	8101	7923	7976
4	8053	8186	7761
5	7839	8038	8123
6	7918	7964	8118
7	7989	7881	8130
8	7977	8050	7973
9	8005	8024	7971

TABLE 3: Sampling set self-sampling results.

Number of the sample set	Number of class A failure samples	Number of class B failure samples	Number of class C failure samples
1	7964	8092	7944
2	7954	8079	7967
3	8153	7902	7945
4	8084	8118	7798
5	7862	8001	8137
6	7848	7937	8215
7	7973	7892	8135
8	7976	8059	7965
9	7959	7963	8078

TABLE 4: Experimental results of the integrated model for safety fault diagnosis based on Bagging.

Training data set	Test data set	Average diagnostic results of individual learners (%)	The average diagnosis result of the integrated learner (%)
A, B, C	D	91.23	94.93
A, B, D	C	91.53	95.79
A, C, D	B	91.95	95.86
B, C, D	A	91.48	95.41

that we proposed can efficiently deal with the dynamic factors of the industrial control system.

6. Conclusion

Aiming at the problem of dynamic changes of actual parameters in the environment of intelligent manufacturing and Industrial Internet of Things, this paper focuses on the safety fault diagnosis of industrial control equipment. This paper studies how to self-adaptively complete the safety fault diagnosis through the existing model and innovatively proposes a safety fault diagnosis method based on the integrated learning model. This effectively improves the adaptive ability of the model and makes the fault diagnosis framework truly intelligent. Furthermore, this method is based on the voting weight adjustment method which effectively improves the reliability of voting.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper is partially supported by the National Key R&D Program of China (No. 2020YFB1805503), Jiangsu Province Modern Education Technology Research Project (84365), National Vocational Education Teacher Enterprise Practice Base "Integration of Industry and Education" Special Project

(Study on Evaluation Standard of Artificial Intelligence Vocational Skilled Level).

References

- [1] T. Chen, D. J. Hill, and C. Wang, "Distributed fast fault diagnosis for multimachine power systems via deterministic learning," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 5, pp. 4152–4162, 2019.
- [2] M. Mousavi, M. Moradi, A. Chaibakhsh et al., "Ensemble-based fault detection and isolation of an industrial Gas turbine," in *Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2351–2358, IEEE, Toronto, Canada, October 2020.
- [3] M. Soleimani, J. Faiz, P. S. Nasab, and M. Moallem, "Temperature measuring-based decision-making prognostic approach in electric power transformers winding failures," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6995–7003, 2020.
- [4] R. Kumar and S. Takai, "Inference-based ambiguity management in decentralized decision-making: decentralized diagnosis of discrete-event systems," *IEEE Transactions on Automation Science and Engineering*, vol. 6, no. 3, pp. 479–491, 2009.
- [5] L. Guan, H. Chen, and L. Lin, "A multi-agent-based self-healing framework considering fault tolerance and automatic restoration for distribution networks," *IEEE Access*, vol. 9, pp. 21522–21531, 2021.
- [6] T. G. Dietterich, "Ensemble learning," *The Handbook of Brain Theory and Neural Networks*, vol. 2, pp. 110–125, 2002.
- [7] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [8] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," *icml*, vol. 96, pp. 148–156, 1996.
- [9] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [10] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, San Francisco, CA, USA, August 2016.
- [11] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, pp. 1189–1232, 2001.
- [12] Q. Li and P. Yin Hai, "Time series association state analysis method for attacks on the smart Internet of electric vehicle charging network," *Transportation Research Record*, vol. 2673, no. 4, pp. 217–228, 2019.
- [13] X. Chen, C. Tang, Z. Li, L. Qi, Y. Chen, and S. Chen, "A pricing approach toward incentive mechanisms for participant mobile crowdsensing in edge computing," *Mobile Networks and Applications*, vol. 25, no. 3, pp. 1220–1232, 2020.
- [14] X. Zhou, B. Wu, and Q. Jin, "User role identification based on social behavior and networking analysis for information dissemination," *Future Generation Computer Systems*, vol. 96, pp. 639–648, New York, NY, USA, 2019.
- [15] Q. Li, S. Meng, S. Wang, J. Zhang, and J. Hou, "CAD: Command-level anomaly detection for vehicle-road collaborative charging network," *IEEE Access*, vol. 7, pp. 34910–34924, 2019.
- [16] Q. Li, S. Meng, S. Zhang, J. Hou, and L. Qi, "Complex attack linkage decision-making in edge computing networks," *IEEE Access*, vol. 7, pp. 12058–12072, 2019.
- [17] X. Zhou, B. Wu, and Q. Jin, "Analysis of user network and correlation for community discovery based on topic-aware similarity and behavioral influence," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 6, pp. 559–571, 2018.
- [18] W. Lin, X. Zhang, L. Qi et al., "Location-aware service recommendations with privacy-preservation in the Internet of Things," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 227–235, 2021.
- [19] Q. Li, S. Meng, S. Zhang et al., "Safety risk monitoring of cyber-physical power systems based on ensemble learning algorithm," *IEEE Access*, vol. 7, pp. 24788–24805, 2019.
- [20] K. Tang, W. Tang, E. Luo, Z. Tan, W. Meng, and L. Qi, "Secure information transmissions in wireless-powered cognitive radio networks for internet of medical things," *Security and Communication Networks*, vol. 2020, Article ID 7542726, 10 pages, 2020.
- [21] J. Hou, Q. Li, S. Meng, Z. Ni, Y. Chen, and Y. Liu, "DPRF: a differential privacy protection random forest," *IEEE Access*, vol. 7, pp. 130707–130720, 2019.
- [22] L. Lam and S. Y. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 27, no. 5, pp. 553–568, 1997.
- [23] P. Emerson, "The original Borda count and partial voting," *Social Choice and Welfare*, vol. 40, no. 2, pp. 353–358, 2013.
- [24] Y. S. Huang and C. Y. Suen, "The behavior-knowledge space method for combination of multiple classifier," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 347p. 347, Institute of Electrical Engineers Inc (IEEE), July 1993.
- [25] L. I. Kuncheva, J. C. Bezdek, and R. P. W. Duin, "Decision templates for multiple classifier fusion: an experimental comparison," *Pattern Recognition*, vol. 34, no. 2, pp. 299–314, 2001.
- [26] K. Sentz and S. Ferson, "Combination of evidence in Dempster-Shafer theory," Technical Report, Sandia National Laboratories, Albuquerque, New Mexico, 2002.
- [27] T. Stathaki, *Image Fusion: Algorithms and Applications*, Elsevier, New York, NY, USA, 2011.
- [28] L. Qi, Q. He, F. Chen et al., "Finding all you need: web APIs recommendation in web of Things through keywords search," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 5, pp. 1063–1072, 2019.
- [29] X. Zhou, J. Chen, B. Wu, and Q. Jin, "Discovery of action patterns and user correlations in task-oriented processes for goal-driven learning recommendation," *IEEE Transactions on Learning Technologies*, vol. 7, no. 3, pp. 231–245, 2014.
- [30] S. Zhao, S. Li, L. Qi, and L. Xu, "Computational intelligence-enabled cybersecurity for the Internet of Things," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 5, pp. 666–674, 2020.
- [31] X. Zhou, W. Liang, K. I.-K. Wang, and S. Shimizu, "Multi-modality behavioral influence analysis for personalized recommendations in health social media environment," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 5, pp. 888–897, 2019.
- [32] T. Wang, Z. Cao, S. Wang et al., "Privacy-enhanced data collection based on deep learning for Internet of vehicles," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6663–6672, 2020.

Research Article

Revisiting a Multifactor Authentication Scheme in Industrial IIoT

Ding Wang ^{1,2}, Shuhong Hong ^{1,2} and Qingxuan Wang ^{1,2}

¹College of Cyber Science, Nankai University, Tianjin 300350, China

²Tianjin Key Laboratory of Network and Data Security Technology, Nankai University, Tianjin 300350, China

Correspondence should be addressed to Qingxuan Wang; wangqingxuan@mail.nankai.edu.cn

Received 23 March 2021; Accepted 22 May 2021; Published 4 June 2021

Academic Editor: AnMin Fu

Copyright © 2021 Ding 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.

Nowadays, as one of the key applications of Internet of Things, Industry IIoT (IIoT) has recently received significant attention and has facilitated our life. In IIoT environments, an amount of data generally requires to be transmitted between the user and sensing devices in an open channel. In order to ensure safe transmission of these data, it is necessary for the user and sensing devices to authenticate each other and establish a secure channel between them. Recently, a multifactor authenticated key agreement scheme for IIoT was proposed, which aims to tackle this problem and provide solutions for user multiple sensing devices' access. This work claims that the proposed scheme is secure against various attacks and has less communication and computational costs than other existing related schemes. Unfortunately, we find that this scheme cannot resist smart card attack and sensing device capture attack. Furthermore, we show that this scheme fails to provide forward secrecy, which is essential for a secure multifactor authentication scheme.

1. Introduction

Internet of Things (IoT) has developed rapidly in recent years, which generally penetrates into people's life, and there are many IoT devices applied to various domains [1, 2]. Due to the superiority in automatic monitoring, efficient control, and intelligent manufacturing, Industry IIoT (IIoT) is widely concerned among these domains. In the IIoT environment, sensing devices can be accessed and controlled by users remotely. During the process of production, sensing devices collect the real-time data, and the data can be obtained by users. The network model for IIoT is described in Figure 1. As a security critical system, IIoT has higher requirements in the secure transmission and communication of data [3, 4]. However, it is vulnerable to an attacker to perform attacks because the collected data is often transmitted through a public channel, and this brings security problems in the IIoT environment. It is possible for an adversary to launch attacks and impersonate an authorized user to obtain the data by accessing sensing devices. The unsatisfactory situation mentioned above will lead to destruction of the industrial production.

Therefore, in order to ensure the safe data transmission between users and sensing devices, many authenticated key

agreement schemes [5–7] in IIoT are proposed. In these schemes, users are only allowed to access one sensing device at a time. When the user accesses multiple devices, his identity must be validated repeatedly. While supporting the critical security features, such as shared session key establishment and user authentication, an authentication scheme for IIoT environments should also be able to reduce the communication and computational costs due to the resource-constrained nature of IoT devices.

Recently, Vinoth et al. [8] proposed a multifactor authenticated key agreement scheme for the IIoT environment, aiming to support the authorized user remotely accessing multiple sensing devices. They claimed that their scheme is suitable for the resource-constrained IIoT and has less cost during communication and computation processes. Besides, they demonstrated the security of the proposed scheme through a formal security analysis, which indicates that their scheme is resistant to various known attacks. Unfortunately, some subtleties are overlooked. In this paper, we find that their scheme cannot resist the smart card attack and the sensing device capture attack. Furthermore, we point out that their scheme cannot support forward secrecy. Although the scheme is a multifactor authentication mechanism

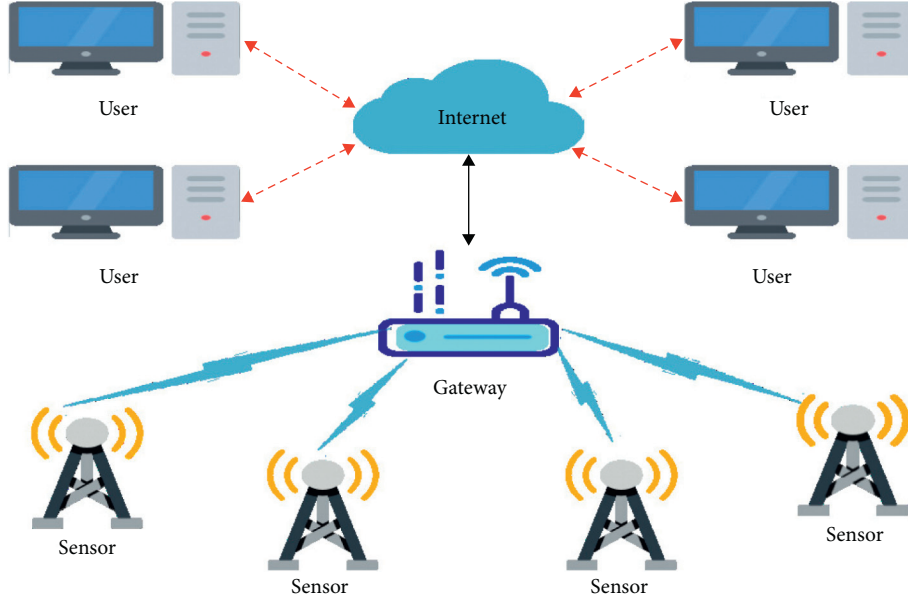


FIGURE 1: Network model for IIoT.

designed by Vinoth et al., we do not think it can provide truly multifactor security.

2. Revisiting Vinoth et al.'s Scheme

In this section, we first revisit Vinoth et al.'s scheme [8] briefly and list some intuitive notations and abbreviates in Table 1 for the convenience of description. Their scheme includes six phases, while we only review the first three phases, which are related to our proposed attacks.

2.1. Offline Sensing Devices' Registration Phase. Each sensing device SD_j is registered by GWN in offline and is distributed a unique identity ID_{SD_j} . In order to calculate the secret, GWN chooses a secret value S and two vectors $Vector_1$ and $Vector_2$. Assume that $S = Vector_1 \cdot x_0$ and $S^2 = Vector_2 \cdot x_0$. GWN then computes $s_j = Vector_1 \cdot x_j$ and $f_j = Vector_2 \cdot x_j$ and picks pair-wise relative positive numbers k_1, \dots, k_n for each sensing device SD_j . GWN computes $Mul = \prod_{j=1}^n k_j$ and $Mul_j = (Mul/k_j)$. Then, GWN generates a random nonce $Nonce_j$, which satisfies $Mul_j \times Nonce_j \equiv 1 \pmod{k_j}$. GWN calculates γ as $\gamma = \sum_{j=1}^n Var_j = \sum_{j=1}^n Mul_j \times Nonce_j$ and stores it. GWN sends $\langle ID_{SD_j}, s_j, f_j, k_j \rangle$ to each sensing device.

2.2. User Registration Phase

- (1) Step URP1: U_i chooses a high-entropy password PW_i and an identity ID_i . U_i imprints the biometrics B_i and uses the generation algorithm to calculate $(BK_i, \tau_i) = Gen(B_i)$. It notes that the algorithm is built into the fuzzy extractor. U_i generates a 128-bit random nonce a and computes TPW_i as $TPW_i = h(ID_i \| PW_i \| BK_i) \oplus a$. Finally, U_i sends a message $\{ID_i, TPW_i\}$ to GWN.

- (2) Step URP2: after receiving the message $\{ID_i, TPW_i\}$, GWN generates a 1024-bit random secret key KEY_{GWN} and further calculates $KEY_{GWN-U_i} = h(ID_i \| KEY_{GWN})$. Then, GWN calculates A_i as $A_i = KEY_{GWN-U_i} \oplus TPW_i$ and C_i as $C_i = ID_{GWN} \oplus TPW_i$. In addition, for each user U_i , GWN generates a 128-bit identity TID_i . Finally, GWN generates the smart card SC_i and sends SC_i to U_i .
- (3) Step URP3: after receiving SC_i , in order to protect A_i , U_i calculates $RPW_i = h(ID_i \| PW_i \| BK_i)$ and $A'_i = A_i \oplus TPW_i \oplus RPW_i$. U_i then computes $D_i = a \oplus h(ID_i \| BK_i)$ and $C'_i = C_i \oplus TPW_i \oplus h(ID_i \| BK_i)$. U_i further calculates $V_i = h(RPW_i \| A_i \| a \| h(ID_i \| BK_i)) \pmod{\omega}$. Finally, U_i needs to store $\{TID_i, A'_i, C'_i, D_i, V_i, Gen(\cdot), \tau_i, \omega, Rep(\cdot), h(\cdot)\}$ into the memory.

2.3. Authenticated Key Agreement Phase. This phase includes the following steps. This phase along with the login phase is summarized in Table 2.

- (1) Step AKAP1: after receiving the message $\{TID_i, M_1, M_2, TS_1\}$, GWN firstly verifies whether $|TS_1 - TS'_1| \leq \Delta TS$ to check the freshness of login request. If it is true, GWN obtains ID_i and KEY_{GWN-U_i} corresponding to TID_i by retrieving the database. GWN calculates $r_i^* = M_1 \oplus KEY_{GWN-U_i}$, $M_3 = h(TID_i \| M_1 \| ID_{GWN} \| r_i^* \| TS_1)$. In order to authenticate the authenticity of U_i , GWN checks whether M_3 is equal to M_2 . If it holds, GWN generates the current timestamp TS_2 and a random nonce r_{GWN} ($r_{GWN} \leq \min\{k_j\}$). GWN calculates M_4 as $M_4 = r_{GWN} \times \gamma$ to securely send r_{GWN} to each sensing device. Then, GWN calculates $M_5 = Enc_{r_{GWN}}(ID_i, ID_{GWN}, r_i^*, r_{GWN} \oplus KEY_{GWN-U_i})$ to encrypt the parameters. After that, GWN computes $M_6 = h(ID_i \| ID_{GWN} \| r_i^* \| M_4 \| TS_2 \pmod{\omega} \| KEY_{GWN-U_i})$

TABLE 1: Notations and abbreviations.

Symbol	Description
GWN	Gateway node
ID_{GWN}	GWN's identity
U_i and SD_j	i^{th} user and j^{th} sensing device
ID_i and ID_{SD_j}	U_i 's and SD_j 's identity, respectively
r_{GWN}	A random nonce
PW_i and B_i	U_i 's password and biometrics
BK_i and τ_i	U_i 's biometrics key and public reproduction parameter
Gen(\cdot) and Rep(\cdot)	Generation and reproduction algorithm of fuzzy extractor, respectively
TS_i	Current timestamp
ΔTS	Maximum transmission delay
KEY_{GWN-U_i}	Symmetric key between U_i and GWN
SK	Session key between the user and sensing devices
S	Secret value utilized for secret sharing
s_j, f_j , and k_j	SD_j 's secret parameters
$h(\cdot)$	Hash function
\oplus and \parallel	Concatenation and bit-wise XOR operation

to help sensing device SD_j authenticate GWN. Finally, GWN broadcasts the message $\{M_4, M_5, M_6, TS_2\}$ via a public channel.

- (2) Step AKAP2: after receiving the broadcast message from GWN, each sensing device SD_j verifies $|TS_2 - TS_2'| \leq \Delta TS$ to check the freshness of the message firstly. If the inequality holds, SD_j uses CRT to obtain $r_{GWN}^* = M_4 \bmod k_j$ by its stored value k_j . SD_j then uses the group key r_{GWN}^* to decrypt M_5 to attain the sensitive parameter ID_i, ID_{GWN}, r_i^* , and $r_{GWN} \oplus KEY_{GWN-U_i}$. SD_j further calculates $M_7 = h(ID_i \parallel ID_{GWN} \parallel r_i^* \parallel M_4 \parallel KEY_{GWN-U_i} \oplus r_{GWN} \oplus r_{GWN}^* \parallel TS_2)$ with the condition $M_7 = M_6$ to verify GWN. If it is true, each SD_j computes $M_8 = \text{Enc}_{r_{GWN}^*}(ID_{SD_j}, s_j f_j)$ to encrypt the legal share s_j and f_j and generates the current timestamp TS_3 . Then, each sensing device SD_j sends the reply message $\{M_8, TS_3\}$ to GWN securely.
- (3) Step AKAP3: when receiving the message, GWN firstly verifies $|TS_3 - TS_3'| \leq \Delta TS$ to check the freshness of the message. If it holds, GWN obtains share s_j and f_j by calculating $\text{Dec}_{r_{GWN}^*}(M_8) = (ID_{SD_j}, s_j, f_j)$. GWN further computes $\theta_1 = \sum_{t=1}^l \lambda_t s_t$ and $\theta_2 = \sum_{t=1}^l \lambda_t f_t$ and checks whether $\theta_1^2 = \theta_2$. If it holds, GWN can reconstruct the secret successfully. Then, GWN computes $M_9 = h(S \parallel r_{GWN})$, $M_{10} = M_9 \times \gamma$, $M_{11} = h(M_9 \parallel M_{10})$, and $M_{12} = \text{Enc}_{KEY_{GWN-U_i}}(r_{GWN}, r_i^*, M_9)$. GWN then generates the current timestamps TS_4 and a new temporal identity TID_i^{new} and calculates $M_{13} = h(ID_i \parallel KEY_{GWN-U_i} \parallel TS_4 \oplus TID_i^{\text{new}})$ and $M_{14} = h(M_{12} \parallel M_9 \parallel r_i^*)$. Finally, GWN broadcasts the message $\{M_{10}, M_{11}\}$ to all the participants and sends the message $\{M_{12}, M_{13}, M_{14}, TS_4\}$ to U_i .
- (4) Step AKAP4: when receiving the message $\{M_{10}, M_{11}\}$ from GWN, each sensing device SD_j

calculates $M_9^* = M_{10} \bmod k_j$ and $M_{15} = h(M_9^* \parallel M_{10})$. If $M_{15} = M_{11}$, each device SD_j calculates SK as $SK = h(ID_i \parallel ID_{GWN} \parallel r_{GWN}^* \parallel r_i^* \parallel M_9^* \parallel KEY_{GWN-U_i})$. GWN validates the shared session key by computing $M_{16} = h(SK \parallel ID_{GWN} \parallel ID_i)$ and sends it to U_i .

- (5) Step AKAP5: after receiving the message, U_i firstly verifies $|TS_4 - TS_4'| \leq \Delta TS$ to check the freshness of the message. If it holds, U_i computes $\text{Dec}_{KEY_{GWN-U_i}}(M_{12}) = (r_{GWN}^*, r_i^*, M_9^*)$. Then, U_i checks whether $r_i^* = r_i$ to validate the session consistency. If it holds, U_i computes $M_{17} = h(M_{12} \parallel M_9^*)$. If $M_{17} = M_{14}$, U_i further calculates SK* as $SK^* = h(ID_i \parallel ID_{GWN} \parallel r_{GWN}^* \parallel r_i^* \parallel M_9^* \parallel KEY_{GWN-U_i})$ with sensing devices. U_i calculates $M_{18} = h(SK^* \parallel ID_{GWN} \parallel ID_i)$ after receiving the message from the sensing device and checks whether $M_{18} = M_{16}$. If it holds, U_i needs to change $TID_i^{\text{new}} = h(ID_i \parallel KEY_{GWN-U_i} \parallel TS_4) \oplus M_{13}$.

3. Cryptanalysis of Vinoth et al.'s Scheme

For a multifactor authentication scheme, it is essential to create a concise and concrete adversarial model. In this section, we propose two attacks, a smart loss attack and a sensing device capture attack to show the vulnerabilities of the scheme. First of all, we refer to the adversary model proposed by Wang et al. [9] which is strict but reasonable. The assumptions below are about the adversary's capabilities:

- (1) There exist two kinds of communication channels: a secure channel and a public channel. The former is mainly used for registration, while the other is mainly used in login and authentication phases. The adversary \mathcal{A} has full control of the public channel, i.e., \mathcal{A} can eavesdrop, intercept, modify, and redirect messages transmitted between communication participants [10, 11].

TABLE 2: Login and authenticated key agreement phase.

	User(U_i)	Gateway(GWN)	Sensing device(SD $_j$)	
L1	<p>Insert SC_i to card reader. Input ID_i, PW_i, and B_i^*.</p> <p>Compute $BK_i^* = \text{Rep}(B_i^*, \tau_i)$, $RPW_i^* = h(ID_i \ PW_i \ BK_i^*)$, $a^* = D_i \oplus h(ID_i \ BK_i^*)$, $A_i^* = A_i \oplus a^*$, and $V_i^* = h(RPW_i^* \ A_i^* \ a^* \ h(ID_i \ BK_i^*)) \bmod \omega$</p> <p>Check whether $V_i = V_i^*$, if so, generate a random nonce r_i and timestamps TS_1. Calculate, $ID_{GWN}^* = C_i \oplus h(ID_i \ BK_i^*)$, $M_1 = A_i \oplus RPW_i^* \oplus r_i$, $M_2 = h(TID_i \ M_1 \ ID_{GWN}^* \ r_i \ TS_1)$ $\rightarrow \text{[open channel]} \langle M_4, M_5, M_6, TS_2 \rangle$</p> <p>Check whether $TS_4 - TS_1 \leq \Delta TS$ if so, compute $(r_{GWN}^*, r_i^*, M_9^*) = \text{Dec}_{\text{KEY}_{GWN-U_i}}(M_{12})$; check whether $r_i^* = r_i$; if so, compute $M_{17} = h(M_{12} \ M_9^* \ r_i)$. Check whether $M_{17} = M_{14}$; if so, compute $M_{18} = h(ID_i \ ID_{GWN}^* \ r_i \ M_9^* \ \text{KEY}_{GWN-U_i})$ and so, the session key is established successfully. Compute $TID_i^{\text{new}} = h(ID_i \ \text{KEY}_{GWN-U_i} \ TS_4) \oplus M_{13}$. Update TID_i with TID_i^{new}</p>	<p>Check whether $TS_1 - TS_3 \leq \Delta TS$. If so, extract ID_j and KEY_{GWN-U_j}. Compute $r_i^* = M_1 \oplus \text{KEY}_{GWN-U_j}$ and $M_3 = h(TID_i \ M_1 \ ID_{GWN}^* \ r_i^* \ TS_1)$. Check whether $M_3 = M_2$; if so, generate a random nonce r_{GWN} and timestamps TS_2. Compute $M_4 = r_{GWN} \times \gamma$, $M_5 = \text{Enc}_{r_{GWN}}(ID_i, ID_{GWN}^*, r_{GWN} \oplus \text{KEY}_{GWN-U_j})$, and $M_6 = h(ID_i \ ID_{GWN}^* \ r_i^* \ M_4 \ \text{KEY}_{GWN-U_j} \ TS_2)$, $\rightarrow \text{[open channel]} \langle M_4, M_5, M_6, TS_2 \rangle$</p> <p>Check whether $TS_3 - TS_3 \leq \Delta TS$; if so, compute $(ID_{SD_j}, s_j, f_j) = \text{Dec}_{r_{GWN}}(M_8)$, $\theta_1 = \sum_{d=1}^l \lambda_d s_d$, and $\theta_2 = \sum_{l=1}^l \lambda_l f_l$. Check whether $\theta_1^2 = \theta_2$; if so, return θ_1 as S. Compute $M_9 = h(S \ r_{GWN})$, $M_{10} = M_9 \times \gamma$, and $M_{11} = h(M_9 \ M_{10})$.</p> <p>Generate a temporal identity TID_i^{new} and timestamp TS_4. Compute $M_{12} = \text{Enc}_{\text{KEY}_{GWN-U_j}}(r_{GWN}, r_i^*, M_9)$, $M_{13} = h(ID_j \ \text{KEY}_{GWN-U_j} \ TS_4) \oplus TID_i^{\text{new}}$, and $M_{14} = h(M_{12} \ M_9 \ r_i^*)$. $\rightarrow \text{[open channel]} \langle M_{10}, M_{11} \rangle \leftarrow \text{[open channel]} \langle M_{12}, M_{13}, M_{14}, TS_4 \rangle$</p>	<p>Check whether $TS_2 - TS_2 \leq \Delta TS$. If so, compute $r_{GWN}^* = M_4 \bmod k_j$, $(ID_i, ID_{GWN}^*, r_i^*, r_{GWN} \oplus \text{KEY}_{GWN-U_j}) = \text{Dec}_{r_{GWN}^*}(M_5)$, and $M_7 = h(ID_i \ ID_{GWN}^* \ r_i^* \ M_4 \ \text{KEY}_{GWN-U_j} \oplus r_{GWN} \oplus r_{GWN}^* \ TS_2)$</p> <p>Check whether $M_7 = M_6$; if so, generate a timestamp TS_3. Compute $M_8 = \text{Enc}_{r_{GWN}^*}(ID_{SD_j}, s_j, f_j)$, $\leftarrow \text{[open channel]} \langle M_8, TS_3 \rangle$</p>	
L2				
V1				
V2				
V3				
V4				
V5				

- (2) The adversary \mathcal{A} can offline exhaust all the items in the Descartes space of identities and passwords which are of low entropy within polynomial time.
- (3) When it comes to multifactor authentication, the scheme should be secure even if one or more factors are compromised, which is called truly multifactor security [12]. Therefore, it is reasonable to make an assumption that \mathcal{A} may (i) obtain a victim's password by performing shoulder surfing or phishing attacks, (ii) extract the secret parameters in the lost smart card by performing side-channel attack, or (iii) attain a victim's biometric information using malicious devices. However, the above assumptions cannot be achieved at the same time; otherwise, it will be a trivial case.
- (4) The adversary \mathcal{A} could be the administrator of the server or a legitimate user in the system.
- (5) The adversary \mathcal{A} can determine victim's identity.

It is worth noting that users can select his/her identity ID and password PW in many protocols. However, the user selected identities and passwords are usually of low entropy ($|\mathcal{D}_{id}| \leq |\mathcal{D}_{pw}| \leq 10^6$) [13, 14]. Therefore, assumption (2) is realistic. Then, assumption (3) specifies truly three-factor security. And, assumption (4) can be used to capture the threats from the system when the server is corrupted or any legitimate users are malicious. Finally, assumption (5) describes the fact that most of the user identity are user's e-mail addresses or phone numbers, which can be easily obtained. The following analysis will take the five assumptions mentioned above into account.

3.1. Smart Card Loss Attack. We employ the user U_i as the victim to show the process of this attack. According to assumption (3), it is reasonable for the adversary \mathcal{A} to get U_i 's smart card SC (stolen or picked up) and corresponding biometrics B_i^* . Besides, as a premeditated adversary, \mathcal{A} has full control of the public channel, and she can collect a past transcript between U_i and gateway node (GWN) (i.e., $\{TID_i, M_1, M_2, TS_1\}$). Then, \mathcal{A} can guess U_i 's password and identity correctly as following steps:

Step 1. \mathcal{A} computes $BK_i = \text{Rep}(B_i^*, \tau_i)$, where τ_i can be extracted from victim's smart card

Step 2. \mathcal{A} chooses a pair (ID_i^*, PW_i^*) from $\mathcal{D}_{id} \times \mathcal{D}_{pw}$, where \mathcal{D}_{id} denotes the identity space and \mathcal{D}_{pw} denotes the password space

Step 3. \mathcal{A} computes $ID_{GWN}^* = C'_i \oplus h(ID_i^* \| BK_i)$

Step 4. \mathcal{A} computes $RPW_i^* = h(ID_i^* \| PW_i^* \| BK_i)$

Step 5. \mathcal{A} computes $r_i^* = M_1 \oplus A'_i \oplus RPW_i^*$, noted that \mathcal{A} can extract A'_i from victim's smart card and collect M_1 from the past transcript

Step 6. \mathcal{A} computes $M_2^* = h(TID_i \| M_i \| ID_{GWN}^* \| r_i^* \| TS_1)$ and verifies the correctness of (ID_i^*, PW_i^*) pair by checking if $M_2^* = M_2$

Step 7. \mathcal{A} executes the steps 2 ~ 6 repeatedly until finding the correct values

As mentioned before, users can choose his/her own ID and PW in most password-based authentication schemes (e.g., References [15–17]) aiming to achieve user-friendliness. And, Vinoth et al.'s scheme is no exception. It makes assumption (2) reasonable that users often select low entropy identities and passwords. Therefore, it is possible for \mathcal{A} to exhaust all the (ID, PW) pairs offline within polynomial time. We can calculate the running time of the attack procedure as $\mathcal{O}(3T_H \times |\mathcal{D}_{id}| \times |\mathcal{D}_{pw}|)$, where $|\mathcal{D}_{id}|$ represents the number of identities, $|\mathcal{D}_{pw}|$ represents the number of passwords, and T_H represents the running time for Hash operation. Note that the operation time of bit-wise XOR operation in Step 3 can be ignored. Since $|\mathcal{D}_{id}|$ and $|\mathcal{D}_{pw}|$ are very limited (e.g., $|\mathcal{D}_{id}| \leq |\mathcal{D}_{pw}| \leq 10^6$) [13, 14], the attack mentioned above is significant and shows a challenge to user authentication protocols.

3.2. Sensing Device Capture Attack. According to Vinoth et al.'s threat model, the adversary \mathcal{A} can compromise a sensing device (SD) and extract the parameters stored in it (i.e., $\{ID_{SD_j}, s_i, f_i, k_j, a\}$). We assume that SD_j is captured by the adversary; then, \mathcal{A} can successfully impersonate the user U_i as follows:

Step 1. Computes $r_{GWN}^* = M_4 \bmod k_j$, where M_4 is received from (GWN)

Step 2. Decrypts the received message M_5 by using the key r_{GWN}^* and obtains the security parameters ($r_{GWN}^* \oplus \text{KEY}_{GWN-U_i}$, ID_{GWN} , ID_i , and r_i^*) of the user U_i who is sending the login request

Step 3. Computes KEY_{GWN-U_i} as $\text{KEY}_{GWN-U_i} = r_{GWN}^* \oplus r_{GWN}^* \oplus \text{KEY}_{GWN-U_i}$

Step 4. Computes $TID_i^{\text{new}} = h(ID_i \| \text{KEY}_{GWN-U_i} \| TS_4) \oplus M_{13}$, where TS_4 and M_{13} are obtained from the public channel

Step 5. Randomly chooses a new nonce r_i^{new} and current timestamp TS_1^{new}

Step 6. Computes M_1^* as $M_1^* = \text{KEY}_{GWN-U_i} \oplus r_i^{\text{new}}$ and M_2^* as $M_2^* = h(TID_i^{\text{new}} \| M_1^* \| ID_{GWN} \| r_i^{\text{new}} \| TS_1^{\text{new}})$

Step 7. Sends the login request $\{TID_i^{\text{new}}, M_1^*, M_2^*, TS_1^{\text{new}}\}$ to (GWN) and finishes the login phase

After receiving the message, GWN first checks the freshness of the received message and computes $r_i^{\text{new}} = \text{KEY}_{GWN-U_i} \oplus M_1^*$, where KEY_{GWN-U_i} is stored in the GWN's database and retrieved according to corresponding TID_i . Then, GWN computes $M_3 = h(TID_i^{\text{new}} \| M_1^* \| ID_{GWN} \| r_i^{\text{new}} \| TS_1^{\text{new}})$ and verifies whether the calculated M_3 is equal to the received M_2 . If it holds, GWN will authenticate the authenticity of U_i . Since the parameters are calculated correctly, the adversary \mathcal{A} can pass the verification of the GWN. So far, the adversary has successfully impersonated user U_i .

3.3. No Forward Secrecy. When a scheme ensures that, even the long-term private keys (or secret) of communication participants are leaked, previously agreed session keys can

still be secure [18], then the scheme is called supporting forward secrecy. It is important for security critical systems to support forward secrecy, especially when there still exist many security and privacy problems in the IIoT environment.

If an attacker \mathcal{A} has captured a sensing device SD_j , extracted the parameters $\{ID_{SD_j}, s_i, f_i, k_j\}$ from SD_j , and intercepted the messages $\{M_4, M_5, M_{10}\}$, the following method can be used to calculate the session key:

Step 1. \mathcal{A} computes $r_{GWN}^* = M_4 \bmod k_j$, where M_4 is received from (GWN)

Step 2. \mathcal{A} decrypts the received message M_5 by using the key r_{GWN}^* and obtains the security parameters ($r_{GWN} \oplus KEY_{GWN-U_i}$, ID_i, r_i^* , and ID_{GWN}) of the user U_i

Step 3. \mathcal{A} computes $M_9^* = M_{10} \bmod k_j$, where M_{10} is received from (GWN)

Step 4. \mathcal{A} computes the session key $SK^* = h(ID_i \| ID_{GWN} \| r_{GWN}^* \| r_i^* \| M_9^* \| KEY_{GWN-U_i})$

With the session key SK^* computed, the entire session will be no secret to the adversary \mathcal{A} .

3.4. Security Vulnerability Discussion. In this section, we highlight again that when considering multifactor security, even if one or more authentication factors are obtained (not all) by the adversary, the scheme should not be broken. Based on this assumption, we proposed the smart card loss attack and the sensing device capture attack. Although Vinoth et al. have employed the fuzzy-verify technique proposed by Wang and Wang [12], the adversary can still obtain victim's password in the way of offline guessing. This disappointing situation is caused that they do not employ the public-key cryptosystem and no public key material is used to construct the login message. To solve this problem, we suggest to use Diffie–Hellman key exchange scheme. Specifically, GWN computes $y = g^x \bmod p$ and stores it into U_i 's smart card SC during the user registration phase, where g is a generator of the group G , p is a large prime number, and x is GWN's secret key. After that, when U_i logs in, she should choose a random number u and compute $Y_1 = g^u \bmod p$ and $C_1 = y^u \bmod p$ first; then, she constructs the login message $M_2 = h(TID_i \| M_1 \| ID_{GWN}^* \| r_i \| C_1 \| Y_1 \| TS_1)$. Since the adversary cannot calculate Y_1 , the aforementioned smart card loss attack can be prevented.

Meanwhile, in the sensing device capture attack, an adversary \mathcal{A} can impersonate user U_i even without her password PW_i . Essentially, when authenticating the identity of user U_i , the GWN only checks whether the user U_i who sends the login request holds the parameter KEY_{GWN-U_i} . Unfortunately, KEY_{GWN-U_i} is encrypted by the group key r_{GWN}^* in the message M_5 , but obtaining the group key r_{GWN}^* is easy for an adversary who has breached the SD_j and extracted k_j . After this, \mathcal{A} could decrypt M_5 to get $KEY_{GWN-U_i}, ID_i, r_i^*$, and ID_{GWN} . With these parameters, \mathcal{A} can bypass the system's user authentication. One possible countermeasure to this problem is that GWN constructs the message $M_5 = Enc_{r_{GWN}}(ID_i, ID_{GWN}, KEY_{GWN-U_i} \oplus r_i^*)$ and

$M_6 = h(ID_i \| ID_{GWN} \| M_4 \| KEY_{GWN-U_i} \| TS_2)$. As a result, when receiving the message, $\{M_4, M_5, M_6, TS_2\}$, the adversary \mathcal{A} who captures the sensing device can only obtain the parameter $KEY_{GWN-U_i} \oplus r_i^*$ by decrypting M_5 . Therefore, \mathcal{A} cannot impersonate U_i since she cannot obtain KEY_{GWN-U_i} .

Note that, one may argue that when the victim user U_i interacts with the GWN, she uses her temporary identity TID_i , and it seems impressible for an adversary to find victim's message from the transcript. However, according to assumption (5), the adversary \mathcal{A} can determine victim's identity. Thus, a premeditated adversary may first compromise a sensing device and wait for the victim chosen by her to send the login request message. Then, \mathcal{A} decrypts M_5 to get ID_i and checks if this ID belongs to the victim. After that, \mathcal{A} continues to monitor the channel until U_i 's session ends. Finally, \mathcal{A} could calculate TID_i^{new} as Step 4 of the sensing device capture attack.

In order to fix the defects of forward secrecy, we also rely on public key cryptography. Specifically, before computing M_8 , SD_j first chooses a random number r_s and computes $Y_2 = g^{r_s} \bmod p$. Then, SD_j computes $M_8 = Enc_{r_{GWN}}(ID_{SD_j}, S_j, f_j, Y_2)$ and sends it to GWN. After that, GWN chooses a random number x' and computes $C_2 = Y_2 Z^{x'} \bmod p$, $Y_{GWN-S_j} = g^{x'} \bmod p$, $M_{10} = Enc_{C_2}(M_9)$, and $M_{11} = h(M_{10} \| M_9 \| Y_{GWN-S_j})$. In Section 3.4, we show that the adversary \mathcal{A} can compute the session key SK , and this is caused by \mathcal{A} to obtain M_9 . However, M_9 is protected by the shared key C_2 now. As a result, \mathcal{A} cannot calculate the previous session key. In order to be consistent with the previous modification, the session key is calculated as $SK = h(ID_i \| ID_{GWN} \| r_{GWN}^* \| M_9^* \| KEY_{GWN-U_i} \oplus r_i^*)$.

4. Conclusion

In this paper, we have revisited and analysed Vinoth et al.'s authentication scheme for IIoT environments. We demonstrate that their scheme suffers from the smart card loss attack and the sensing device capture attack although they claimed that their scheme has the ability to defend various known attacks. We have also briefly discussed the potential causes of these defects. It is hoped that the proposed attacks can help inspire new designs of secure and efficient multifactor authentication protocols for IIoT.

Data Availability

Data sharing is not applicable to this article as no new data was created or analysed in this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Z. Meng, Z. Wu, C. Muvianto, and J. Gray, "A data-oriented M2M messaging mechanism for industrial iot applications,"

- IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 236–246, 2017.
- [2] 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, 2018.
- [3] W. Z. Khan, M. H. U. Rehman, H. M. Zangoti, M. K. Afzal, N. Armi, and K. Salah, “Industrial internet of things: recent advances, enabling technologies and open challenges,” *Computers & Electrical Engineering*, vol. 81, Article ID 106522, 2020.
- [4] W. Ali, I. Ud Din, A. Almogren, M. Guizani, and M. Zuair, “A lightweight privacy-aware IoT-based metering scheme for smart industrial ecosystems,” *IEEE Transactions on Industrial Informatics*, vol. 9, p. 1, 2020.
- [5] J. Seto, Y. Wang, and X. Lin, “User-habit-oriented authentication model: toward secure, user-friendly authentication for mobile devices,” *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 1, pp. 107–118, 2015.
- [6] C. Wang, L. Xiao, J. Shen, and R. Huang, “Neighborhood trustworthiness-based vehicle-to-vehicle authentication scheme for vehicular ad hoc networks,” *Concurrency and Computation: Practice & Experience*, vol. 31, no. 21, Article ID e4643, 2019.
- [7] J. Shen, T. Zhou, X. Liu, and Y.-C. Chang, “A novel Latin-square-based secret sharing for M2M communications,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3659–3668, 2018.
- [8] R. Vinoth, L. J. Deborah, P. Vijayakumar, and N. Kumar, “Secure multi-factor authenticated key agreement scheme for industrial IoT,” *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3801–3811, 2020.
- [9] C. Wang, D. Wang, Y. Tu, G. Xu, and H. Wang, “Understanding node capture attacks in user authentication schemes for wireless sensor networks,” *IEEE Transactions on Dependable and Secure Computing*, p. 1, 2020.
- [10] X. Li, M. H. Ibrahim, S. Kumari, A. K. Sangaiah, V. Gupta, and K.-K. R. Choo, “Anonymous mutual authentication and key agreement scheme for wearable sensors in wireless body area networks,” *Computer Networks*, vol. 129, pp. 429–443, 2017.
- [11] D. He, S. Zeadally, L. Wu, and H. Wang, “Analysis of handover authentication protocols for mobile wireless networks using identity-based public key cryptography,” *Computer Networks*, vol. 128, pp. 154–163, 2017.
- [12] D. Wang and P. Wang, “Two birds with one stone: two-factor authentication with security beyond conventional bound,” *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 708–722, 2018.
- [13] D. Wang, Z. Zhang, P. Wang, J. Yan, and X. Huang, “Targeted online password guessing: an underestimated threat,” in *CCS’16: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1242–1254, Vienna, Austria, October 2016.
- [14] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, “Zipf’s law in passwords,” *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.
- [15] R. Amin, N. Kumar, G. P. Biswas, R. Iqbal, and V. Chang, “A light weight authentication protocol for iot-enabled devices in distributed cloud computing environment,” *Future Generation Computer Systems*, vol. 78, pp. 1005–1019, 2018.
- [16] F. Wu, X. Li, A. K. Sangaiah et al., “A lightweight and robust two-factor authentication scheme for personalized healthcare systems using wireless medical sensor networks,” *Future Generation Computer Systems*, vol. 82, pp. 727–737, 2018.
- [17] Q. Jiang, J. Ma, F. Wei, Y. Tian, J. Shen, and Y. Yang, “An untraceable temporal-credential-based two-factor authentication scheme using ECC for wireless sensor networks,” *Journal of Network and Computer Applications*, vol. 76, pp. 37–48, 2016.
- [18] C. Wang and G. Xu, “Cryptanalysis of three password-based remote user authentication schemes with non-tamper-resistant smart card,” *Security and Communication Networks*, vol. 2017, Article ID 1619741, 14 pages, 2017.