Wireless Communications and Mobile Computing

Rethinking Authentication on Smart Mobile Devices

Lead Guest Editor: Ding Wang
Guest Editors: Jian Shen, Joseph Liu, and Raymond Kim-Kwang Choo

Rethinking Authentication on Smart Mobile Devices

Wireless Communications and Mobile Computing

Rethinking Authentication on Smart Mobile Devices

Lead Guest Editor: Ding Wang
Guest Editors: Jian Shen, Joseph Liu,
and Raymond Kim-Kwang Choo

Editorial Board

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 7079037, 4 pages
https://doi.org/10.1155/2018/7079037

Editorial

Rethinking Authentication on Smart Mobile Devices

Ding Wang,[1] **Jian Shen**,[2] **Joseph K. Liu**,[3] **and Kim-Kwang Raymond Choo**[4]

[1]*School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China*
[2]*School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, China*
[3]*Faculty of Information Technology, Monash University, Melbourne, Australia*
[4]*Department of Information Systems and Cyber Security, The University of Texas at San Antonio, San Antonio, TX 78249, USA*

Correspondence should be addressed to Ding Wang; wangdingg@pku.edu.cn

Received 14 November 2018; Accepted 14 November 2018; Published 18 December 2018

1. Introduction

Rapid advances in wireless technologies (e.g., LTE, LTE-A, WiMAX, 3G, Bluetooth, ZigBee, Z-Wave, and LoRaWAN) have partly contributed to the proliferation of smart mobile devices (e.g., sensors), unmanned vehicles, wearable and embedded devices, and so on. The amount and nature of communications and transactions on such devices and the underpinning systems require a secure and effective authentication mechanism to prevent unauthorized access from illegitimate entities (including both devices and users).

Authentication has been widely deployed to prevent unauthorized access and, in many cases, is also the primary line of defense. A large number of authentication mechanisms and schemes exist for conventional systems, but they may not be suitable for the smart mobile computing paradigm. Firstly, smart mobile devices generally have limited computation and storage and energy capabilities (in comparison to servers, personal computers and laptops), and thus deploying authentication schemes that employ expensive cryptographic primitives will not be viable. Secondly, smart mobile devices are typically small devices with a small screen, keyboard, and so forth, and thus existing authentication schemes may not be sufficiently user-friendly. Thirdly, smart mobile devices often deal with sensitive applications, activities, and data (e.g., location, preferences, and physical condition), and thus privacy demands are much more stringent than traditional authentication schemes. Consequently, it is necessary to perform a critical rethink the way we perform authentication for smart mobile devices and promote new methods that are both robust and easy to use, in order to minimize impact on the user's primary task.

This special issue aims to provide a forum for researchers to publish and exchange their recent research ideas and results about authentication on smart mobile devices. The following 20 papers were selected for inclusion in this special issue after several rounds of reviews by experts in the respective domains. The topics covered in the accepted papers range from attacks against fingerprint sensor hardware, biometric template protection, privacy-preserving message authentication, new lightweight cryptographic primitives (e.g., random number generation and oblivious transfer from lattice-based cryptography) for authentication, to various authentication schemes (one-factor, two-factor, and three-factor) designed for varied specific environments (such as Cloud, RFID, Vehicular, and Wireless Sensor Networks).

2. In This Special Issue

The paper entitled "Attack Potential Evaluation in Desktop and Smartphone Fingerprint Sensors: Can They Be Attacked by Anyone?" by I. Goicoechea-Telleria et al. explains how one can hack into a fingerprint sensor using fake fingers made of Play-Doh and other easy-to-obtain materials. The authors also perform three evaluations in order to demonstrate the attacks on desktop fingerprint sensors and smartphones with embedded sensors, using 15 simulated attackers with no prior background in biometrics. The authors also analyze the attack potential of each of the presented case, based on ISO/IEC 30107-3.

The paper entitled "Biometrics Based Privacy-Preserving Authentication and Mobile Template Protection" by W. Yang et al. presents a new cancelable fingerprint template, which not only mitigates the negative effect of nonlinear distortion by combining multiple feature sets, but also defeats ARM attacks through a proposed feature decorrelation algorithm. Experimental results on public databases and security analysis show the validity of the proposed cancelable template.

The paper entitled "Muscle Activity-Driven Green-Oriented Random Number Generation Mechanism to Secure WBSN Wearable Device Communications" by Y. Cao et al. presents a muscle activity-driven green-oriented random number generation mechanism for wireless body sensor network (WBSN). Specifically, the mechanism uses the human muscle-activity as the green energy resource to generate random numbers (RNs). In comparison to other methods, their scheme could generate random numbers with comparable performance but at a higher speed (128 bits per second).

The paper entitled "Oblivious Transfer via Lossy Encryption from Lattice-Based Cryptography" by Z. Li et al. presents an efficient authentication protocol by improving the performance of the UC-secure OT protocol. The authors first design a multibit lossy encryption under the decisional learning with errors (LWE) assumption and then develop a new variant of UC-secure OT protocol for authenticated protocol via lossy encryption scheme. Additionally, the proposed OT protocol is shown to be secure against semihonest (static) adversaries in the common reference string (CRS) model within the UC framework.

The paper entitled "Efficient Message Authentication Scheme with Conditional Privacy-Preserving and Signature Aggregation for Vehicular Cloud Network" by Y. Xie et al. gives an efficient message authentication in a Vehicular cloud network (VCN) setting. The scheme is shown to be secure and achieves conditional privacy-preserving. Compared with other similar conditional privacy-preserving authentication schemes, the proposed scheme has better performance for both computation and communication. Simulation analysis further demonstrates that the new scheme has reduced verification loss rate and message delay.

The paper entitled "Secure Certificateless Authentication and Road Message Dissemination Protocol in VANETs" by H. Tan et al. presents a certificateless authentication and road message dissemination protocol for a VANET environment. In their scheme, the certificateless signature and the relevance feedback mechanism are adapted for authentication and group key distribution. Subsequently, a message evaluating and ranking strategy is introduced. The security analysis shows that the proposed protocol achieves the desirable security properties.

The paper entitled "An Anonymous Authentication Protocol Based on Cloud for Telemedical Systems" by W. Li et al. introduces an anonymous authentication protocol for cloud-based telemedical systems. Compared with similar related works, the proposed scheme allows patients to remotely access medical services with privacy and achieves better efficiency. A formal security proof is also presented and the performance evaluation suggests better efficiency.

The paper entitled "A Provably Secure Anonymous Authenticated Key Exchange Protocol Based on ECC for Wireless Sensor Networks" by K. Zhang et al. presents an anonymous authenticated key exchange protocol, based on Elliptic Curves Cryptography (ECC). This protocol provides strong user anonymity, such that even the gateway node and the sensor nodes do not know the real identity of the user. The security of the proposed protocol is proven in a well-defined security model under the CDH assumption. Compared with other related protocols, their protocol is efficient in terms of communication and enjoys stronger security.

The paper entitled "An Enhanced User Authentication Protocol Based on Elliptic Curve Cryptosystem in Cloud Computing Environment" by C. Wang et al. revealed security weaknesses in Amin et al's protocol. Then, the authors design a secure authentication protocol and use BAN logic and heuristic analysis method to prove the security of the proposed protocol.

In the paper "Cryptanalysis and Security Enhancement of Three Authentication Schemes in Wireless Sensor Networks", W. Li et al. cryptanalyze and enhance three password-based user authentication schemes designed for WSNs (i.e., several security vulnerabilities loopholes are in the first protocol, the second protocol is not able to achieve the claimed security goal of forward secrecy and is vulnerable to user anonymity violation and offline password guessing attacks, and the third anonymous scheme does not provide forward secrecy and user-friendliness). In addition, by adopting the "perfect forward secrecy (PFS)" principle proposed by Ding Wang et al. (IJCS, 2014), the authors provide several effective countermeasures to mitigate the identified weaknesses. To test the necessity and effectiveness of their countermeasures, the authors conduct a comparison of 10 representative schemes in terms of the underlying cryptographic primitives used for realizing forward secrecy.

The paper entitled "Trusted Authority Assisted Three-Factor Authentication and Key Agreement Protocol for the Implantable Medical System" by D. Mao et al. proposed an improved AKA scheme which achieves strong security features including user anonymity and known key security. It is provably secure under the Real-Or-Random model. Moreover, a comprehensive heuristic security analysis shows that their scheme can resist various attacks and satisfy the desired requirements. Finally, the performance analysis shows that the superiority of their protocol is suitable for the implantable medical system.

The paper entitled "A Secure Three-Factor Multiserver Authentication Protocol against the Honest-But-Curious Servers" by H. Gao et al. took Chengqi Wang et al.'s protocol (PLoS ONE, 2016) as an example, to exhibit how an honest-but-curious server can attack their protocol. To remedy this weakness, a novel three-factor multiserver authentication protocol is presented. By introducing the registration centre into the authentication process, the new protocol can resist the passive attack from the honest-but-curious servers. Security analysis is given to demonstrate the correctness and validity of the new protocol. Compared with related protocols, the proposed protocol possesses more security features

and practical functionalities than others at a relatively low computation cost and communication cost.

The paper entitled "Efficient Multifactor Two-Server Authenticated Scheme under Mobile Cloud Computing" by Z. Han et al. proposed a Kerberos-like authentication method using two servers (adding another authentication server in addition to the Web server) and considering multiple factors to avoid the leakage of users' private data stored on the server side. The proposed scheme aims at the security issue of fingerprint information in the mobile payment environment. The main idea of the proposed solution is to separate certain security-related functions from a Web server to an independent server.

The paper entitled "The Research of Mobile Location Privacy Protection Access Control Method Based on Game Theory" by L. Zheng et al. develops a mobile location privacy access control method based on game theory aiming at the leakage of private information in the mobile location of the Internet of Things users. It controls access behaviour of the privacy information according to the specified location access policy from the perspective of the service provider. The access control can guarantee the server to make a dynamic response to illegal access behaviour to the private information and at the same time, according to tolerance setting, avoid the indirect leakage of mobile location and privacy information caused by the superposition of information.

The paper "A Data Leakage Prevention Method Based on the Reduction of Confidential and Context Terms for Smart Mobile Devices" by X. Yu et al. suggests a model of Data Leakage Protection Based on Context (or CBDLP), a data leakage prevention model based on confidential terms and their context terms, which can detect the rephrased confidential contents effectively. In CBDLP, a graph structure with confidential terms and their context involved is adopted to represent documents of the same class, and then the confidentiality score of the document to be detected is calculated to justify whether confidential contents is involved or not. Based on the attribute reduction method from rough set theory, the authors further present a pruning method. According to the importance of the confidential terms and their context, the graph structure of each cluster is updated after pruning.

The paper "LIP-PA: A Logistics Information Privacy Protection Scheme with Position and Attribute-Based Access Control on Mobile Devices" by Q. Gao et al. constructs a logistics information privacy protection scheme with position and attribute-based access control on mobile devices. First, in order to realize fine-grained access control of encrypted logistics information, the authors adopt ciphertext-policy attribute-based encryption (CP-ABE) scheme, which encrypts segmented logistics information in different access policies. Different couriers can only decrypt different segments of the express order in accordance with their respective attributes. Second, the authors apply position-based key exchange, which uses the courier's physical position information as the credential, to realize position-based access control on couriers. Third, the authors utilize public key encryption to achieve the confidentiality of personal information. Meanwhile, the authors use the digital signature to ensure the verifiability of the parcel and the undeniability of customers.

The paper "Multidevice Authentication with Strong Privacy Protection" by J. Hajny focuses on the card-based physical access control systems and proposed a novel cryptographic scheme based on efficient zero-knowledge proofs and Boneh-Boyen signatures. The proposed scheme is provably secure and provides the full set of privacy-enhancing features that is the anonymity, untraceability, and unlinkability of users. Furthermore, the proposed scheme supports distributed multidevice authentication with multiple RFID (Radio-Frequency IDentification) user devices. This feature is particularly important in applications for controlling access to dangerous sites where the presence of protective equipment is checked during each access control session.

The paper "Identity-Based Fast Authentication Scheme for Smart Mobile Devices in Body Area Networks" by C. Wang et al. proposes an identity-based fast authentication scheme for smart mobile devices in wireless body area networks (WBANs). The scheme can shorten the time of device authentication in an emergency to achieve fast authentication. The analysis of the scheme shows the security and efficiency of the proposed scheme.

The paper "An SDN-Based Connectivity Control System for Wi-Fi Devices" by T. Nguyen-Duc and T. Kim introduces a remote connectivity control system for Wi-Fi devices based on software-defined networking (SDN) in a wireless environment. The main contributions of the proposed system are twofold: (i) it enables network owner/administrator to manage as well as approve connection request from Wi-Fi devices through remote services, which is essential for easy connection management across diverse IoT devices; it also allows fine-grained access control at the device level through remote control. They describe the architecture of SDN-based remote connectivity control of Wi-Fi devices.

The paper "Lightweight Cryptographic Techniques for Automotive Cybersecurity" by A. K. Jadoon et al. presents a survey about developments in vehicular networks from the perspective of lightweight cryptographic protocols and privacy preserving algorithms. Meanwhile, the authors increase awareness about the possible threats to the future automotive industry and give an interesting overview of lightweight cryptographic solutions to these threats. In all, the paper deals with a very interesting and up to date subject: Cryptographic Techniques for Automotive Cyber Security.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

In closing, the guest editors would like to thank all authors who have submitted their papers to this special issue. We would also like to appreciate all the precious time and efforts that the reviewers devoted to the review process of these submissions. The launch of this special issue was in part

supported by the National Nature Science Foundation of China under Grant no. 61802006, no. U1836115, and no. 61672295. It is our hope that this special issue will advance the understanding and research of User Authentication on Smart Mobile Devices. We hope you enjoy the papers.

Ding Wang
Jian Shen
Joseph K. Liu
Kim-Kwang Raymond Choo

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 1847890, 9 pages
https://doi.org/10.1155/2018/1847890

Research Article

The Research of Mobile Location Privacy Protection Access Control Method Based on Game Theory

Lijuan Zheng[1], **Linhao Zhang**[1], **Meng Cui**[1], **Ning Cao**[2], **Jianrui Ding**[3], **Leul Yalemshet**[1], **Tsepo Nyakonda**[1] and **Shepard Musasike**[1]

[1]*School of Information Science and Technology, ShiJiazhuang Tiedao University, ShiJiazhuang 050043, China*
[2]*College of Information Engineering Qingdao Binhai University, Qingdao 266555, China*
[3]*Harbin Institute of Technology, School of Computer Science and Technology, Harbin 150001, China*

Correspondence should be addressed to Lijuan Zheng; zhenglijuan@stdu.edu.cn

Received 28 April 2018; Accepted 1 November 2018; Published 12 December 2018

Guest Editor: Jian Shen

In recent years, the Internet of things has developed rapidly. And the location-based service (LBS) is becoming more and more extensive. Service providers hold a large number of users' information. In order to improve the quality of service, service providers increasingly use big data technology to provide more accurate services for users. At the same time, it aggravates the information disclosure of users' privacy. From the perspective of the service provider, a mobile location privacy access control method based on game theory is proposed to solve the access control problem of mobile location privacy information. Firstly, the weight coefficient is set according to the location privacy influence factors, and then the access control threshold is calculated according to the privacy location leakage situation of the mobile location. Different visitor levels are set according to the threshold. In the process of access control, the prejudgement of the access behaviour is performed, and then the privacy information amount of the information requested for access is calculated according to the weights of the different information. Compare the result with the threshold and get the access control strategy. The strategy set is selected based on strategy matrix of game theory and thresholds are adjusted based on the calculation returns of strategy matrix. The effectiveness and practicability of the method are verified through the security analysis.

1. Introduction

In recent years, the Internet of things (IoT) has developed rapidly, and applications based on the IoT have gradually emerged [1]. It involves military, industry, agriculture, power grid and water networks, transportation, logistics, energy conservation, environmental protection, medical care and healthcare, smart home appliances, and so on [2]. In the IoT, applications based on location services are also increasingly widespread. Location-based services provide location services to the user by determining the geographical location of the mobile device, which can easily lead to leakage of user location privacy. For example, a single point location will be changed in real time when a user requests a location service. An unauthorized visitor accesses an application server to obtain user location information. After obtaining a certain amount of information, the location privacy information can infer the user's track information, resulting in leakage of user location privacy information which will also lead to attacks against users [3]. The continuous development of big data technology also brings greater development space for enterprises [4]. More and more location-based service providers can make data mining based on existing user data to improve user experience. JP Tang et al. studied the platform through intelligent traffic data sensing information processing platform and used the existing traffic big data to improve the intelligent traffic service [5]. The leakage of private information greatly affects the user's service experience. Solving the problem of privacy information protection in the application process of Internet of things is one of the necessary conditions for the wide application of the Internet of things [6].

The user's mobile location privacy information is divided into single point location privacy information and trajectory privacy information. Single point location privacy refers to an independent location point related to private information. Different from single point location privacy, the track mainly refers to the order of the position information of a moving object in chronological order. Trajectory privacy focuses on the temporality of location information. The mobile location privacy contains single point location privacy and trajectory privacy, which refers to the temporality, spatiality, and frequency characteristics of location information.

This paper proposes a mobile location privacy access control method based on game theory aiming at the leakage of privacy information in the mobile location of Internet of things users. It controls the privacy information access behaviour according to the specified location privacy information access policy from the perspective of the service provider. The access control can guarantee the server to make a dynamic response to illegal access behaviour about privacy information. At the same time, according to tolerance setting, avoid the indirect leakage of mobile location and privacy information caused by the superposition of information.

2. Research Status

IoT location privacy protection has received more and more attention in recent years, and the related research work has achieved a lot of results. It is mainly aimed at the query and release of location information. Related location privacy protection technologies include location anonymous technologies, among which application research of k-anonymous technology proposed by Samarati and Sweeney is the most widely used [7]: encryption technologies, such as secure multiparty computing technology, homomorphic encryption technology, fake convergence node protocol, and RFID privacy protection technology [8]. IoT location privacy protection technology has gradually matured, but protection technology related to mobile location privacy is still in its infancy. The existing methods are mostly based on the traditional anonymization, fuzzy, and cryptography. Wang YJ et al. proposed a real incentive mechanism based on location privacy protection in mobile crowdsourcing systems under the online network environment. The improved two-stage auction algorithm based on trust degree and privacy sensibility is proposed. The differential privacy based on Gauss white noise is applied to k-anonymity to prevent user location information leakage [9]. Zhang HT. et al. proposed novel online spatial temporal k-anonymity method for location privacy protection from sequence rules-based inference attacks. It predicts the destination location from the privacy-sensitive sequence rules excavated from large scale anonymous datasets. This method extends the original sequence database of anonymous datasets. It adopts specific generalization and avoidance principles to gradually hide privacy-sensitive rules and resist reasoning attacks [10]. There is also someone in conjunction with other technologies for privacy protection research. Huguenin K. et al. use machine learning methods to predict the user's motivation to check-in and quantify utility implications to protect location privacy [11]. Based on mobile cloud computing, Gong Y. et al. proposed a framework to protect location privacy when assigning tasks to mobile devices, allowing mobile devices to contribute resources to the ad hoc mobile cloud without revealing location information [12]. Mobile location privacy involves trajectory privacy information. There are three main aspects of protection technology for trajectory privacy: fuzzification, release suppression, and addition of fake data. Zhao P. et al. proposed ILLIA to solve location injection attacks. On the basis of exploring the similarity of user mobility in continuous LBS queries, a trust-based k-anonymity scheme is developed. ILLIA can defend against LIA without knowing how to manipulate false locations in advance, while still maintaining high-quality services [13]. Suppression method is often used in trajectory privacy protection. Al-Hussaenik et al., based on the (K, C) L standard, used the local and global suppression methods to avoid attackers obtaining track privacy information [14]. The fake data method is also widely used in track privacy protection. Gao Sheng et al. disturb the actual trajectory data to generate false trajectory data and then added it to the original trajectory data [15]. Su Qing et al. proposed a hierarchical role-based access control method for location privacy protection on mobile terminals and defined access permissions through secondary role definitions [16]. Related researchers have introduced game theory for privacy protection, but these methods are still in deep exploration [17]. In addition, the access control method for the service information provider to private information is not perfect. In order to improve the service quality, data mining and cloud computing technologies are used, but the privacy protection of the user privacy data is not controlled.

Game theory is not only a new branch of modern mathematics, but also an important discipline in operations research. It is a mathematical theory and method for studying the phenomena of struggle or competition [18]. Game theory has many applications in security, privacy, and wireless networks. It is a new research method to solve the problem of privacy protection based on game theory. Unlike traditional privacy protection methods, privacy protection methods based on game theory describe participants' benefits and costs, simulate their rational choice process through mechanism design and development strategy, and find the best solution for each party by analysing game equilibrium. There are many privacy protection models based on game theory that have been proposed in succession, such as the privacy protection model based on game theory proposed by Zhang Yixuan [19]. A game theory analysis for location privacy protection based on P-destination in mobile social networks is proposed by Bidi Ying and Amiya Nayak [20]. Consider users that collaborate with each other in the anonymous group and choose game strategies to provide their own data for other users so as to obtain desired services. The research methods for mobile location privacy are not perfect, and the issues of point location privacy and trajectory privacy are gradually exposed, which will also be the new direction of the research on privacy protection model based on game theory in the future.

3. Mobile Location Privacy Protection Scheme Based on Game Theory

This article starts with the characteristics of mobile location privacy and establishes a mobile location privacy access control model based on game theory. It implements dynamic control and protection of mobile location privacy information through access control.

3.1. Basic Concepts. Single point location refers to sensitive location information with frequency characteristics generated by individual activities. For example, someone who often appears at a certain location at a certain moment, such as family location, work place. The position information D may be represented as $D = \{u_i, (x_i, y_i, t_i)\}$. The privacy information of a single point of location can easily expose a person's fixed position. Once the disclosure is captured by criminals, the security threat to individuals is extremely high. Therefore, it is necessary to prevent attackers from obtaining such sensitive information.

Trajectory privacy refers to a series of sequential and frequent sensitive location information generated by individual activities, such as where persons are, where they are during a certain period of time, and what activity area is. The track information T can be represented as $T = \{u_i, (x_1, y_1, t_1), (x_2, y_2, t_2), (x_3, y_3, t_3), \ldots, (x_n, y_n, t_n)\}$. According to the trajectory privacy, it is easy to deduce personal habits, action route preferences, and even infer private information such as home address and work place. Therefore, it is necessary to prevent malicious attackers from obtaining such sensitive information.

In the position indication, id indicates the user identifier, t_i indicates the time of the user position information, and (x_i, y_i) indicates the two-dimensional position information of the user at time t_i.

Historical access record refers to the visitor's access record to a user's privacy information.

3.2. Problem Description

3.2.1. Location Privacy Leakage Caused by Information Superposition. The location information itself contains abundant temporal and spatial information, and the single location information has no substantial value, but the superposition of information can make the information integrated to obtain more valuable and specific information. Currently, due to the increasing location services, a large amount of user location information has been generated. Many enterprises and scientific research institutions have obtained a lot of analysis results related to user habits through the research and analysis of a large number of location information. Many of the location information are intrinsically related, especially for personal location information analysis, which can easily expose the user's privacy.

The model calculates the maximum tolerance of privacy information leakage by weighting the time, space, and frequency factors of the user's mobile location and restricts the amount of access to private information of the visitor.

3.2.2. Location Server Cannot Identify Malicious Access Behaviour and Implement Access Control. The user's location request information and real-time location information are stored on the location server. Malicious visitors can gain certain access rights through certain means and access the user's location record information. When the gained information is gradually increased, the amount of private information is also increasing. It is difficult for the server to detect such malicious behaviour and make timely response measures.

The above two key issues can be summarized as follows:

①Information superposition can cause indirect leakage of mobile location privacy.

②Identification of access behaviour and real-time dynamic control realized by the location server.

Different levels of thresholds are set for different visitors to solve the problem of dynamic access control.

This article focuses on privacy protection based on game theory for mobile location privacy. It mainly studies the characteristics of mobile location privacy and proposes a model suitable for protecting mobile location privacy access control using game theory in combination with current privacy information protection methods.

3.3. Model Structure Design. There are three parts in the model: the strategy execution part, the strategy deployment part, and the historical visit record. And there are five modules in total.

①Request acquisition module: module obtains a request for privacy information from a visitor request, and then it extracts critical information from access.

②Key information extraction module: it integrates visitor's mobile location privacy information request.

③Visitor's historical visit record module of mobile location privacy information: it records each visitor's historical mobile location privacy information request and provides reference for the threshold calculation module.

④Threshold calculation module: it sets the privacy location information threshold based on the game theory by the existing factor weight. Module performs weighted calculation based on historical record feedback results and feeds it back to the decision execution module.

⑤Decision execution module: it executes feedback made through the game strategy.

The design composition is shown as in Figure 1.

The basic element of model is as follows.

User set: $user = \{u_1, u_2, \ldots, u_n\}$, represented as a collection of all users who own information in the system.

The visitor set: $visitors = \{v_1, v_2, \ldots, v_n\}$, represented as a collection of all visitors in the system.

Access the role set: $roles = \{r_1, r_2, \ldots, r_n\}$, represented as a hierarchical collection of visitors in the system.

The system operation set: $operate = \{op_1, op_2, \ldots, op_n\}$, represented as a collection of response actions for the system for access requests.

3.4. Scheme Design

3.4.1. Factor Weight Setting. Firstly, it is necessary to analyse the key factors of privacy protection in mobile location. It

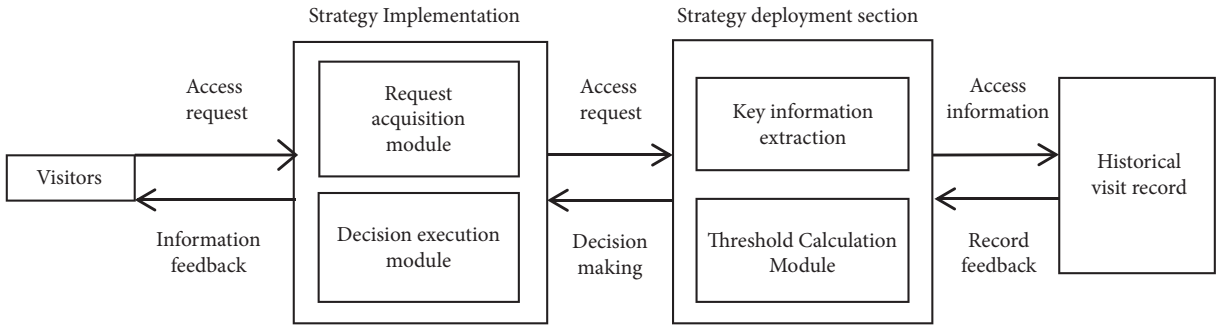


Figure 1: Access control method design composition.

performs weighted calculations and comparisons for time, space, and frequency. Summarizing the weight of each factor and then the weight setting suitable for this model is obtained. According to the result, the amount of information is calculated. Whether to respond to the access behaviour of location privacy information is decided based on the calculation results about the amount of information. The weight values of time, space, and frequency factor are set to be w_t, w_s, w_f.

3.4.2. Game Theory Strategy Implementation Plan. The scheme uses the noncooperative game theory related technology to control the visitors' behaviours. The recent behaviours of all visitors will be recorded and calculated for the degree of privacy mastered by the visitors. The server sets a threshold for the users' mobile location privacy information as the standard for that server controls visitors' access behaviour. Besides, the standard can also be regarded as the maximum value for allowing the visitor to grasp user's mobile location privacy information. When it is exceeded, the privacy of mobile users' location will be disclosed.

Assuming that the weighted result of accessing private information is *ac-R*, according to the single point location information $D = \{u_1, (x_1, y_1, t_1)\}$ given above, the time and space information can be known, and the frequency information can be obtained based on the historical access records. After integrating the three factor information, the weighted results of visit are calculated according to the following.

$$\begin{aligned} ac - R_{point} = &-\mathrm{w}_t P(x_t)\log_2 P(x_t) \\ &-\mathrm{w}_s P(x_s)\log_2 P(x_s) \\ &-\mathrm{w}_f P(x_f)\log_2 P(x_f) \end{aligned} \tag{1}$$

This equation aims at the request behaviour of single point location privacy information, calculates the amount of request information according to different factor request information conditions, and sums to obtain the amount of private information of this access behaviour.

According to track information $T = \{u_1, (x_1, y_1, t_1), (x_2, y_2, t_2), (x_3, y_3, t_3), \ldots, (x_n, y_n, t_n)\}$, the weighted result of accessing private information is obtained as follows.

$$ac - R_{trace} = -\sum_{i=1}^{k} \frac{1}{n}\log_2 \frac{1}{n} \tag{2}$$

This equation is for access behaviour of tracking private information. The number of single points contained in the track information is n. The number of location points for the request access behaviour is k, and $ac\text{-}R_{trace}$ indicates the amount of private information that the access behaviour can obtain.

3.4.3. Threshold Setting Scheme. This threshold is mainly composed of three factors: time, space, and frequency. Through analysis and summary of the impact of each factor on the privacy of the users' mobile location, the weights are calculated according to the result setting factors. The weighted result is calculated by the strategy deployment module for different weighted mobile location privacy information. Once the degree of visitor's privacy mastery exceeds the initial threshold, privacy information is no longer provided to the visitor. The threshold is set as *ac-threshold*. The formula for calculating the threshold is shown in

$$\begin{aligned} ac - threshold = &-\mathrm{w}_t P_1 \log_2 P_1 - \mathrm{w}_s P_2 \log_2 P_2 \\ &- \mathrm{w}_f P_3 \log_2 P_3 \end{aligned} \tag{3}$$

Suppose that the value of the access information extracted from the historical access record is *ac-V*. Whether to allow the visitors to access the privacy information of users' mobile location is determined by comparing *ac-threshold* with *ac-V+ac-R*. If *ac-threshold* is more than *ac-V+ac-R*, access is allowed; if *ac-threshold* is less than *ac-V+ac-R*, access is denied.

3.4.4. Dynamic Access Control Scheme. Scheme sets different access permissions for different visitors and judges whether the visitor has malicious access intention based on the access behaviour. Then it classifies the malicious access intentions as single point location privacy and trajectory privacy. This solution is designed to set access rights for the server's managers with different levels and applications of location privacy access behaviours.

Combine the visitor's recent history records and this visit request to determine the intention of the request and suppose the requested location information is $\{v_1, (u_1, x_1, y_1, t_1), (u_2, x_2, y_2, t_2) \ldots\ldots, (u_n, x_n, y_n, t_n)\}$. If the visitor's visit is successful, the recent visit record, location information (x, y), and time information t have more than 3 records with

Table 1: Game strategy matrix table.

System	Visitors	
	Goodwill visit (G)	Malicious visits (E)
Allow access (P)	Ssin-ac, Fper-ac	-Sevil-ac, Fper-ac
Access denied (R)	-Ssin-ac, 0	0, -Fre-ac

Table 2: System threshold adjustment parameters table.

	System threshold adjustment parameters
-Sevil-ac	-Sevil-ac
-Ssin-ac	Ssin-ac

similar information at a single day; then it is judged as a single point of privacy malicious visit. If there are more than 6 access records at the above single point and the time t for a single-day access record is time-series, then it is determined as the trajectory privacy malicious access behaviour. According to the system operation set $operate = \{op_1, op_2, \ldots, op_n\}$, in this model, four kinds of system operations are set up, that is, the system's permission and denial of access operations in the case of visitor's goodwill access and the system's permission and denial of access operations in the case of visitor's malicious access.

The game-theoretic strategy matrix is set as Table 1. The system revenue is divided into good-faith access revenue to *Ssin-ac* and malicious access revenue to *Sevil-ac*. Visitors' income is allowed access revenue to *Fper-ac* and denied access revenue to *Fre-ac*.

The system first makes early judgments on access behaviour, executes access control through late threshold comparisons, and then combines the previous judgment results with the access control execution results. Referring to the strategy matrix, the system revenue and visitor income are summarized; then adjust the corresponding system thresholds and visitor's permission through the revenue summary and make model access control more accurate. When the system revenue is positive, the system threshold is not adjusted. Instead ofwhen the system revenue is negative, the system threshold adjustment parameters are shown in Table 2.

4. Security Theoretical Analysis

4.1. Adversary Model. There are two types of visitor types: goodwill visitors and malicious visitors.

Instead of visiting a single user in depth and trying to get the user's identity information, goodwill visitors only access location information related to their needs.

A malicious visitor may attempt to obtain more user privacy information, including accessing more user records or attempting to obtain more location information for the same user record.

Malicious visitors can have the following capabilities:

(A) Information other than identity information about the user can be obtained.

Table 3: Privacy information transaction set table.

Transaction set number	Event collection
E1	{t p f s}
E2	{p f b s}
E3	{t f b s}
E4	{t p b s}
E5	{t f s}
E6	{t p f b s}

(B) An access request can be issued regardless of the type of access information.

(C) User partial location privacy information (less than the system-defined threshold) can be obtained.

4.2. Factor Weight Analysis. Supposing a user's location information transaction set is organized and a conclusion can be made using theories related to data mining association rules, the results are shown in Table 3.

t represents time information (assuming t = 21:00), p represents location information, f represents frequency information (assuming f is the situation where frequency is greater than 5), b represents background knowledge, and s represents the situation of privacy leakage (assuming s is the home address information disclosure situation).

According to Table 3, there are 6 transactions, $E = \{E_1, E_2, E_3, E_4, E_5, E_6\}$, the item set $I = \{t, p, f, b, s\}$, and $E\ (t, s)$ represents the transaction that has both event t and event s.

According to data mining related knowledge, $support_{x,y} = N_{E(x,y)}/N_E$, $confidence_{x,y} = N_{E(x,y)}/N_{E(x)}$, $N_{E(x,y)}$ represents the number of transaction containing both events x and y, and $N_{E(x)}$ represents the number of transactions containing event x. Something can be obtained as follows:

$$\text{Rule } t \text{ —>} s,\ support_{t,s} = 5/6,\ confidence_{t,s} = 1.$$

$$\text{Rule } p \text{ —>} s,\ support_{p\ s} = 2/3,\ confidence_{p,s} = 1.$$

$$\text{Rule } f \text{ —>} s,\ support_{f,s} = 5/6,\ confidence_{f,s} = 1.$$

The proportion of time factor $w_t = support_{t,s}/(support_{t,s} + support_{t,s} + support_{t,s}) = 0.357$. Similarly, the weight coefficient $w_t = 0.357$, $w_p = 0.286$, $w_f = 0.357$.

From the above analysis and calculation results, it can be seen that time, location, and frequency factors account for different proportions of location privacy information. That is, information sensitivity is different. The method of this scheme can distinguish the weights of different influencing factors of location privacy information and express them in numerical form.

4.3. Access Hypothesis Analysis

4.3.1. Prior Judgment of Access Behaviour. First, when a visitor accesses private information, the access control system firstly retrieves the historical access record and guesses whether the access behaviour is legal to the privacy of single point or trajectory privacy according to the historical access. If there is no history access record, the system will think of it as

Table 4: Threshold setting table.

Visitors	Single point privacy threshold	Trajectory privacy threshold
First level	0.654	0.764
Second level	0.818	0.955
Third Level	0.981	1.146

point location privacy access by default. According to role set $roles = \{r_1, r_2, \ldots, r_n\}$, the total level of visitors is 3, and the number of records that visitors in each level can visit increases by levels. The first-level visitors can access 4 records. The second-level visitors can access 5 records, and the third-level visitors can access 6 records. If this condition is satisfied, it will be judged as normal goodwill access behaviour. If this condition is not satisfied, it will be judged as malicious access behaviour.

The threshold is set for different visitors. The single point location data contains thirty records, and the track data is taken 24 times a day. According to the analysis of test, when the single point location privacy disclosure is more than five records, the sensitive information of the user is considered to be leaked; when the trace data disclosure is more than 5 pieces, the privacy information of the user is considered to be leaked. The threshold setting table is shown in Table 4.

4.3.2. Visit by Peer Visitors. Assume that visitors A and B are first-level visitors. A and B visit the same single point privacy information of a mobile user. The query information is $\{v_A, (u_1, x_1, y_1, t_1), v_B, (u_1, x_1, y_1, t_1)\}$, set the visitor A's history access record value as $ac\text{-}V(A) = 0.432$ and visitor B's history access record value as $ac\text{-}V(B) = 0.537$.

According to (1) we can get the following.

$$\begin{aligned} ac\text{-}R_{point}(A) &= 0.153 \\ ac\text{-}R_{point}(B) &= 0.153 \\ ac\text{-}V(A) + ac\text{-}R_{point}(A) &= 0.585 \\ ac\text{-}V(B) + ac\text{-}R_{point}(B) &= 0.690 \end{aligned} \tag{4}$$

The first-level visitor's access single point privacy threshold $ac\text{-}threshold_1$ is 0.654. According to $ac\text{-}threshold_1$ being more than $ac\text{-}V(A) + ac\text{-}R_{point}(A)$, $ac\text{-}threshold_1$ is less than $ac\text{-}V(B) + ac\text{-}R_{point}(B)$. Visitor A's privacy information access request can be answered while visitor B's privacy information access request is denied.

4.3.3. Visit by Different Levels Visitor. Assuming visitor C is a first-level visitor, visitor D is a second-level visitor. They visit the same trajectory privacy information of a mobile user, and the query information is $\{v_C, v_D, (u_2, x_2, y_2, t_2), (u_2, x_3, y_3, t_3)\}$; set the history access record value of visitor C as $ac\text{-}V(C) = 0.373$ and the historical access record value of visitor D as $ac\text{-}V(D) = 0.576$.

According to formula (2) we can get the following.

$$\begin{aligned} ac\text{-}R_{trace}(C) &= 0.382 \\ ac\text{-}R_{trace}(D) &= 0.382 \\ ac\text{-}V(C) + ac\text{-}R_{trace}(C) &= 0.755 \\ ac\text{-}V(D) + ac\text{-}R_{trace}(D) &= 0.958 \end{aligned} \tag{5}$$

The value of the first-level visitor's access trajectory privacy threshold $ac\text{-}threshold_1$ is 0.764, and the value of the second-level visitor's access threshold $ac\text{-}threshold_2$ is 0.955; then according to condition, $ac\text{-}threshold_1$ is more than $ac\text{-}V(C)$ plus $ac\text{-}R_{trace}(C)$ and $ac\text{-}threshold_2$ is less than $ac\text{-}V(D)$ plus $ac\text{-}R_{trace}(D)$. Visitor C's privacy information access request can be answered while visitor D's privacy information access request is denied.

From the above calculation and analysis, we can see that, for the same access request of the peer visitors, the access control model can make different decisions according to different historical access records. For the same access request of nonpeer visitors, the access control model can also make a more appropriate judgment based on historical access records. There are no fixed restrictions on the visitors' access behaviour, but the protection of location privacy information can still be achieved. And the information entropy theory focuses on the calculation of information amount of the whole information set. Applying this theory in access control is easy to lead to missing the information of single access. The system prejudgment method set by this model can make up for this defect, and, as a reference condition of the game strategy matrix, it also can help to adjust the access control model.

5. Game Strategy Analysis

5.1. Threshold Adjustment. The adjustment parameter of the system is $Ssin\text{-}ac = Sevil\text{-}ac = 0.1$.

(1) Assume that visitor E is a first-level visitor. He accesses the track privacy information of a mobile user. The query information is $\{v_E, (u_1, x_1, y_1, t_1), (u_1, x_2, y_2, t_2)\}$. Suppose visitor E's historical access record value is $ac\text{-}V(E) = 0.472$. According to the prejudgment criteria, E can access up to four records as a first-level visitor. So E's access request meets the goodwill visit standard.

$$\begin{aligned} ac\text{-}R_{trace}(E) &= 0.299 \\ ac\text{-}V(E) + ac\text{-}R_{trace}(E) &= 0.771 \end{aligned} \tag{6}$$

The first-level visitor's access threshold is $ac\text{-}threshold_1 = 0.764$, so $ac\text{-}threshold_1 < ac\text{-}V(E) + ac\text{-}R_{trace}(E)$. Visitor E's request of privacy information access can be denied. It satisfies the situation that the goodwill visitor be refused to access by system in the policy matrix. At this time, the benefit of the system is negative, which satisfies the adjustment condition of system parameter. According to Table 2, adjust the parameter as *Ssin-ac*. After adjustment, the system threshold is adjusted to $ac\text{-}threshold_1 + Ssin\text{-}ac = 0.774$. At this point, adjust this policy to allow the access requests of E.

(2) Assume that visitor F is a first-level visitor. He accesses the track privacy information of a mobile user. The query information is $\{v_F, (u_1, x_1, y_1, t_1), (u_1, x_2, y_2, t_2), (u_1, x_3, y_3, t_3), (u_1, x_4, y_4, t_4), (u_1, x_5, y_5, t_5)\}$, and suppose visitor F's historical access record value is $ac\text{-}V(F) = 0.285$. According to the prejudgment criteria, as a first-level visitor, F can access up to four records, so F's access request cannot meet the goodwill visit standard and will be judged as malicious access.

$$\begin{aligned} ac\text{-}R_{trace}(F) &= 0.471 \\ ac\text{-}V(F) + ac\text{-}R_{trace}(F) &= 0.756 \end{aligned} \tag{7}$$

The first-level visitor's access threshold is $ac\text{-}threshold_1 = 0.764$, so $ac\text{-}threshold_1 > ac\text{-}V(F) + ac\text{-}R_{trace}(F)$. Visitor F's request of privacy information access can be responded. It satisfies the situation that the malicious visitor be allowed to access by system in the policy matrix. At this time, the benefit of the system is negative, which satisfies the adjustment condition of system parameter. According to Table 2, adjust the parameter as *Sevil-ac*. After adjustment, the system threshold is adjusted to $ac\text{-}threshold_1\text{-}Sevil\text{-}ac = 0.754$. At this point, adjust this policy to deny the access requests of F.

5.2. Game Theory Analysis. When a visitor conducts a malicious visit, the gain of visitor is divided into *Fre-ac* and *–Fre-ac* according to the system whether access is allowed. At the same time, the system's gain is divided into *-Sevil-ac* and 0. On the contrary, when the visitor carries out a goodwill visit, the gain of visitor is divided into *Fper-ac* and 0 while the gain of the system is divided into *Ssin-ac* and *-Ssin-ac*, according to whether the system access is allowed. This is the game of incomplete information.

Assume that the probability of a visitor who performs a good visit is x and a malicious visit is $1\text{-}x$. Assuming that the system can accurately identify the visitor's access intent, the probability of the system allowing access is also x, and the probability of denying access is $1\text{-}x$. The combination of system and visitor strategies is $S_1 = (P, G)$, $S_2 = (P, E)$, $S_3 = (R, G)$, $S_4 = (R, E)$. The visitor's income value *income* is shown in

$$\text{Income} = \sum_{k} \text{income}_k(S_k)\, p_f \tag{8}$$

$Incomek(S_k)$ is the gain when selecting the S_k strategy; p_f is the probability of choosing the k strategy. According to the above assumptions, (9) is as follows:

$$\text{Income}_f(G) = x * \text{Fper} - ac \tag{9}$$

Incomef(G) is the expected value of income when the visitor makes a good visit;

$$\text{Income}_f(E) = x * \text{Fper} - ac + (1 - x)(-\text{Fre} - ac) \tag{10}$$

Incomef(E) is the expected value of income when a visitor visits maliciously;

$$\text{Income}_s(P) = x * \text{Ssin} - ac + (1 - x)(-\text{Sevil} - ac) \tag{11}$$

Incomes(P) is the expected value of the income when the system adopts an allowed access strategy;

$$\text{Income}_s(R) = x(-\text{Ssin} - ac) \tag{12}$$

Incomes(R) is the expected value of the income when the system adopts the deny access policy.

Informef(G)> Incomef(E) can be derived from (9) and (10). That is, regardless of the probability of a good-faith visit of the visitor, the visitor's good-faith access behaviour will have a greater expected total return value than a malicious visit. So the visitors should choose good-faith visits as they wish to maximize their own interests.

In order to ensure the normal services of system location access, the system should try its best to meet visitors' access needs and ensure that privacy information is not revealed, it should satisfy the following: *Incomes(P)* is larger than *Incomes(R)*. It is shown as

$$x * \text{Ssin} - ac + (1 - x)(-\text{Sevil} - ac) > x(-\text{Ssin} - ac) \tag{13}$$

When x is more toward the value of 1, the fuller condition is established in (13). That is to say, when the visitor adopts a good-faith access behaviour, the system's profit expectation also tends to be maximum.

From this, it can be concluded that good-faith access behaviour taken by visitors and access allowed by the systems can be seen as the Nash equilibrium strategy. And, at this point, both parties can get the most benefit.

From the analysis of the first two parts, we can see that the game theory strategy matrix set by this model can adjust the threshold parameters when the system strategy is deviant, so that the system can make more appropriate judgments. This makes the model access control to be better for location privacy information and ensures the dynamic of access control. At the same time, considering the first case, this paper argues that the threshold cannot be adjusted every time to avoid being used by malicious visitors to increase the threshold and reduce the system access control conditions. It ensures the stability of the system.

6. Comparative Analysis

Most of the existing mobile location privacy protection methods are for location information publishing and query operations, but these methods still cannot fully protect user location privacy information. For example, the (K, C) L privacy standard and its improvement methods are mainly to protect location privacy information against location trajectory tracking attacks. The continuous anonymity and its related improvement methods are mainly to protect the privacy of trajectory reconstruction attacks. The method proposed in this paper controls the amount of information accessed by normal visitors and attackers using the standard of information volume, so the attacker can only obtain limited private information and the difficulty of tracking and reconstructing attacks is increased.

Existing mobile location privacy protection methods have their own advantages, but, with the development of big data environment, they also have many drawbacks. For example,

Table 5: Comparative analysis table.

	Considering different information factors	No curing access operation is required	Strategic adjustment mechanism
This model	√	√	√
[9]	×	×	√
[12]	×	×	×
[14]	×	×	×
[16]	×	√	×
[20]	×	√	×

anonymity is effective for single-private data, but it cannot avoid the privacy leakage caused by the joint reconstruction attacks of multiple data sources. At the same time, there are many privacy protection methods for emerging technologies. For example, the location privacy protection development model and task allocation strategy in mobile cloud computing of ad hoc networks are proposed in [12]. A truthful incentive mechanism to protect location privacy in mobile crowdsourcing system is proposed in document [9]. And location service providers have a lot of user privacy information. Some service providers use existing data for data mining. However, they did not conduct more comprehensive access control over private data, which led to internal data leakage or external attack access leakage. The method proposed in this paper is mainly aimed at access control of data in the above two kinds of privacy leakage, and a game theory strategy matrix is introduced. It can complement the existing mobile location privacy protection method from the perspective of a location service provider and improve the effectiveness of location privacy protection and data utilization.

Compared with the game analysis of location privacy protection based on P-destination proposed in literature [20], the game theory is applied to the defence strategy game between the attacker and the system to maximize the system benefits. The location privacy protection method based on game theory proposed in literature [16] uses game theory to embody the behaviour of user cooperation in probability. The scheme proposed in this paper not only considers the execution of policy and the prediction of visitor behaviour, but also pays attention to the control of the amount of information that the visitor grasps and the adaptability of the scheme. Game theory strategy matrix adopted in this paper participates in the access control parameter adjustment, which is more conducive to the realization of dynamic and accurate and feasible access control strategies. So the scheme is more secure. Because this method does not identify and control collusion attacks, it can resist simple collusion attacks because of its restriction on the amount of information of visitors.

In Table 5, we compare and summarize three aspects: location privacy influencing factor, access operation, and adjustment mechanism.

7. Concluding Remarks

This paper provides a new access control method for single point location privacy access and trajectory privacy access in mobile location and achieves dynamic access control by setting visitor level and threshold for accessing privacy information. After security analysis, this method has certain practicality and scientificity.

Data Availability

The data used to support the findings of this study are included within the article.

Disclosure

An earlier version of this study was presented as a conference paper in ICCCS 2018: International Conference on Communication and Computing Systems.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was partially supported by Hebei Education Department (QN2015231, QN2017132), Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, Research and Practice of Higher Education Teaching Reform of Hebei Education Department (no. 2016GJJG112), Doctoral Scientific Research Foundation of Shijiazhuang Tiedao University (Z991015137), and Natural Science Foundation of Hebei Province (no. F2017210161, no. F2018210109).

References

[1] J. Kaur and K. Kaur, "A fuzzy approach for an IoT-based automated employee performance appraisa," *Computers Materials and Continua*, vol. 53, no. 1, pp. 24–38, 2015.

[2] M. Guo, S. B. Zhang, and X. D. Li, "Research on location privacy protection technology in IoT," *Journal of Chinese Computer Systemspp*, no. 9, pp. 1961–1965, 2017.

[3] M. Rong, X. H Chen, and H. Liu, "Research on user privacy measurement and privacy protection in mobile crowdsensing," *Netinfo Security*, no. 8, pp. 1671-1122, 2018.

[4] C. R. Wu, E. Zapevalova, and Y. W. Chen, "Time optimization of multiple knowledge transfers in the big data environment," *Computers Materials and Continua*, vol. 54, no. 3, pp. 269–285, 2015.

[5] J. P Tang and L. Li, "Big data sensing information processing platform for intelligent traffic," *Applied Mechanics Materials*, vol. 667, pp. 324–327, 2014.

[6] H. T. Chi, *Location Privacy Protection Based on Location Service*, Xidian University, 2014.

[7] P. Samarati and L. Sweeney, *Protecting Privacy when Disclosing Information: k-Anonymity and Its Enforcement through Generalization and Suppression*, 1998.

[8] 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.

[9] Y. Wang, Z. Cai, X. Tong, Y. Gao, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems," *Computer Networks*, vol. 135, pp. 32–43, 2018.

[10] H. Zhang, C. Wu, Z. Chen, Z. Liu, and Y. Zhu, "A novel online spatial-temporal k-anonymity method for location privacy protection from sequence rules-based inference attacks," *PLoS ONE*, vol. 12, no. 8, 2017.

[11] K. Huguenin, I. Bilogrevic, J. S. Machado et al., "A predictive model for user motivation and utility implications of privacy-protection mechanisms in location check-ins," *IEEE Transactions on Mobile Computing*, vol. 17, no. 4, pp. 760–774, 2018.

[12] Y. Gong, C. Zhang, Y. Fang, and J. Sun, "Protecting location privacy for task allocation in ad hoc mobile cloud computing," *IEEE Transactions on Emerging Topics in Computing*, vol. 6, no. 1, pp. 110–121, 2018.

[13] P. Zhao, J. Li, F. Zeng et al., "ILLIA: enabling k-Anonymity-based privacy preserving against location injection attacks in continuous LBS queries," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1033–1042, 2018.

[14] K. Al-Hussaeni, B. C. M. Fung, and W. K. Cheung, "Privacy-preserving trajectory stream publishing," *Data & Knowledge Engineering*, vol. 94, pp. 89–109, 2014.

[15] S. Gao, J. Ma, W. Shi, G. Zhan, and C. Sun, "TrPF: A trajectory privacy-preserving framework for participatory sensing," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 874–887, 2013.

[16] Q. Su and Y. Z Li, "Research of privacy protection method based on access control," *Journal of Yanbian University*, vol. 42, no. 1, pp. 69–74, 2016.

[17] D. D. Zhou, W. W. Li, and Y. Q. Sun, "Survey on game theory based privacy protection," *Journal of Chinese Computer Systems*, vol. 36, no. 12, pp. 2696–2700, 2015.

[18] S. K. Zhang and Y. D. Zhang, *Methods of Modern Game Theory and Mathematics in Economic*, Dongbei University of Finance and Economics Press, 2016.

[19] Y. X. Zhang, J. S. He, B. Zhao, and N. F. Zhu, "A privacy protection model based on game theory," *Chinese Journal of Computers*, vol. 39, no. 3, pp. 615–627, 2016.

[20] B. Ying and A. Nayak, "Location privacy-protection based on p-destination in mobile social networks: A game theory analysis," in *Proceedings of the 2017 IEEE Conference on Dependable and Secure Computing*, pp. 243–250, Taiwan, August 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 5823439, 11 pages
https://doi.org/10.1155/2018/5823439

Research Article

A Data Leakage Prevention Method Based on the Reduction of Confidential and Context Terms for Smart Mobile Devices

Xiang Yu,[1] Zhihong Tian,[2] Jing Qiu,[2] and Feng Jiang[3]

[1]*School of Electronics and Information Engineering, Taizhou University, Taizhou 318000, China*
[2]*Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China*
[3]*College of Science and Technology, Harbin Institute of Technology, Harbin 150001, China*

Correspondence should be addressed to Zhihong Tian; tianzhihong@gzhu.edu.cn and Jing Qiu; qiujing@gzhu.edu.cn

Received 27 April 2018; Revised 30 August 2018; Accepted 24 September 2018; Published 21 October 2018

Guest Editor: Ding Wang

Early data leakage protection methods for smart mobile devices usually focus on confidential terms and their context, which truly prevent some kinds of data leakage events. However, with the high dimensionality and redundancy of text data, it is difficult to detect the documents which contain confidential contents accurately. Our approach updates cluster graph structure based on CBDLP (Data Leakage Protection Based on Context) model by computing the importance of confidential terms and the terms within the range of their context. By applying CBDLP with pruning procedure which has been validated, we further remove the redundancy terms and noise terms. Actually, not only can confidential terms be accurately detected but also the sophisticated rephrased confidential contents are detected during the experiments.

1. Introduction

With the development of Internet and information technology, smart mobile devices appear in our daily lives, and the problem of information leakage on smart mobile devices will follow which has become more and more serious [1, 2]. All kinds of private or sensitive information, such as intellectual property and financial data, might be distributed to unauthorized entity intentionally or accidentally. And that it is impossible to prevent from spreading once the confidential information has leaked.

According to survey reports [3, 4], most of the threats to information security are caused by internal data leakage. These internal threats consist of approximate 29% private or sensitive accidental data leakage, approximate 16% theft of intellectual property, and approximate 15% other thefts including customer information, and financial data. Further, the consensus of approximate 67% organizations shows that the damage caused from internal threats is more serious than those form outside.

Although laws and regulations have been passed to punish various behaviors of intentional data leakage, it is still hard to prevent data leakage effectively. Confidential data can be easily disguised by rephrasing confidential contents or embedding confidential contents in nonconfidential contents [5, 6]. In order to avoid the problems arising from data leakage, lots of software and hardware solutions have been developed which are discussed in the following chapter.

In this paper, we present CBDLP, a data leakage prevention model based on confidential terms and their context terms, which can detect the rephrased confidential contents effectively. In CBDLP, a graph structure with confidential terms and their context involved is adopted to represent documents of the same class, and then the confidentiality score of the document to be detected is calculated to justify whether confidential contents is involved or not. Based on the attribute reduction method from rough set theory, we further propose a pruning method. According to the importance of the confidential terms and their context, the graph structure of each cluster is updated after pruning. The motivation of the paper is to develop a solution which can prevent intentional or accidental data leakage from insider effectively. As mixed-confidential documents are very common, it is very important to accurately detect the documents containing confidential contents even when most of the confidential contents have been rephrased.

The remainder of this paper is organized as follows. In Section 2, we introduce previous related work on data leakage prevention. In Section 3, we present CBDLP model together with the corresponding clustering, decision, and calculation algorithms. The experiments conducted to evaluate CBDLP in all circumstance are discussed in Section 4. Finally, Section 5 concludes this paper and discusses the directions of our future research.

2. Related Work

In this section, we review clustering of textual documents, attribute reduction method, and graph representation of textual documents, respectively.

2.1. Clustering of Textual Documents. The problem of clustering textual documents is similar to high dimensional clustering. In general, each term of a textual document is considered as an independent dimension and then each document is considered as a vector consists of thousands of terms. By calculating the angle cosine measure between documents, textual documents can be classified in terms of similarity which is reflected by the angle cosine value [7–10].

Vector space model, VSM, is one of the most widely used text representation models [11], which is first presented by Salton in the 1960s and successfully applied in SMART, a system for the manipulation and retrieval of text. In VSM model, a textual document is represented as $D = D((T_1, W_1), (T_2, W_2) \ldots, (T_n, W_n))$, where T_i and W_i denote the ith term and its weight in the document, respectively, and then the classification of a textual document is determined by calculating the similarity between the textual document to be classified and the textual documents whose classification are already known.

Term frequency and inverse document frequency, TF-IDF, is a frequently employed and effective statistics method which is used to evaluate the importance of a term for a documents collection [12]. As is well known, the importance of a term is proportional to the frequency of its occurrence in a document and is inversely proportional to the frequency of its occurrence in the whole corpus. Till now, TF-IDF has been widely used in various fields, such as text mining, search engine, and information retrieval.

On the basis of VSM model and TF-IDF method, existing textual documents clustering algorithms can be divided into five main categories: partitioning methods, hierarchical methods, density-based methods, grid-based methods, and model-based methods. Partitioning methods, which are efficient and insensitive to the sequence of documents, divide n documents into k clusters in terms of clustering criteria. The representative partitioning methods include k-means and k-medoids [13, 14]. Hierarchical methods disintegrate documents into different clusters or integrate different documents together into one cluster in terms of the similarity with a top-down or bottom-up hierarchical manner. The representative hierarchical methods include BIRCH and CURE [15, 16]. Other than partitioning methods, density-based methods focus on the density of a certain area. When the density of the documents within a certain area exceeds a predefined threshold, they are incorporated into the same cluster. The representative density-based methods include DBSCAN and OPTICS [17, 18]. Grid-based methods partition data space into limited cells in advance and integrate adjacent cells whose density exceed the density threshold into the same cluster. The representative grid-based methods include STING and CLIQUE [19, 20]. In model-based methods, different models are bound up with each cluster respectively, and the objective is to find all data subsets that fit each model best. Statistics solution, such as SVM [21], and neural network solution are adopted extensively in model-based methods [22]. The support vector clustering algorithm created by Hava Siegelmann and Vladimir Vapnik applies the statistics of support vectors, developed in the support vector machines algorithm, to categorize unlabeled data, and is one of the most widely used clustering algorithms in industrial applications.

In this paper, we calculate the angle cosine value which reflects the similarity between documents and cluster documents with DBSCAN. DBSCAN, proposed by Martin Ester in 1996, is a density-based clustering algorithm which is widely cited in scientific literature [23], and it is awarded the test of time award in 2014 [24]. When clustering, other than k-means, DBSCAN does not need to specify the number of clusters and it can find the clusters of arbitrary shape. In addition, DBSCAN is robust to outliers as opposed to k-means.

2.2. Graph Representation of Textual Documents. Graphs have already been used in many text-related tasks, which employ graph as the model for text representation instead of the existing methods [25]. As an alternative method to the vector space model for representing textual information, graphs can be created from documents and be further used in the text-related tasks such as information retrieval [26], text mining [27], and topic detection [28].

In general, graph-based model is usually employed in the domain of information retrieval such as PageRank [29] and HITS [30]. When determining the similarity, graph matching, which is generally used to detect similar documents, is NP in complexity [31], whereas the methods based on vector space model perform efficiently by calculating the Euclidean distance or Cosine measure between document vectors [32]. The main advantage of graph-based model is that it can not only capture the contents and structure of a document but also represent the terms together with their context. To the best of our knowledge, graph-based model is seldom employed in text-related tasks. Schenker presents a graph-related algorithm with its several variants [33] in which a graph is presented and the terms connected with edges are considered as nodes. The differences between the variants are related to term-based techniques. Gilad Katz presents CoBAn, a context based model for data leakage prevention, which enlightens us a lot [34]. However, CoBAn is partly influenced by the limitation of k-means which is employed in CoBAn. Moreover, there might exist some redundancy nodes in the graph generated in CoBAn. Xiaohong Huang et al. propose an Adaptive weighted Graph Walk model (AGW) to

solve the problem of transformed data leakage by mapping it to the dimension of weighted graphs [35].

In this paper, we employ a hybrid approaches which combines graph and vector representations. When clustering documents, we employ DBSCAN with cosine measure. When representing the confidential textual content and its context of each cluster, the graph of each cluster which includes only confidential and contextual nodes is created.

2.3. Redundancy Information Reduction. When dealing with text-related tasks, redundancy information is generally useless and even worse, it might decrease the efficiency of task execution. There exist many representative redundancy information reduction methods such as PCA [36], SVD [37], LSI [38], etc. The principle of PCA is to transform multiple attributes into a few primary attributes, which can reflect the information of original data effectively. However, the complexity of PCA is generally high and there might be part of original information loss. More than characteristics, SVD has almost the same advantages and disadvantages as PCA does. LSI represents textual data with latent topics that consists of specific terms, but in most cases, the influence of specific terms are ignored. In this paper, the reduction method from rough set theory, as shown in Section 3, is employed and partly recomposed to meet requirements.

2.4. Data Leakage Prevention. With the number of leakage incidents and the cost they inflict continues to increase, the threat of data leakage posed to companies and organizations has become more and more serious [39–41]. Considering the enormity of data leakage prevention, various models and approaches have been developed to address the problem of data leakage prevention. Tripwire is a more recent prototype system proposed by Joe DeBlasio et al. in 2017; it registers honey accounts with individual third-party websites, and thus access to an email account provides indirect evidence of credentials theft at the corresponding website [42]. However, Tripwire is more suitable for forensics rather than confidential data leakage prevention. In 2018, Wenjia Xu et al. propose a new promising image retrieval method in ciphertext domain by block image encrypting based on Paillier homomorphic cryptosystem which can manage and retrieve ciphertext data effectively [43]. Nevertheless, the method focus on data encryption rather than data detection. Since smart devices based on ARM processor become an attractive target of cyberattacks, Jinhua Cui et al. present a scheme named SecDisplay for trusted display service in 2018, it protects sensitive data displayed from being stolen or tampered surreptitiously by a compromised OS [44]. But it pays less attention to the scenarios of intentional or accidental data leakage from insider. According to the work of Ding Wang et al., lots of authentication schemes have been proposed to secure data access in industrial wireless sensor networks, however, they do not work well [45]. In addition, Ding Wang et al. develop a series of practical experiments to evaluate the security of four main suggested honeyword-generation methods and further prove that they all fail to provide the expected security [46].

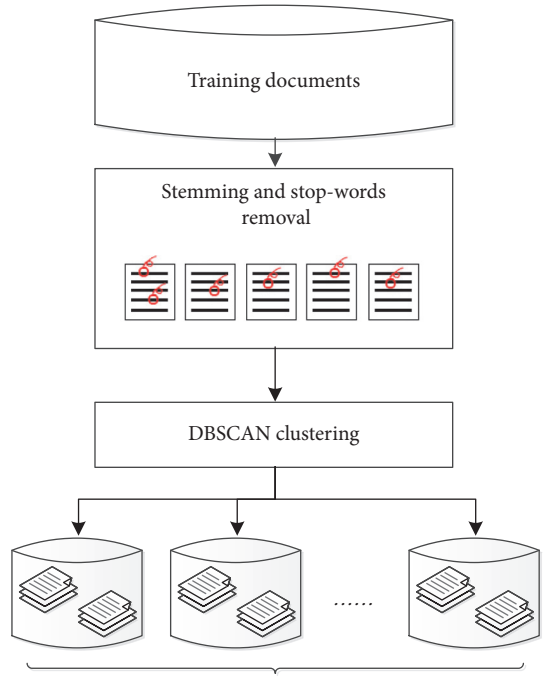


Figure 1: DBSCAN method.

3. CBDLP Model

CBDLP consists of training phrase and detection phrase. The training phrase can be further divided into three steps, clustering step, graph building step, and pruning step. During the training phrase, the training documents are first classified into different clusters, then each cluster is represented by graph, and finally the nodes of each graph are pruned in terms of their importance. During the detection phrase, documents are matched to the graphs of clusters respectively and the confidential scores are calculated. A document is considered as confidential only if its confidential score exceeds a predefined threshold. The detail of CBDLP training phrase is presented in Algorithm 1.

3.1. Clustering Documents with DBSCAN. In the first step, we apply stemming and stop-words removal to all documents in training set, and transform the processed documents into vectors of weighted terms. After applying DBSCAN with cosine measure to the vectors, which represent the training documents, each resulting cluster represents an independent topic of training documents and there might exist both confidential and nonconfidential documents. As shown in Figure 1.

The procedure of DBSCAN is described as follows:

Step 1. A data set is given with n documents and ε as the threshold of minimal similarity between documents of the same cluster and $MinPts$ as the threshold of minimal number of documents in a cluster.

Input: C - Confidential documents set
N - Non-confidential documents set
TR_{min} - The minimum similarity threshold
Output: CR - The set of clusters, each with the centroid and corresponding graph
CT - The set of confidential terms in clusters
$ContextT$ - The set of context terms

(1) $T \longleftarrow C \cup N$
(2) $CR \longleftarrow DBSCAN\,(T)$ % The result of clustering T is saved in CR
(3) Initializing $CT[CR]$ %The scores of confidential terms are saved in CT
(4) Initializing $ContextT$ %The context terms set of each confidential term is saved in $ContextT$
(5) **for** (each cr in CR)
(6) { Calculate the similarity between cr and the other clusters
(7) Create language model for cr, and calculate the scores for each confidential term
(8) $TR \longleftarrow$ initial the threshold of cluster similarity
(9) **while** $(TR > TR_{min})$
(10) { $C_{temp} \longleftarrow$ All clusters whose similarity to $cr > TR$
(11) Create language model for the documents of C_{temp}
(12) $CT[cr] \longleftarrow$ Based on new language model, Update the scores of confidential terms
(13) **for**(each confidential term ct in cr)
(14) { Detect the occurrence of ct in $C \cup N$
(15) $P_{confidential\_doc}(term, ct) \longleftarrow$ For each context term of ct, calculate the probability of the appearance both ct and the context term in confidential documents.
(16) $P_{non\_confidential\_doc}(term, ct) \longleftarrow$ For each context term of ct, calculate the probability of the appearance both ct and the context term in non-confidential documents.
(17) $ContextT \longleftarrow$ Calculate the value of $P_{confidential\_doc}(term, ct)/P_{non\_confidential\_doc}(term, ct)$ for each confidential term ct
(18) Detect all clusters whose similarity is greater than TR, and detect the occurrences of all terms in the clusters.
(19) $ContextT \longleftarrow$ Update the probability of the context terms that appear in the scopes of different confidential terms
(20) }
(21) Reduce the value of TR
(22) }
(23) }

Algorithm 1: CBDLP.

Step 2. Start with an arbitrary document that has not been visited and find all the documents in its ε-neighborhood. If the number of documents in the neighborhood exceeds $MinPts$, incorporate the documents into the same cluster and label them.

Step 3. If not all documents have been visited, start from another arbitrary document which has not been visited.

Step 4. Mark the documents which are not labelled as noise.

3.2. Representing Clusters with Graph. In this step, the confidential contents in all clusters, which include not only the confidential terms but also their context, need to be represented by graphs. The procedure of creating graph representation for the clusters which include confidential contents is described as follows:

(1) Detect the confidential terms provided by domain experts or inferred from the key terms of training documents.

(2) Analyze the context of each confidential terms.

(3) Create the graph representation for confidential terms and their contexts on the cluster level.

3.2.1. Detect Confidential Terms. In general, a term, which appears in confidential documents with high probability and appears in nonconfidential documents with low probability, is considered as confidential term. We first build language models for the confidential and non-confidential documents of the same cluster, which are denoted by *cVM* (confidential vector model) and *ncVM* (nonconfidential vector model). Then the confidentiality score can be represented by the ratio of its probability in confidential documents to that in non-confidential documents as shown in as follows, where $P_{cVM}(t)$ and $P_{ncVM}(t)$ denote the probability of term t in confidential and nonconfidential language models, respectively:

$$\forall t \in cVM,$$
$$score+ = \frac{P_{cVM}(t)}{P_{ncVM}(t)} \tag{1}$$

However, there may exist the following problem. If a cluster includes only few nonconfidential documents or possibly none at all, its language model cannot fully represent the nonconfidential documents in it. The solution we proposed follows an expanding manner; we first predefine the minimal similarity threshold TR_{min} and iteratively expand the *ncVM* to include more clusters. TR is referred to as the similarity threshold of cosine measure of the cluster with few

nonconfidential documents. Note that not all clusters need to be expanded. After each iteration, we lower the value of TR. Unless TR is greater than TR_{min}, the nonconfidential documents of the expanding clusters are included to recalculate the scores of terms in original cluster.

When the adjacent clusters are included and the scores of confidential terms are recalculated, each term whose score is greater than 1 is considered as confidential term, which means the term is more likely to appear in confidential documents than in non-confidential documents. After this phase, the set of confidential terms, CT, is obtained.

3.2.2. Analyze the Context of Confidential Terms. After confidential terms detection, we further analyze the context of confidential terms. Apparently, a term is more likely to be considered as confidential if it appears in the similar contexts in other confidential documents. Inversely, if the context of a confidential term frequently appears in nonconfidential documents, the probability of the confidential term being part of confidential contents is lower.

As a predefined parameter, context span η determines the number of terms that precede and follow the confidential term. Context span with high value might increase the computational cost, inversely, and context span with low value could not provide adequate context information of confidential terms. Experimental results show that $\eta = 10$ tends to be the optimal value of context span in our experiments, which means that the context of a confidential term consists of the five terms preceding it and the other five following it. Apparently, only the context of the confidential terms in confidential documents needs to be taken into account.

The probabilities of a confidential term together with its context appearing in confidential documents and nonconfidential documents, which are denoted by $P_c(key_{context}/key)$ and $P_{nc}(key_{context}/key)$, are calculated separately. If the former is higher than the latter, the corresponding confidential contents can be well represented by the confidential term with its context. $P_c(key_{context}/key)$ is defined as the number of confidential documents in which the confidential term with its context appears divided by the number of confidential documents in which only the confidential term appears. And $P_{nc}(key_{context}/key)$ is defined as the number of nonconfidential documents in which the confidential term with its context appears divided by the number of nonconfidential documents in which only the confidential term appears.

As mentioned above, we predefine the similarity threshold of minimum cosine measure TR_{min}, and iteratively expand to include more clusters. TR is referred to as the similarity threshold of cosine measure between the cluster with few non-confidential documents and its expanding cluster. After each iteration, we lower the value of TR at a certain rate μ which is predefined, namely $TR = \mu^* TR$. Unless TR is greater than TR_{min}, the non-confidential documents of the expanding cluster are included to recalculate the scores of context terms in original cluster. By including more adjacent clusters, we can accurately estimate which terms are most likely to indicate the confidentiality of the document.

By subtracting the probability of the appearance of each context term with confidential term in non-confidential documents from the probability of the appearance of them in confidential ones, the score of each context term is calculated, as shown in (2).

The reason for employing subtraction rather than division is to avoid large fluctuations in the values of the context terms. When employing division, even a single document can dramatically change the probabilities as only the documents including confidential terms are taken into account.

$$score+ = P_c\left(\frac{key_{context}}{key}\right) - P_{nc}\left(\frac{key_{context}}{key}\right) \quad (2)$$

We iteratively expand to include more clusters. After each iteration, we lower the value of TR until TR is less than TR_{min}, and the score of each context term is calculated, as shown in (3) in which n_{clu} denotes the number of clusters involved.

$$score+ = \frac{1}{n_{clu}}\left(P_c\left(\frac{key_{context}}{key}\right) - P_{nc}\left(\frac{key_{context}}{key}\right)\right) \quad (3)$$

After this phase, the set of context terms with their scores, $ContextT$, is obtained. For each confidential term, its context terms whose scores are positive are more likely to appear in confidential documents with the confidential term.

3.2.3. Create Graph Representation. After the operations described in previous section, confidential terms and their context can be easily represented as nodes and connected together according to their interrelation. As shown in Figure 2, for each cluster, a set of confidential terms and a set of its context terms are obtained after the training phase, and confidential terms and its context terms are represented as confidential nodes and context nodes respectively. Confidential nodes are connected together as long as there exists at least one common context node between them.

3.3. Pruning Nodes of Graph. Due to the calculation of confidential terms and their context terms are based on statistics scores, there might exist occasional case of a nonconfidential term with high score because of term abuse. In the pruning phase, we employ the method of term reduction in rough set theory to remove the redundancy nodes in graph.

With the information of confidential and nonconfidential documents, we evaluate the importance of nodes in graph for each cluster. A node in graph can be pruned only if the removal of the term represented by the node does not influence the results of identifying the confidential documents in this cluster. As shown in (4), $Impt(t_i)$ denotes the importance measure of term t_i which is represented as node i in graph. And $Pos_G(C)$ denotes the portion in confidential documents set C can be identified correctly by graph G. Similarly, $Pos_{\{G-n_i\}}(C)$ denotes the portion in confidential documents set C can be identified correctly by graph $\{G - n_i\}$, which means node n_i is removed from graph G. The detail of the pruning procedure for graph is presented in Algorithm 2.

$$Impt(t_i) = \frac{\left(\left|Pos_G(C)\right| - \left|Pos_{\{G-n_i\}}(C)\right|\right)}{|C|} \quad (4)$$

Input: C - Confidential documents set
N - Non-confidential documents set
CR - Clusters set, each with the centroid and corresponding graph
$ContextT$ - The set of context terms
Output: CT' - The updated set of confidential terms in clusters
$ContextT'$ - The updated set of context terms
(1) $irr\_terms \longleftarrow CT \cup ContextT$
(2) **for** (each t in irr_terms)
(3) **If** $(Impt(t) = (|Pos_T(C_{cr} \wedge N_{cr})| - |Pos_{T-t}(C_{cr} \wedge N_{cr})|)/|cr| < Impt\_threshold)$
(4) $irr\_terms = irr\_terms - t$
(5) **for** (each t in irr_terms)
(6) { $CT' \longleftarrow$ Update confidential terms in CT
(7) $ContextT' \longleftarrow$ Update context terms in $ContextT$ }

Algorithm 2: Pruning.

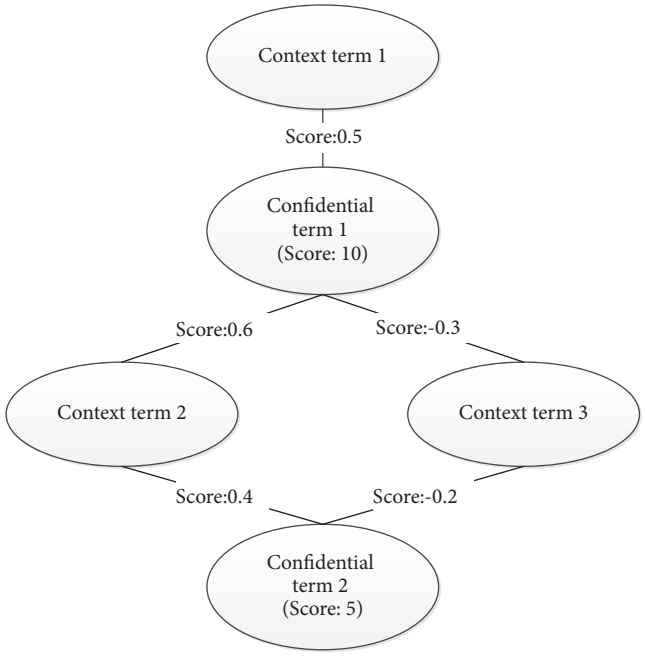


Figure 2: An example of confidential nodes connected through their context nodes.

3.4. Detection Phrase. Obviously, a confidential document without any modification is easy to be detected according to confidential terms. However, the confidential documents, which are rephrased or partitioned into portions and further concealed in different nonconfidential documents as most plagiarizers often do, can hardly be detected. Once the confidential contents detection fails, it is more likely to lead to data leakage or copyright infringement.

In the detection phase, we employ CBDLP model to deal with three scenarios that could possibly happen. The three different scenarios are described as follows:

(i) Each confidential document is detected as a whole.

(ii) Each confidential document is divided into portions and embedded in nonconfidential ones.

(iii) The confidential terms in confidential documents are rephrased completely.

The detection method we employed includes three steps as shown in Figure 3, which are described as follows:

(1) Classify the documents to be tested to the corresponding clusters.

(2) Identify the confidential terms and their context terms according to the graphs of the corresponding clusters.

(3) Calculate the confidentiality scores for the documents and draw the conclusion that whether a document is confidential or not.

Then, the security model, which combines the training phrase and the detection phrase, is shown in Figure 4.

4. Experiments

In this section, we evaluate the performance of CBDLP on Reuters-21578 dataset. As testing dataset, Reuters-21578 consists of 21578 pieces of news distributed by Reuters in 1987 which are saved in 22 files. Reuters-21578 dataset is manually classified as five categories, each of which can be subdivided into different number of subcategories. For example, the news of economy includes the inventories subset, gold subset, and money-supply subset.

4.1. Performance Experiments. In the experiments, we present the data leakage prevention method based on CBDLP model, and also present a modified model without pruning step which is represented as CBDLP-Pr. Since SVM has been proved to be an excellent classifier with high accuracy and CoBAn performs well in the scenario where confidential contents are embedded in nonconfidential documents or rephrased, we compare the performance of CBDLP, CBDLP-Pr, SVM, and CoBAn. We evaluate the performance of the methods in this paper with true positive rate (TPR) and false positive rate (FPR), and our goal is to maximize TPR and minimize FPR concurrently.

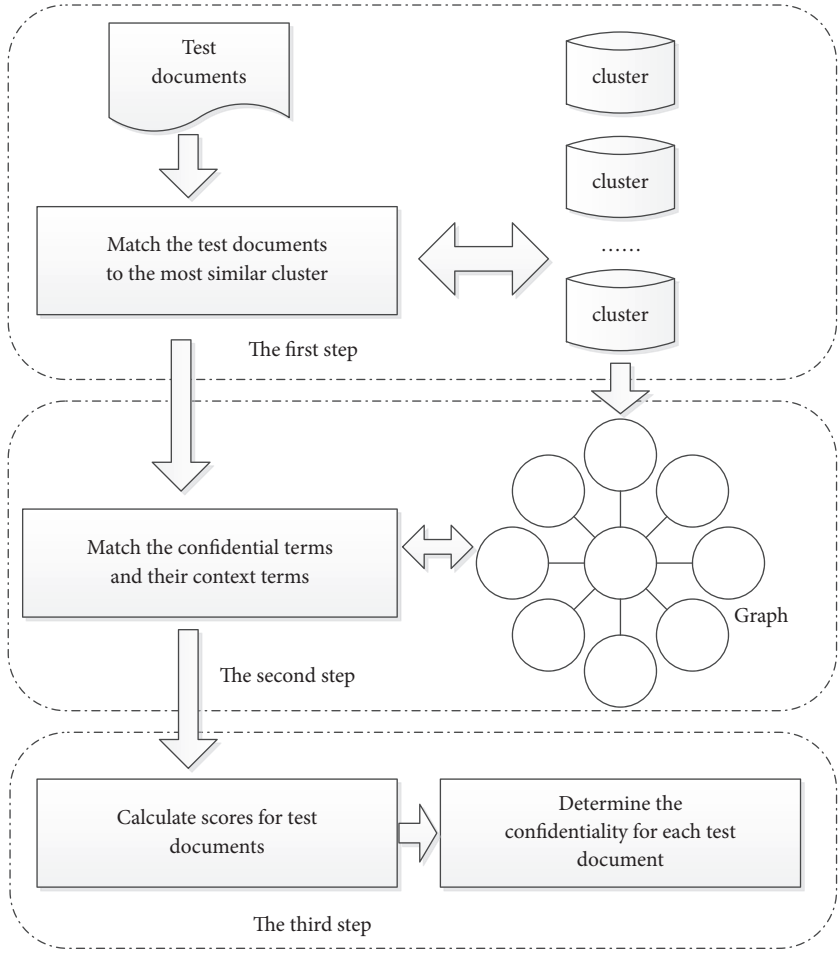


Figure 3: The detection of test documents.

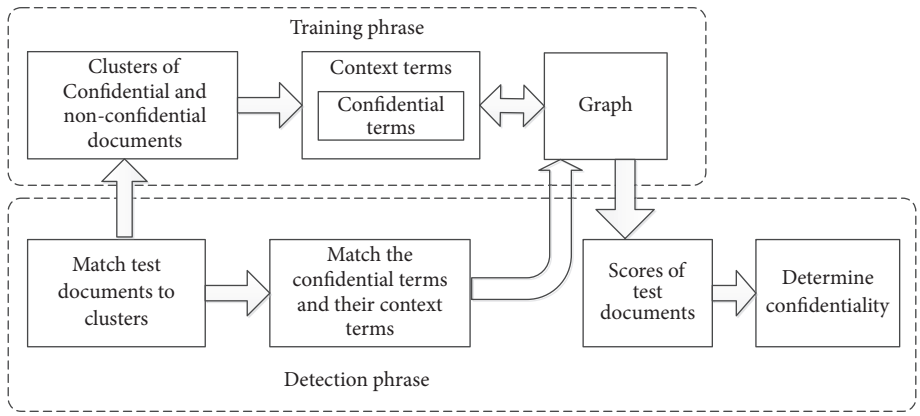


Figure 4: The security model.

We conduct experiments on the three scenarios which are described above. As for the first type of scenario, we select the news of "earn" as the carrier for confidential contents and mix them with the news from other economy subsets as training dataset and testing dataset separately. As for the second type of scenario, we extract the contents from the documents of "earn" subset and embed them in the documents from other subsets. The embedded portions are detected as confidential contents. As for the third type of scenario, we manually rephrase the contents in the documents of "earn" subset and embed them in the documents from other subsets.

4.1.1. Confidential Documents as a Whole. The experimental result of the first scenario is presented in Figure 5. As shown in

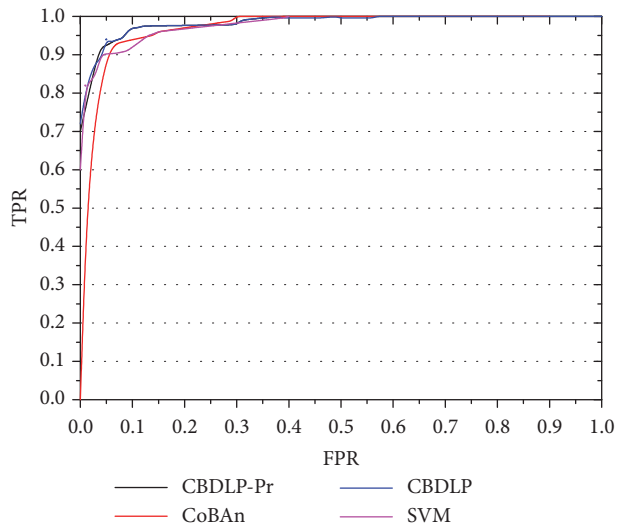


Figure 5: Performance of detecting confidential documents as a whole.

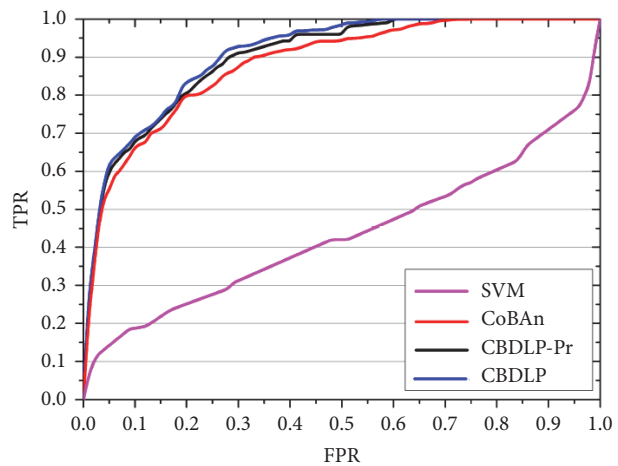


Figure 6: Performance of detecting the confidential contents embedded in nonconfidential documents.

Figure 5, when dealing with the scenario where confidential documents are considered as a whole, the performance of the four detection algorithms has no much difference. In spite of that, CBDLP and CBDLP-Pr still perform slightly better than CoBAn and SVM, which can be explained as that the performance of CoBAn is partly influenced by the limitation of k-means that it cannot deal with the clusters of various shapes effectively, and SVM only focuses on the confidential terms nevertheless ignores the context terms. In this scenario, since the documents containing confidential terms are explicitly detected as confidential documents, the performance of the four methods has no much difference.

4.1.2. Confidential Portions Embedded in Nonconfidential Documents. The result of the second scenario is presented in Figure 6. As shown in Figure 6, when dealing with the scenario where the confidential portions are embedded in nonconfidential documents, CBDLP, CBDLP-Pr, and CoBAn

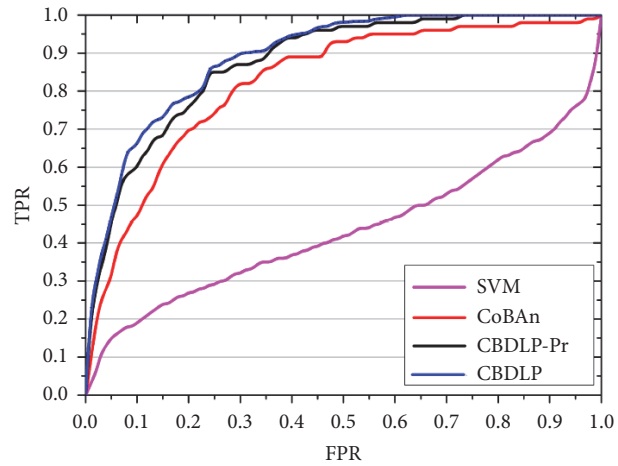


Figure 7: Performance of detecting the rephrased confidential contents embedded in nonconfidential documents.

perform better than SVM, which can be explained as that SVM is deceived by the scenario due to its statistics nature. As expected, the performance of CBDLP is slightly better than CBDLP-Pr and CoBAn due to its pruning step which removes the redundancy nodes in graph that might deteriorate the results of detection.

In this scenario, confidential portions are extracted from the documents defined as confidential and then embedded in the nonconfidential documents whose length are at least ten times larger than the extracted portions. Due to the statistical nature, most documents containing confidential portions are incorrectly detected as nonconfidential by SVM, which result in dramatic decline in the accuracy of SVM. Other than SVM, CBDLP, CBDLP-Pr, and CoBAn take the confidential terms together with their context into account, and most nonconfidential documents containing embedded confidential portions are detected as confidential.

4.1.3. Rephrased Confidential Contents in Nonconfidential Documents. The result of the third scenario is presented in Figure 7. As shown in Figure 7, when dealing with the scenario where the confidential contents are rephrased and embedded in nonconfidential documents, the performance of SVM deteriorates considerably due to its statistics nature. Since the rephrased contents do not deviate much from its original meaning, CBDLP, CBDLP-Pr, and CoBAn perform well. In addition, the performance of CBDLP is better than CBDLP-Pr and CoBAn due to its pruning step which removes the redundancy nodes in graph.

In this scenario, the rephrased confidential terms are embedded in nonconfidential documents which confuse SVM greatly, and most documents containing rephrased confidential contents are incorrectly detected as nonconfidential. Other than SVM, with the context of confidential terms taken into account, CoBAn detects most documents containing confidential contents; however, the accuracy of CoBAn is partially influenced by the cluster's terms graph which depends on the quality of clusters generated by k-means. As a result, CBDLP clusters documents with DBSCAN which

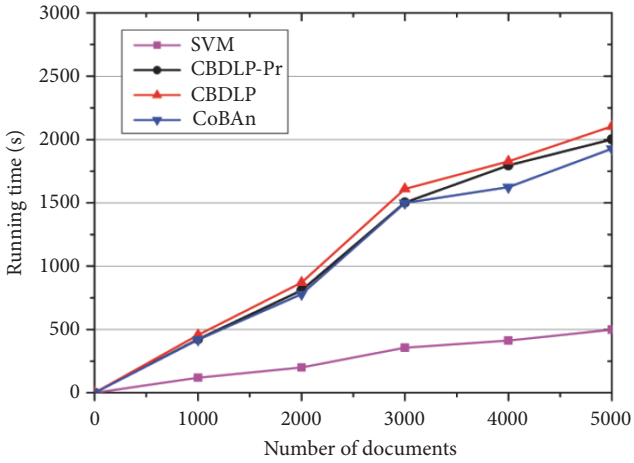


Figure 8: The scores of confidential documents and nonconfidential documents.

improves the quality of clusters and the cluster's terms graph; meanwhile, the pruning method removes the redundancy nodes in graph and further improves the performance of CBDLP.

4.2. Running Time Comparisons. In this experiment, we mixed non-confidential documents together with the three type of confidential documents, which are the whole confidential documents, the confidential contents embedded in nonconfidential documents and the rephrased confidential contents embedded in non-confidential documents. The experiment is conducted by using 10 fold cross validation. To Compare the running time of CBDLP, CBDLP-Pr, CoBAn, and SVM, we conduct the experiment on the datasets of different size. The result is as shown in Figure 8, the running time of training phase and testing phase are exhibited as line graph in which the running time of CBDLP, CBDLP-Pr, CoBan and SVM increase as more documents are added to the dataset. Although the additional steps of CBDLP, CBDLP-Pr, and CoBan result in more running time than SVM needs, their running time is still an order of magnitude; more than that, CBDLP performs much better than SVM does.

5. Conclusion and Future Work

In this paper, we present a new method for data leakage Prevention based on CBDLP model, which has the following advantages:

(1) It clusters the documents with DBSCAN and cosine measure which have been verified to be effective.

(2) It represents confidential terms and their context terms in graph.

(3) It presents a pruning method based on the attribute reduction method of rough set theory.

Up to now, some designated commercial DLP solutions can reduce the risk of most accidental leakage; however, they cannot provide sufficient protection against intentional leakage. And the other DLP solutions, such as firewalls, IDS, antimalware software, and management policies, which can provide assistance in detection intrusion or malicious software and enforce policies to protect data, still do not prevent intentional leaks perfectly. To the best of our knowledge, there might be two main future research topics on DLP, data leakage from mobile devices and accidental data leakage by insider.

Since accidental data leakage may be part of a larger attack in which their role will be mainly to activate an advanced persistent threat inside the organization, it is expected to continue to be one of the most challenging research topics. And our future work will focus on accidental data leakage in two directions. First, try to improve the efficiency and effectiveness of CBDLP on confidential contents detection. Second, adjust the model dynamically according to the changes of training dataset.

Data Availability

All data generated or analysed during this study are included in this published article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research is supported by the National Natural Science Foundation of China under Grant nos. 61871140 and 61572153.

References

[1] A. Goyal, F. Bonchi, and V. S. Lakshmanan, "On minimizing budget and time in influence propagation over social networks," *Social Network Analysis and Mining*, pp. 1–14, 2012.

[2] C. Aggarwal, *Social Network Data Analytics*, Springer, Berlin, Germany, 2011.

[3] "Information week global security survey," Information Week, 2004.

[4] D. Alassi and R. Alhajj, "Effectiveness of template detection on noise reduction and websites summarization," *Information Sciences*, vol. 219, pp. 41–72, 2013.

[5] D. Holmes, *Using language models for information retrieval [Ph.D. Thesis]*, Center for telematies and information technology, University of Twente, 2001.

[6] R. Böhme, "Security metrics and security investment models," in *Proceedings of the 5th international conference on advances in information and computer security*, 2010.

[7] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, Boston, Mass, USA, 1990.

[8] S. M. Katz, "Estimation of probabilities from sparse data for the language model component of a speech recognizer," *IEEE Transactions on Signal Processing*, vol. 35, no. 3, pp. 400-401, 1987.

[9] R. Jin, L. Si, A. G. Hauptmann, and J. Callan, "Language model for IR using collection information," in *Proceedings of the the 25th annual international ACM SIGIR conference*, pp. 419-420, 2002.

[10] W. W. Cohen, "Learning rules that classify e-mail," in *Proceedings of the AAAI Spring Symposium on Machine Learning in Information Access*, pp. 18–25, 1996.

[11] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[12] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *Journal of the Association for Information Science and Technology*, vol. 41, no. 6, pp. 391–407, 1990.

[13] C. Ordonez, "Clustering binary data streams with K-means," in *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, DMKD '03*, pp. 12–19, USA, June 2003.

[14] B. Babcock, M. Datar, R. Motwani, and L. O'Callaghan, "Maintaining variance and k-medians over data stream windows," in *Proceedings of the Twenty second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2003*, pp. 234–243, June 2003.

[15] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: a new data clustering algorithm and its applications," *Data Mining and Knowledge Discovery*, vol. 1, no. 2, pp. 141–182, 1997.

[16] S. Guha, R. Rastogi, and K. Shim, "Cure: an efficient clustering algorithm for large databases," in *Proceedings of 1998 ACM SIGMOD International Conference Management of Data*, pp. 73–84, 1998.

[17] J. W. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2006.

[18] A. Hinneburg and D. A. Keim, "Optimal grid-clustering: towards breaking the curse of dimensionality in high-dimensional clustering," in *Proceedings of the 25th VLDB Conference*, pp. 506–517, 1999.

[19] W. Wang, J. Yang, and R. Muntz, "Sting: a statistical information grid approach to spatial data mining," in *Proceedings of the 23rd VLDB Conference*, pp. 186–195, 1997.

[20] R. Agrawal, J. Gehrke, and D. Gunopulos, "Automatic subspace clustering of high dimensional data for data mining applications," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 94–105, 1998.

[21] A. McCallum, K. Nigam, and L. H. Ungar, "Efficient clustering of high-dimensional data sets with application to reference matching," in *Proceedings of the KDD 2000*, pp. 169–178, ACM, New York, NY, USA, 2000.

[22] R. Hyde, P. Angelov, and A. R. MacKenzie, "Fully online clustering of evolving data streams into arbitrarily shaped clusters," *Information Sciences*, vol. 382-383, pp. 96–114, 2017.

[23] F. Jiang, Y. Fu, B. B. Gupta et al., "Deep learning based multichannel intelligent attack detection for data security," *IEEE Transactions on Sustainable Computing*, 2018.

[24] R. Agrawal and R. Srikant, "Privacy-preserving data mining," in *Proceedings of the the 2000 ACM SIGMOD international conference*, pp. 439–450, May 2000.

[25] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.

[26] Salton, "Automatic text processing: the transformation, analysis and retrieval of information by computer," Tech. Rep., Addison-Wesley Inc., 1989.

[27] M. N. Islam, M. Seera, and C. K. Loo, "A robust incremental clustering-based facial feature tracking," *Applied Soft Computing*, vol. 53, pp. 34–44, 2017.

[28] A. Shabtai and Y. Elovici, *A Survey of Data Leakage Detection and Prevention Solutions*, Springer, Berlin, Germany, 2012.

[29] S. R. Kalidindi, S. R. Niezgoda, G. Landi, S. Vachhani, and T. Fast, "A novel framework for building materials knowledge systems," *Computers, Materials and Continua*, vol. 17, no. 2, pp. 103–125, 2010.

[30] A. J. Wang, "Information security models and metrics," in *Proceedings of the 43rd annual southeast regional conference on ACMSE43*, pp. 178–184, 2005.

[31] P. Mitra, C. A. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 301–312, 2002.

[32] M. Dash, K. Choi, P. Scheuermann, and H. Liu, "Feature selection for clustering - A filter solution," in *Proceedings of the 2nd IEEE International Conference on Data Mining, ICDM '02*, pp. 115–122, December 2002.

[33] C.-S. Liu, "An analytical method for computing the one-dimensional backward wave problem," *Computers, Materials and Continua*, vol. 13, no. 3, pp. 219–234, 2010.

[34] G. Katz, Y. Elovici, and B. Shapira, "Coban a context based model for data leakage prevention," *Information Sciences*, vol. 262, pp. 137–158, 2014.

[35] X. Huang, Y. Lu, D. Li, and M. Ma, "A novel mechanism for fast detection of transformed data leakage," *IEEE Access*, vol. 1, pp. 1–11, 2018.

[36] M. Porter, "The porter stemming algorithm," 2006.

[37] F. Pacheco, M. Cerrada, R.-V. Sánchez, D. Cabrera, C. Li, and J. Valente de Oliveira, "Attribute clustering using rough set theory for feature selection in fault severity classification of rotating machinery," *Expert Systems with Applications*, vol. 71, pp. 69–86, 2017.

[38] "Information week global security survey," Information Week, 2004.

[39] C. M. Praba, "A technical review on data leakage detection and prevention approaches," *Journal of Network Communications and Emerging Technologies (JNCET)*, 2017.

[40] F. Ullah, M. Edwards, R. Ramdhany, R. Chitchyan, M. A. Babar, and A. Rashid, "Data exfiltration: A review of external attack vectors and countermeasures," *Journal of Network and Computer Applications*, 2017.

[41] K. Thomas, F. Li, A. Zand et al., "Data Breaches, phishing, or malware? understanding the risks of stolen credentials," in *Proceedings of the 24th ACM SIGSAC Conference on Computer and Communications Security, CCS '17*, pp. 1421–1434, November 2017.

[42] J. DeBlasio, S. Savage, G. M. Voelker, and A. C. Snoeren, "Tripwire: Inferring internet site compromise," in *Proceedings of the IMC '17*, pp. 1–14, 2017.

[43] W. Xu, S. Xiang, and V. Sachnev, "A cryptograph domain image retrieval method based on paillier homomorphic block encryption," *Computers Materials and Continua*, pp. 1–11, 2018.

[44] J. Cui, Y. Zhang, Z. Cai, A. Liu, and Y. Li, "Securing display path for security-sensitive applications on mobile devices," *Computers, Materials and Continua*, vol. 55, no. 1, pp. 17–35, 2018.

[45] D. Wang, W. Li, and P. Wang, "Measuring two-factor authentication schemes for real-time data access in industrial wireless

sensor networks," *IEEE Transactions on Industrial Informatics*, pp. 1–12, 2018.

[46] D. Wang, H. Cheng, P. Wang, J. Yan, and X. Huang, "A security analysis of honeywords," in *Proceedings of the Network and Distributed Systems Security (NDSS) Symposium*, pp. 18–21, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 3048697, 13 pages
https://doi.org/10.1155/2018/3048697

Research Article

An Enhanced User Authentication Protocol Based on Elliptic Curve Cryptosystem in Cloud Computing Environment

Chenyu Wang[1], **Ke Ding**[2], **Bin Li**[3], **Yiming Zhao**[3], **Guoai Xu**[1], **Yanhui Guo**[1] and **Ping Wang**[2,3]

[1]*School of Cyberspace Security, Beijing University of Posts and Telecommunications, China*
[2]*School of EECS, Peking University, China*
[3]*School of Software and Microelectronics, Peking University, China*

Correspondence should be addressed to Guoai Xu; xga@bupt.edu.cn

Received 13 April 2018; Revised 24 June 2018; Accepted 12 July 2018; Published 1 October 2018

Academic Editor: Jian Shen

With the popularity of cloud computing, information security issues in the cloud environment are becoming more and more prominent. As the first line of defense to ensure cloud computing security, user authentication has attracted extensive attention. Though considerable efforts have been paid for a secure and practical authentication scheme in cloud computing environment, most attempts ended in failure. The design of a secure and efficient user authentication scheme for cloud computing remains a challenge on the one hand and user's smart card or mobile devices are of limited resource; on the other hand, with the combination of cloud computing and the Internet of Things, applications in cloud environments often need to meet various security requirements and are vulnerable to more attacks. In 2018, Amin et al. proposed an enhanced user authentication scheme in cloud computing, hoping to overcome the identified security flaws of two previous schemes. However, after a scrutinization of their scheme, we revealed that it still suffers from the same attacks (such as no user anonymity, no forward secrecy, and being vulnerable to offline dictionary attack) as the two schemes they compromised. Consequently, we take the scheme of Amin et al. (2018) as a study case, we discussed the inherent reason and the corresponding solutions to authentication schemes for cloud computing environment in detail. Next, we not only proposed an enhanced secure and efficient scheme, but also explained the design rationales for a secure cloud environment protocol. Finally, we applied BAN logic and heuristic analysis to show the security of the protocol and compared our scheme with related schemes. The results manifest the superiority of our scheme.

1. Introduction

With the development of IT technology, cloud computing has become one of the hottest research directions in recent years. As a new type of service, cloud computing is rapidly integrated into our daily lives with its high scalability, high service efficiency, and low-cost charge [1]. It fundamentally changed the traditional model of service providers providing services and consumers' access to resources: as a service provider (such as Google, Microsoft, Amazon) of cloud computing, it effectively improves the utilization of resources by centralizing the demands; consumers not only gain the convenience of using resources, but also reduce the using cost through paying on demand. Therefore, more and more big firms build their own cloud platforms and provide services, including Google App Engine, Amazon Web services, and IBM SmartCloud [2]. Furthermore, both the individuals and small companies enjoy the benefits of cloud services. Generally speaking, there are three kinds of cloud services: (1) Iaas, Infrastructure as a Service, which means providing user with the infrastructure such as storage and networks to use; (2) Paas, Platform as a Service, which means providing user with the platform to develop various applications; (3) Saas, Software as a Service, which means providing user with software applications [3, 4].

However, with the increasing popularity of the cloud services, the security issues have become more prominent, how to protect user privacy and restrict data from being

illegally accessed has become a challenging problem and research hotspot. The first step to solve these issues is user authentication which can verify the authenticity of communication participants. A secure user authentication scheme will firstly verify the authenticity of the user when he/she applies to access the cloud data; then to prevent a malicious cloud server trick users, the validity of the cloud server should be checked; once confirming the identity of the user and the cloud server, a session key will be established to encrypt the communication messages.

Generally speaking, there are three ways to authenticate a user, which are based on the following: (1) what you know (such as the password); (2) what you have (such as the smart card); (3) who you are (such as the biometric characteristic: fingerprint and iris). Due to its simplicity and practicality, passwords have been used more widely. While a password-based authentication protocol has natural flaws, it cannot resist against offline dictionary guessing attacks. Consequently, as a factor to help enhance security, the smart card gets used [5–9]. A scheme combined two factors (such as the password and the smart card) is called two-factor user authentication scheme. The participants of the two-factor authentication scheme in cloud computing environment involves a user, a cloud server and a register authority. Note that among the three participants, only the register authority is trusted. At first, the user and the cloud server register to the register authority, respectively. Then the cloud server will send the user a smart card with some sensitive information and negotiate a shared secret parameter with the cloud server. Later on, when the user initiates an access request to the cloud server in login phase, the three participants will authenticate themselves to each other. If they all are authenticated, the user will be allowed to access the cloud server.

Motivations. In 2013, Yang et al. [10] devoted to design a secure authentication scheme for cloud computing environment, while their scheme is vulnerable to dictionary attack. Then Yang et al. [11] proposed a new scheme; unfortunately, Chen et al. [12] then showed this scheme is not secure to insider attack and impersonation attack and proposed a new version which once again is broken by Wang et al. [13]. Wang et al. [13] pointed out that Chen et al.'s scheme [12] is subject to offline dictionary attack and impersonation attack. Most recently, Amin et al. [3] identified the security weaknesses in the schemes of Xue et al. [14] and Chuang et al. [15] by revealing the two schemes fail to provide user anonymity and forward secrecy while being not able to resist against offline password guessing attack and so on. Therefore, they designed a new scheme that claims to overcome the security flaws of the two schemes and be secure to various attacks. However, after a scrutinization of Amin et al.'s scheme, we found their scheme still cannot overcome their identified security threats.

In these years, considerable efforts have been paid for a secure and practical authentication scheme in cloud computing environment, some typical schemes including [16–19], yet most of them are found having security flaws more or less. Designing of a secure authentication scheme for cloud computing environment is still a challenge. With the widespread use of cloud computing, the potential security threats will lead to greater harm. This unsatisfactory situation motivates us to explore the inherent reasons of the failure in those schemes, find the basic method to fix the security flaws, and design a robust and efficient user authentication protocol for cloud computing environment.

Our contributions. Amin et al.'s scheme [3] is a very typical scheme which suffers from the common attacks, while the scheme's structure is widely accepted. So we take Amin et al.'s protocol as a study case to elaborate the common issues (and its corresponding solutions) in most authentication schemes and provide rationales for designing a secure cloud environment protocol. In addition, based on the analysis, we design a secure authentication protocol. In a short, our contributions can be summarized as follows:

(1) We demonstrated that Amin et al.'s scheme [3] fails to achieve user anonymity and forward secrecy while being not able to resist against offline dictionary attack.

(2) We discussed the inherent reasons of the identified flaws and its corresponding solutions; furthermore, we realized the way of deploying a public key algorithm rightly is challenging. Therefore, we showed the essential points for deploying public key algorithms.

(3) We improved Amin et al.'s scheme from security and effectiveness two aspects, proved the security of our scheme via BAN logic and heuristic analysis and, finally, compared our scheme with other related schemes. The results show that our scheme is more suitable for cloud computing environment.

The remainder of this paper is organized as follows: Section 2 sketches complexity assumptions and extends adversary model; then, Amin et al.'s scheme is reviewed and analyzed in Section 3; in Section 4, we propose a secure scheme and elaborate on design rationales; Section 5 proves the security of our scheme; in Section 6, we compare our scheme with other related schemes; finally, the conclusion is drawn in Section 7.

2. Preliminary

This section introduces the preliminary of the whole paper, including complexity assumptions in designing a scheme and some notations and abbreviations.

2.1. Computational Problems. Given two large primes p and q, let $\mathbb{F}_p$ be a finite field, $E/\mathbb{F}_p$ be an elliptic curve over $\mathbb{F}_p$, and $\mathbb{G}$ be a q-order subgroup of $E/\mathbb{F}_p$. For $\alpha, \beta \in Z_p^*$ and a point P in $\mathbb{G}$, we can define the discrete logarithm problem as follows:

(1) Elliptic curve discrete logarithm (ECDL) problem: given $(P, \alpha P)$, it is impossible to compute α within polynomial time.

(2) Elliptic curve computational Diffie-Hellman (ECCDH) problem: given $(\alpha P, \beta P)$, it is impossible to compute $\alpha\beta P$ within polynomial time.

2.2. Adversary Models. Understanding the adversary models is the most basic step to design and analyze a protocol. In

Table 1: Notations and abbreviations.

Symbol	Description
U_i	i^{th} user
CS_j	j^{th} cloud server
RA	the register authority
$\mathscr{A}$	the adversary
x	the long term secret key of RA
y	the secret key of RA
ID_i	identity of U_i
PW_i	password of U_i
SID_j	identity of CS_j
$Skey_j$	shared key between G_k and CS_j
$\oplus$	bitwise XOR operation
$\|$	concatenation operation
$h(\cdot)$	one-way hash function
$\longrightarrow$	a common channel
$\Longrightarrow$	a secure channel

2015, Wang et al. [20] proposed the capabilities of adversary in distributed systems: (1) exhaust passwords and identities; (2) learn ID_i when evaluating security strength; (3) control of the open communication channel; (4) learn PW_i or extract information in the smart card; (5) acquire previous session keys; (6) know the long-term secret key x when considering forward secrecy. As both the distributed systems and the cloud computing systems have similar network environment, their adversary models are also similar too. Therefore, we adopt Wang et al.'s adversary models [20] which have been accepted by various schemes [21–23].

2.3. Notations and Abbreviations. As shown in Table 1, we summarize the notations and abbreviations used in this paper.

3. Cryptanalysis of Amin et al.'s Scheme

After identifying the security pitfalls in other two user authentication schemes, Amin et al. [3] attempted to design a new light weight protocol in cloud computing environment. After analyzing their scheme using AVISPA tool, they claimed the new scheme achieves forward security while being resistant to various attacks. However, this section will show that, under the assumptions on adversary capabilities in Section 2.2, their scheme cannot provide forward security while being subject to two kinds of offline dictionary attacks [27] and so on. Thus their scheme is not a truly two-factor scheme. To address these issues, this section first reviews the scheme of Amin et al. and then analyzes Amin et al.'s scheme [3].

3.1. Review of Amin et al.'s Scheme. This section briefly reviews the scheme of Amin et al. [3]; their scheme consists of five phases. As the password change phase and identity update phase have little relevance, we omit them. Furthermore, we adjust some symbols of their scheme for the ease of reading and the unity of the paper.

Registration Phase

(1) Cloud Server Registration Phase

Step 1. $CS_j \Longrightarrow RA$: $\{SID_j, d\}$. d is a random number.

Step 2. $RA \Longrightarrow CS_j$: $\{Skey_j\}$. RA computes $PSID_j = h(SID_j \| d)$, $Skey_j = h(PSID_j \| y)$.

Step 3. CS_j keeps $\{Skey_j, d\}$.

(2) User Registration Phase

Step 1. $U_i \Longrightarrow RA$: $\{A_i, PID_i\}$. $A_i = h(PW_i \| b_1)$, $PID_i = h(ID_i \| b_2)$, $bb_i = b_2 \oplus A_i$ where b_1, b_2 are two random number chosen by U_i.

Step 2. $RA \Longrightarrow U_i$: smart card $\{C_i, E_i, bb_i, DP, h(\cdot)\}$. RA calculates $C_i = h(A_i \| PID_i)$, $D_i = h(PID_i \| x)$, $E_i = D_i \oplus A_i$.

Step 3. U_i inputs DP, bb_i into the card where $DP = h(ID_i \| PW_i) \oplus b_1$.

Login and Authentication Phase

Step 1. $U_i \longrightarrow CS_j$: $\{G_i, F_i, Z_i, PID_i, TS_i\}$. After U_i inputs ID_i and PW_i, the card computes $b_1^* = DP \oplus h(ID_i^* \| PW_i^*)$, $A_i^* = h(PW_i^* \| b_1^*)$, $b_2^* = bb_i \oplus A_i^*$, $PID_i^* = h(ID_i^* \| b_2^*)$, $C_i^* = h(A_i^* \| PID_i^*)$. If $(C_i^* == C_i)$, the card produces a random number N_i (128 bit) and computes $D_i = E_i \oplus A_i$, $G_i = h(PID_i \| SID_j \| N_i \| TS_i \| D_i)$, $F_i = D_i \oplus N_i$, $Z_i = SID_j \oplus h(D_i \| N_i)$.

Step 2. $CS_j \longrightarrow RA$: $\{J_i, K_i, PSID_j, G_i, F_i, Z_i, PID_i, TS_i, TS_j\}$. CS_j first checks $(TS_j - TS_i < \Delta T)$, then selects a 128 bit random number N_j, and computes $J_i = Skey_j \oplus N_j$, $K_i = h(N_j \| Skey_j \| G_i \| TS_j)$.

Step 3. $RA \longrightarrow CS_j$: $\{P_{cs}, R_{cs}, Q_{cs}, V_{cs}\}$. RA checks $(TS_{cs} - TS_j < \Delta T^*)$ and computes $D_i = h(PID_i \parallel x)$, $N_i^* = F_i \oplus D_i$, $SID_j^* = Z_i \oplus h(D_i \parallel N_i^*)$, $G_i^* = h(PID_i \parallel SID_j^* \parallel N_i^* \parallel D_i \parallel TS_i)$.

If $(G_i^* == G_i)$, compute $Skey_j^* = h(PSID_j \parallel y)$, $N_j^* = Skey_j^* \oplus J_i$, $K_i^* = h(Skey_j^* \parallel N_j^* \parallel G_i \parallel TS_j)$.

If $(K_i^* == K_i)$, compute $P_{cs} = N_j \oplus N_{cs} \oplus h(N_i \parallel D_i)$, $R_{cs} = N_i \oplus N_{cs} \oplus h(Skey_j^* \parallel N_j^*)$, $SK_{cs} = h(N_i \oplus N_j \oplus N_{cs})$, $Q_{cs} = h((N_j \oplus N_{cs}) \parallel SK_{cs})$, $V_{cs} = h((N_i \parallel N_{cs}) \parallel SK_{cs})$

Step 4. $CS_j \longrightarrow U_i$: $\{P_{cs}, Q_{cs}\}$. CS_j computes $W_j = h(Skey_j \parallel N_j)$, $N_i \oplus N_{cs} = R_{cs} \oplus W_j$, $SK_j = h(N_i \oplus N_{cs} \oplus N_j)$, $V_{cs}^* = h((N_i \oplus N_{cs}) \parallel SK_j)$. If $(V_{cs}^* \neq V_{cs})$, send $\langle P_{cs}, Q_{cs} \rangle$ to U_i.

Step 5. U_i computes $L_i = h(N_i \parallel D_i)$, $N_j \oplus N_{cs} = P_{cs} \oplus L_i$, $SK_i = h(N_j \oplus N_{cs} \oplus N_i)$, $Q_{cs}^* = h((N_j \oplus N_{cs}) \parallel SK_i)$. If $(Q_{cs}^* == Q_{cs})$, the whole authentication phase finishes successfully.

3.2. Cryptanalysis of Amin et al.'s Scheme. Amin et al.'s scheme [3] does not point out the adversarial model, while, according to their attack on the scheme of Xue et al. [14] and Chuang et al. [15], we can infer their adversary model which is included into our model (see Section 2.2). Although Amin et al.'s scheme [3] provides many admirable features, such as changing password locally and high efficiency, it still suffers from various attacks like most authentication protocol in cloud computing environment. Therefore, their scheme is a typical case to show the security threat in cloud environment. Through Amin et al.'s scheme, we can get insight into the inherent reasons of the failure in other authentication protocols for cloud and, based on it, learn to design a secure one. In brief, this section, on one hand, demonstrates that Amin et al.'s scheme [3] is vulnerable to various attacks and, on the other hand, indicates the failure reasons of their scheme.

Off-Line Dictionary Attack I. ***(i) The adversary's capability***: it obtains the message $\{C_i, E_i, bb_i, DP\}$ in U_i's smart card.

(ii) The attack steps:the steps are as follows.

Step 1. Guess PW_i to be PW_i^*, ID_i to be ID_i^*. Note that, it is quite realistic for $\mathcal{A}$ to obtain the password and identity simultaneously, because their spaces are limited [28].

Step 2. Compute $b_1^* = DP \oplus h(ID_i^* \parallel PW_i^*)$.

Step 3. Compute $A_i^* = h(PW_i^* \parallel b_1^*)$.

Step 4. Compute $b_2^* = bb_i \oplus A_i^*$.

Step 5. Compute $PID_i^* = h(ID_i^* \parallel b_2^*)$.

Step 6. Compute $C_i^* = h(A_i^* \parallel PID_i^*)$.

Step 7. Verify the correctness of PW_i and ID_i by checking if $(C_i^* \stackrel{?}{=} C_i)$.

Step 8. Repeat steps 1 ~ 6 until the correct values of PW_i and ID_i are found.

(iii) The time complexity: $\mathcal{O}(|\mathcal{D}_{pw}| * |\mathcal{D}_{id}| * 4T_H)$, where T_H is the time of hash-function.

Remark. Generally speaking, achieving two-factor security is the most essential requirement of a two-factor authentication protocol; that is, any one of the factors being broken will not trigger the security of another factor, which in turn threatens the entire system. In recent years, many protocols have tried to propose a secure two-factor security protocol, but most have failed. It was not until the work of Ma et al. [29] and Wang et al. [20, 30] did such a stagnant situation completely changed. In 2012, Ma et al. [29] pointed out that public key algorithm is necessary to design a secure two-factor authentication scheme; in 2015, Wang et al. [20] found that there is a conflict between changing password locally and resisting against smart card loss attack under the current technique; therefore, Wang et al. [30] put forward a way of "honeywords"+"fuzzy-verifier" to solve the conflict; in 2016, Wang et al. [27] further pointed out that there are two offline dictionary attacks and then combined with the results of [29, 30] and matched the corresponding solutions for each attack.

In this paper, we follow the classification method of Wang et al. [27] and demonstrate that Amin et al.'s scheme [3] cannot resist against the two kinds of dictionary attack. Looking back at the above attack process, we can find that the key to the problem is that $\mathcal{A}$ can find the verification value C_i to check the correctness of the guessed result. According to Wang et al. [30], this issue can be settled with the integration of "fuzzy-verifier" and "honeywords": let $C_i = h(A_i \parallel PID_i \parallel D_i) \bmod n_0$ where n_0 is a integer $(2^4 \leq n_0 \leq 2^8)$. The detailed explanation on this method can be found in Section IV of [30] or Section 5.2 of this paper.

Off-Line Dictionary Attack II. ***(i) The adversary's capability***: (1) it eavesdrops on one of U_i's login requests $\{G_i, F_i, Z_i, PID_i, TS_i\}$ and (2) obtains the message $\{C_i, E_i, bb_i, DP\}$ in U_i's smart card.

(ii) The attack steps:the steps are as follows.

Step 1. Guess PW_i to be PW_i^*, ID_i to be ID_i^*.

Step 2. Compute $b_1^* = DP \oplus h(ID_i^* \parallel PW_i^*)$.

Step 3. Compute $A_i^* = h(PW_i^* \parallel b_1^*)$.

Step 4. Compute $D_i^* = E_i \oplus A_i^*$.

Step 5. Compute $N_i^* = F_i \oplus D_i^*$.

Step 6. Compute $G_i^* = h(PID_i \parallel SID_j \parallel N_i^* \parallel TS_i \parallel D_i^*)$. Note that Amin et al. [3] view SID_j as a secret only known to the legitimate user. However, it not practical: $\mathcal{A}$ at least can register as a legitimate user to get SID_j.

Step 7. Verify the correctness of PW_i and ID_i by checking if $(G_i^* == G_i)$.

Step 8. Repeat steps 1 ~ 6 until the correct values of PW_i and ID_i are found.

(iii) The time complexity: $\mathcal{O}(|\mathcal{D}_{pw}| * |\mathcal{D}_{id}| * 3T_H)$.

Remark. In this attack, the pivotal parameter is G_i. To adversary $\mathcal{A}$, the only challenge in computing G_i is the value of D_i which can be derived from (ID_i, PW_i), so once $\mathcal{A}$ guesses the value of (ID_i, PW_i), he/she can check the correctness of them via G_i. Now considering a situation where G_i consists of the secret shared D_i and an another nonpublic dynamic parameter which should not be derived from (ID_i, PW_i). In this situation, $\mathcal{A}$ cannot use G_i to check the guessed value anymore, since there is another uncertain parameter besides D_i. Consequently, constructing such a dynamic parameter which is known to U_i and RA is our critical step to address this attack. Taking into account the fact that Ma et al.'s emphasizes [29] on the necessity of lightweight public-key algorithm in designing a secure authentication protocol, we then apply a lightweight public-key algorithm to construct such a dynamic parameter. In addition, our specific ideas on solving this attack are shown in Section 4.

User anonymity: these days user anonymity has become one of the security issues that people are widely concerned about, especially in the case of cloud computing and the Internet of Things that involve massive data. The adversary can acquire people's sensitive personal information via various ways including analyzing the session transcript in the open channel when the services are accessed [31]. Moreover, with the development of the technology, the adversary may even trace users' movement and learn the location of their home or company, which triggers a huge potential threat [32]. Under these circumstances, user anonymity is a pivotal attribute of the authentication scheme to protect user privacy.

Generally speaking, user anonymity covers two aspects [31]: (1) user identity protection; (2) user untraceability. The former requires the scheme does not expose users' identity; and the latter prevents an adversary from linking the session transcripts to a specific user or distinguishes the sessions sent by different users. This definition on user anonymity is widely applied in most authentication schemes [8, 24, 33, 34]. Unfortunately, in Amin et al.'s scheme, the parameter PID_i that identifies the user identity is a static value exposed in the insure channel, which means the adversary can trace U_i via PID_i. Consequently, Amin et al.'s scheme cannot provide user anonymity. As we can see, one of the keys to achieve user anonymity is concealing the real identity with a dynamic parameter. The way to implement this is called dynamic-ID technique [20]. According to Wang et al.'s suggestion [31], we can employ the dynamic-ID technique to protect user anonymity via applying a lightweight public-key algorithm to the authentication schemes as described in Section 4.

Forward Secrecy. ***(i) The adversary's capability***: (1) it eavesdrops on $\{F_i, PID_i\}$ and $\{P_{cs}\}$ and (2) obtains the long-term key x.

(ii) The attack steps:the steps are as follows.

Step 1. Compute $D_i^* = h(PID_i \parallel x)$.

Step 2. Compute $N_i^* = F_i \oplus D_i^*$.

Step 3. Compute $L_i^* = h(N_i^* \parallel D_i^*)$.

Step 4. Compute $(N_j \oplus N_{cs})^* = L_i^* \oplus P_{cs}$.

Step 5. Compute $SK = h((N_j \oplus N_{cs})^*) \oplus N_i^*$.

(iii) The time complexity: $O(|\mathcal{D}_{pw}| * |\mathcal{D}_{id}| * 3T_H)$.

Remark. In Amin et al.' scheme [3], there are two secret keys (x, y) in the register authority. We have shown that the leakage of the long-term key x will lead to the exposure of previous sessions key. In the following attack, we can see that the leakage of y leading to the same question too. As a result, the use of two system parameters is of little significance but consumes resource.

(i) The adversary's capability: (1) it eavesdrops on $\{J_i, PSID_j\}$ and $\{R_{cs}\}$ and (2) obtains the secret key y.

(ii) The attack steps:the steps are as follows.

Step 1. Compute $Skey_j^* = h(PID_j \parallel y)$.

Step 2. Compute $N_j^* = J_i \oplus Skey_j^*$.

Step 3. Compute $W_j^* = h(Skey_j^* \parallel N_j)$.

Step 4. Compute $(N_i \oplus N_{cs})^* = W_j^* \oplus R_{cs}$.

Step 5. Compute $SK = h(N_j \oplus (N_i \oplus N_{cs})^*)$.

(iii) The time complexity: $\mathcal{O}(|\mathcal{D}_{pw}| * |\mathcal{D}_{id}| * 3T_H)$.

Remark. When considering the forward secrecy, the adversary almost has the same capacity with RA except that $\mathcal{A}$ does not know the verifier-table. As a result, if RA can compute the session key according to the processes of the scheme, then $\mathcal{A}$ is very likely to break the session key. For the above considerations, we do not recommend that RA have the ability to calculate session keys. To achieve this, a public-key algorithm is suggested too [29]. In addition, the more concrete improved methods will be explained in Section 4.

Other flaws: using timestamps to resist replay attack is not recommended. As we all know, due to the network congestion, network latency, or other issues, maintaining a consistent network clock between different systems is very difficult, which often results in the desynchronization attacks. As a matter of fact, many papers [20, 30] in their evaluation criteria pointed out that a protocol using timestamps cannot resist against desynchronization attacks. Furthermore, determining an appropriate value of Δt always faces many challenges in practice: if this value is too big, a replay attack occurs; if it is too small, a valid participant may be stopped. Therefore, in protocol design, the use of random numbers is usually a more recommended way. Unfortunately, the timestamps method was applied in Amin et al.'s scheme.

Insecure identity update phase: similar to the process in "offline dictionary attack II" of Section 3.2, an adversary can carry out an offline dictionary attack via using DD_i or DD_s as the verification parameter when U_i tries to update his/her identity.

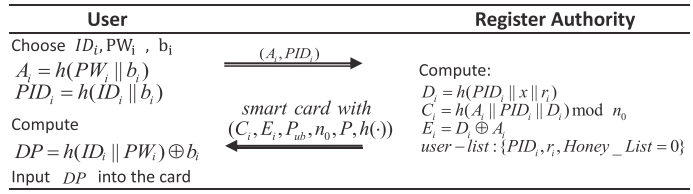


Figure 1: User registration phase.

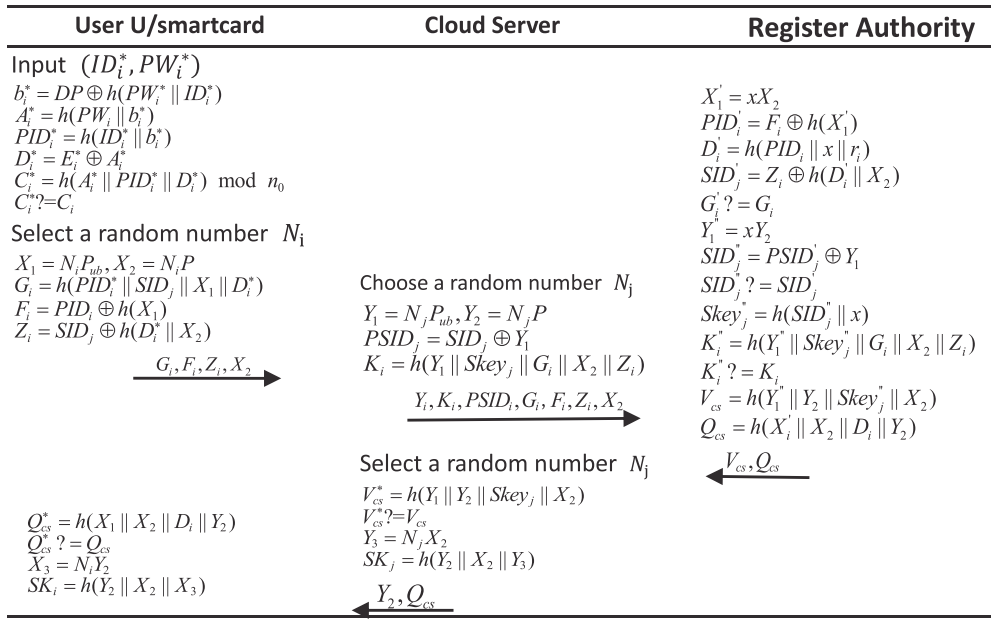


Figure 2: Login and authentication phase.

4. The Proposed Scheme

In this section, we design a secure, simple, and efficient user authentication scheme for cloud computing environment (as shown in Figures 1 and 2) which overcomes all the flaws of Amin et al.'s scheme [3] but provides more attractive attributes, such as updating password and identity and user reregistering with same identity. As a matter of fact, we improve Amin et al.'s scheme from two aspects, efficiency and security. The designing rationales are sketched as follows.

(i) Improvements in Security. According to our discussion in Section 3.2, the public-key algorithm is indispensable to a secure two-factor user authentication scheme [20, 29, 31]. As a matter of fact, this theory has been widely accepted by massive new authentication schemes [26, 30, 35, 36], while the main difficulty lies in deploying the public key algorithm properly. Consequently, we will show the subtleties of deploying a public key algorithm in detail as follows.

Note that, as its high efficiency, the elliptic curve cryptosystem has been used widely in authentication schemes [21, 37–39]. Therefore, our scheme deploys the elliptic curve algorithm to achieve our secure authentication scheme. Under this circumstance, compared with Amin et al.'s scheme [3], our scheme adds a parameter initialization process to set the parameters of elliptic curve cryptosystem.

(1) Apply public-key algorithm to resist against offline dictionary guessing attack II. In Section 3.2, we have shown the pivotal point in addressing such an attack is to add a dynamic nonpublic parameter in G_i. Here we show the ideas about the way of achieving that, to U_i, the shared parameter D_i between RA and U_i is an outcome derived by (ID_i, PW_i). So any parameter based on the transform of (ID_i, PW_i) is futile, as such an parameter is equivalent to D_i. While, with the help of the public-key algorithm, U_i and RA can share a new parameter in every session. More specifically, U_i can choose a random number N_i, compute $X_1 = N_i Pub$, $X_2 = N_i P$ where $Pub = xP$, then let X_1 be the dynamic parameter concealed in G_i. Note that, to U_i and RA, it is easy to compute X_1, yet to $\mathcal{A}$, there is another uncertain parameter in G_i, which stops $\mathcal{A}$ carrying out the offline dictionary attack II.

(2) Apply public-key algorithm to provide user anonymity. With the help of X_1, we conceal the identity related parameter PID_i in F_i as $PID_i \oplus h(X_1)$. Since F_i

is changed with N_i, user anonymity is provided. Note that X_1 is the dynamic parameter we mention in "user anonymity" of Section 3.2.

(3) Apply public-key algorithm to achieve forward secrecy. We set the ECCDH problem in session key to achieve forward secrecy as follows: let SK consist of X_3/Y_3 where $X_3/Y_3 = N_iY_2 = N_iN_jP = N_jX_2$, $Y_2 = N_jP$. As a result, to the one (including RA) who does not know N_i/N_j, computing SK is equivalent to solve the ECCDH problem which cannot be solve in the polynomial time. Therefore, the forward secrecy is achieved.

(4) Following Wang et al.'s way [30] of resisting against offline dictionary attack I, as the detailed explanation on this method can be found in Section IV of [30], Section 6.2.4 of [24], or Section 5.2 of this paper, we do not repeat here.

(ii) Improvements in Efficiency. Note that Amin et al.'s scheme [3] only involves some hash operation, while our scheme deploys a public-key algorithm, so the performance of our scheme is certainly not as efficient as Amin et al.'s scheme. However, except the increased cost of the public-key algorithm, we try to optimize other aspects of the performance in Amin et al.'s scheme through reducing unnecessary parameters or calculations as follows:

(1) Reduce the number of random numbers selected by U_i to one during the user registration process. In Amin et al.' scheme, there are two random numbers in the smart card. While they actually can derive from each other, in addition, the "ability" of computing them is the same, which means they are "equivalent". As a result, using one random number is enough, which saves the storage space and computing resources.

(2) Reduce the number of secret keys in RA. As we see in Section 3.2, the secret key y is of little effect in improving the security. Furthermore, it makes the register phase of CS_j and the authentication of CS_j more complex. Therefore, we only set one system secret parameter and simplify the register phase of CS_j. Such changes also bring other improvements on computing performance.

4.1. Registration. RA selects two large primes $\{p, q\}$ and a medium integer n_0 ($2^4 \le n_0 \le 2^8$). Let $\mathbb{F}_p$ be a finite field, $E/\mathbb{F}_p$ be an elliptic curve over $\mathbb{F}_p$, and $\mathbb{G}$ be a q-order subgroup of $E/\mathbb{F}_p$, then RA chooses a point P in $\mathbb{G}$ and a long-term secret key $x \in Z_p^*$ and computes its public key Pub as xP. In our cloud computing environment, the cloud server and the user registration phases are conducted as follows.

For the Cloud Server CS_j

Step 1. $CS_j \Longrightarrow RA$: $\{SID_j\}$.

Step 2. $RA \Longrightarrow CS_j$: $Skey_j = h(SID_j \parallel x)$.

Step 3. CS_j stores $Skey_j$ as a secret key.

For the UserU_i

Step 1. $U_i \Longrightarrow RA$: $\{A_i, PID_i\}$. A new user U_i firstly selects the password PW_i, identity ID_i, and a random number b_i as his/her personal information and then computes the registration parameters as follows: $A_i = h(PW_i \parallel b_i)$, $PID_i = h(ID_i \parallel b_i)$, and it initiates the registration phase via submitting the request $\{A_i, PID_i\}$ to RA.

Step 2. $RA \Longrightarrow U_i$: a smart card with $\{C_i, E_i, Pub, n_0, h(\cdot)\}$. To guarantee the uniqueness of user identity, RA will firstly check whether the PID_i has been used via traverse $user-list$. If it is available, RA chooses a unique random number r_i for U_i and computes $D_i = h(PID_i \parallel x \parallel r_i)$, $C_i = h(A_i \parallel PID_i \parallel D_i) \bmod n_0$, $E_i = D_i \oplus A_i$, then inputs U_i related parameters $\{PID_i, r_i, Honey\_List = 0\}$ into $User-list$. Note that $Honey\_List$ will record the number of login failures. Finally RA accepts U_i's registration through issuing him/she a smart card with $\{C_i, E_i, Pub, n_0, h(\cdot)\}$.

Step 3. U_i inputs DP into the card where $DP = h(ID_i \parallel PW_i) \oplus b_i$.

4.2. Login phase. As shown in Figure 1, if the user wants to access a cloud server, U_i will submit the information to prove his/her legitimacy in login phase. If the smart card has verified U_i's legality, it initiates the access request to RA for U_i. In short, our login phase involves two aspects: (1) verifying the validity of U_i; (2) initiating the access request.

Step 1. $U_i \longrightarrow CS_j$: login request $\{G_i, F_i, Z_i, X_2\}$. U_i enters $\{ID_i^*, PW_i^*\}$, then the smart card computes $b_i^* = DP \oplus h(ID_i^* \parallel PW_i^*)$, $A_i^* = h(PW_i^* \parallel b_i^*)$, $PID_i^* = h(ID_i^* \parallel b_i^*)$, $D_i^* = E_i \oplus A_i^*$, $C_i^* = h(A_i^* \parallel PID_i^* \parallel D_i^*) \bmod n_0$. The card compares C_i^* with the stored C_i to verify the valid U_i. If $C_i^* \neq C_i$, exit the session.

Step 2. Otherwise, the card accepts U_i's legitimacy and initiates an access request for U_i: select a random number $N_i \in [1, q-1)$, computes $X_1 = N_iPub$, $X_2 = N_iP$, $G_i = (PID_i^* \parallel SID_j \parallel X_1 \parallel D_i^*)$, $F_i = PID_i^* \oplus h(X_1 \parallel X_2)$, $Z_i = SID_j \oplus h(D_i^* \parallel X_2)$, finally transmits $\{G_i, F_i, Z_i, X_2\}$ to CS_j.

4.3. Authentication Phase. Once getting the access request, the CS_j will do some calculation to embed its unique parameters and transmit the request to RA for it is unable to check the authenticity of the request. Then RA will check the validity of the user and the cloud server, respectively, and help them to negotiate the session key. The whole authentication steps are as follows.

Step 1. $CS_j \longrightarrow RA$: $\{Y_2, K_i, PSID_j, G_i, F_i, Z_i, X_2\}$. CS_j selects a random number N_j, computes $Y_1 = N_jPub$, $Y_2 = N_jP$, $PID_j = SID_j \oplus Y_1$, $K_i = h(Y_1 \parallel Skey_j \parallel G_i \parallel X_2 \parallel Z_i)$, then it sends $\{Y_2, K_i, PSID_j, G_i, F_i, Z_i, X_2\}$ to RA.

Step 2. $RA \longrightarrow CS_j$: $\{V_{cs}, Q_{cs}\}$. RA will firstly verify U_i via computing the predefined shared parameter D_i and potential shared secret parameters X_1: $X_1' = xX_2$, $PID_i' = F_i \oplus h(X_1')$, $D_i' = h(PID_i' \parallel x \parallel r_i)$ where r_i is retrieved from the $User - list$ through PID_i', $SID_j' = Z_i \oplus h(D_i' \parallel X_2)$. If U_i's $Hoeny - List$ exceeds a predetermined secure value or the received $G_i \neq (PID_i' \parallel SID_j' \parallel X_1' \parallel D_i')$, RA views U_i as an adversary and let $Hoeny - List = Hoeny - List + 1$ (if it exceeds the preset value, RA will suspend the card till U_i reregisters), then rejects the request.

Otherwise, RA will continue to verify the authenticity of the cloud server as follows: calculate $Y_1'' = xY_2$, $SID_j'' = PSID_j \oplus Y_1$. If $SID_j'' \neq SID_j$, exit (RA thinks CS_j is not the server which U_i actually desires to access); otherwise, continue to compute $Skey_j'' = h(SID_j'' \parallel x)$, $K_i'' = h(Y_1'' \parallel Skey_j'' \parallel G_i \parallel X_2 \parallel Z_i)$. Finally, RA tests the authenticity of CS_j by checking whether $K_i'' \stackrel{?}{=} K_i$. If they are not equal, CS_j does not pass RA's authentication, the session will be terminated.

Otherwise, RA authenticates CS_j and then help them to establish the session key as follows: computes $V_{cs} = h(Y_1'' \parallel Y_2 \parallel Skey_j'' \parallel X_2)$, $Q_{cs} = h(X_1' \parallel X_2 \parallel D_i | Y_2)$, then responds $\{V_{cs}, Q_{cs}\}$ to CS_j.

Note that V_{cs} and Q_{cs} are used to CS_j and U_i, respectively. They are used to verify the authenticity of RA and then convince CS_j/U_i that X_2/Y_2 is truly generated by U_i/CS_j.

Step 3. $CS_j \longrightarrow U_i$: $\{Y_2, Q_{cs}\}$. CS_j computes $V_{cs}^* = h(Y_1 \parallel Y_2 \parallel Skey_j \parallel X_2)$, then check whether $V_{cs}^* \stackrel{?}{=} V_{cs}$. If the equation is not satisfied, CS_j rejects U_i's request. Otherwise, CS_j believes that U_i is a legitimate user who generates X_2, then CS_j computes the session key as $SK_j = h(Y_2 \parallel X_2 \parallel Y_3)$ where $Y_3 = N_j X_2$. Finally, CS_j transmits $\{Y_2, Q_{cs}\}$ to U_i.

Step 4. On receiving this message, U_i will firstly check the valid of RA via comparing $h(X_1 \parallel X_2 \parallel D_i^* \parallel Y_2)$ with the received Q_{cs}. If they are equal, U_i trusts RA and CS_j, then he/she computes their shared session key as $X_3 = N_i \cdot Y_2 \ (= N_i \cdot (N_j \cdot P) = Y_3)$, $SK_i = h(Y_2 \parallel X_2 \parallel X_3)$. Till now, the authentication phase finishes successfully.

4.4. Password Change Phase. Considering the security and user friendliness, our password change phase is conducted locally, which guarantees the efficiency. In other words, the user can change his/her password freely even when he/she does not connect the Internet. All in all, our password change phase is performed as follows.

Step 1. U_i enters ID_i, PW_i, and new password PW_i^{new}.

Step 2. The smart card verified U_i's authenticity as step 1 of Section 4.2. If U_i is authenticated, the card will carry out the password change process as step 3; otherwise, the card will reject this request.

Step 3. The smart card computes some new parameters for new password as follows: $A_i^{new} = h(PW_i^{new} \parallel b_i^*)$, $E_i^{new} = E_i \oplus A_i^* \oplus A_i^{new}$, $DP^{new} = h(ID_i \parallel PW_i^{new}) \oplus b_i^*$, $C_i^{new} = h(A_i^{new} \parallel PID_i^* \parallel D_i^*) \bmod n_0$. Note that $b_i^*, A_i^*, D_i^*, PID_i^*$ is acquired in step 2. Finally, replace $\{C_i, E_i, DP\}$ with $\{C_i^{new}, E_i^{new}, DP^{new}\}$.

4.5. Identity Update Phase. Considering the following occasions, a user sets the phone number as his/her identity, then when the phone number is changed, the user may also want to update the identity. Therefore, similar to password change, the user also needs to change the identity in practice, although it happens less often. Thus we provide the identity update phase as follows.

Step 1. U_i enters ID_i, PW_i, and new password ID_i^{new}.

Step 2. The smart card verified U_i's authenticity as step 1 of Section 4.2. If U_i is not authenticated, the card will reject the request; otherwise the identity update process proceeds.

Step 3. As the PID_i stored in RA is related to ID_i, the identity update phase involves the interaction with RA. On this occasion, the smart card submits $\{X_2, G_i, F_i, NPID_i\}$ to RA for requesting update identity, where N_i is a random number, $X_1 = N_i Pub$, $X_2 = N_i P$, $G_i = h(PID_i^* \parallel X_1 \parallel D_i^* \parallel PID_i^{new})$, $F_i = PID_i^* \oplus h(X_1 \parallel X_2)$, $PID_i^{new} = h(ID_i^{new} \parallel b_i^*)$, $NPID_i = PID_i^{new} \oplus h(X_1 \parallel X_2)$.

Step 4. RA shall verify the valid of U_i as follows: compute $X_1' = xX_2$, $PID_i' = F_i \oplus h(X_1')$, $D_i' = h(PID_i' \parallel x \parallel r_i)$ where r_i is from $User - list$, $PID_i^{new'} = NPID_i \oplus h(X_1 \parallel X_2)$. If the $Hoeny - List$ exceeds a predetermined value or the received $G_i \neq (PID_i' \parallel X_1' \parallel D_i^* \parallel PID_i^{new'})$, RA rejects the request and sets $Hoeny - List = Hoeny - List + 1$. Once it exceeds the preset value, suspend the card.

Otherwise, RA updates PID_i with $PID_i^{new'}$ in $User - list$ and sends $\{M_{cs}, ND_i\}$ to U_i where $M_{cs} = h(PID_i^{new'} \parallel PID_i' \parallel D_i' \parallel X_1')$, $D_i^{new} = h(PID_i^{new'} \parallel x)$, $ND_i = D_i^{new} \oplus h(X_1')$.

Step 5. After receiving M_{cs} from RA, the smart card authenticates RA via testing $M_{cs} \stackrel{?}{=} h(PID_i^{new} \parallel PID_i^* \parallel D_i^* \parallel X_1)$. If the equation does not hold, exit the session; otherwise, finish the identity update process: $DP^{new} = h(ID_i^{new} \parallel PW_i) \oplus b_i^*$, $D_i^{new'} = ND_i \oplus h(X_1)$, $E_i^{new} = D_i^{new'} \oplus A_i^*$, $C_i^{new} = h(A_i^* \parallel PID_i^{new} \parallel D_i^{new'}) \bmod n_0$. Finally, replace $\{C_i, E_i, DP\}$ with $\{C_i^{new}, E_i^{new}, DP^{new}\}$.

4.6. Re-Register Phase. If U_i's smart card is suspended, then he/she shall reregister to RA:

Step 1. $U_i \Longrightarrow RA$: $\{A_i, PID_i, re - register\}$.

Step 2. RA firstly finds PID_i in $User - list$ and checks whether U_i's card is suspended. If so, RA accepts the request and performs the register phase as Section 4.1.

5. Security Analysis

In this section, we analyze the security of our scheme via two popular methods. The results demonstrate that our protocol is secure and effective for the cloud computing environment.

5.1. Formal Analysis Based on BAN Logic. In this section, we apply the BAN logic [40] which is a widely accepted way to analyze the design logic and security of the authentication scheme. Its particular notions to depict protocols are shown in Table 2.

In BAN logic, the goals of our authentication scheme are defined as follows:

(i) Goal 1: $U_i \mid\equiv CS_j \mid\equiv (U_i \overset{SK}{\longleftrightarrow} CS_j)$.

(ii) Goal 2: $U_i \mid\equiv (U_i \overset{SK}{\longleftrightarrow} CS_j)$.

(iii) Goal 3: $CS_j \mid\equiv U_i \mid\equiv (U_i \overset{SK}{\longleftrightarrow} CS_j)$.

(iv) Goal 4: $CS_j \mid\equiv (U_i \overset{SK}{\longleftrightarrow} CS_j)$.

According to the proof steps in BAN logic, we redescribe our scheme into an idealized form:

(i) $Message_1$: $U_i \longrightarrow CS_j$: $\langle X_2, G_i, F_i, Z_i, U_i \overset{X_1}{\longleftrightarrow} RA\rangle_{U_i \overset{D_i}{\longleftrightarrow} RA}$.

(ii) $Message_2$: $CS_j \longrightarrow RA$: $\langle Message_1, Y_2, K_i, PSID_j\rangle_{CS_j \overset{Skey_j}{\longleftrightarrow} RA}$.

(iii) $Message_3$: $RA \longrightarrow CS_j$: $\langle X_2, V_{cs}, \langle Q_{cs}\rangle D_i\rangle_{CS_j \overset{Skey_j}{\longleftrightarrow} RA}$.

(iv) $Message_4$: $CS_j \longrightarrow U_i$: $\langle Y_2, \langle Q_{cs}\rangle D_i\rangle_{U_i \overset{D_i}{\longleftrightarrow} RA}$.

Then, some assumptions are defined as follows:

(i) H_1: $U_i \mid\equiv \sharp(X_2)$.

(ii) H_2: $CS_j \mid\equiv \sharp(Y_2)$.

(iii) H_3: $RA \mid\equiv \sharp(X_2)$.

(iv) H_4: $RA \mid\equiv \sharp(Y_2)$.

(v) H_5: $RA \mid\equiv CS_j \overset{Skey_j}{\longleftrightarrow} RA$.

(vi) H_6: $CS_j \mid\equiv CS_j \overset{Skey_j}{\longleftrightarrow} RA$.

(vii) H_7: $RA \mid\equiv U_i \overset{D_i}{\longleftrightarrow} RA$.

(viii) H_8: $U_i \mid\equiv U_i \overset{D_i}{\longleftrightarrow} RA$.

(ix) H_9: $U_i \mid\equiv CS_j \mid\Longrightarrow U_i \overset{SK}{\longleftrightarrow} CS_j$.

(x) H_{10}: $CS_j \mid\equiv U_i \mid\Longrightarrow U_i \overset{SK}{\longleftrightarrow} CS_j$.

Based on the definition above, we perform the BAN logic proof as follows:

From $Message_2$ ($Message_2$ includes $Message_1$), it is easy to get S_1: $RA \triangleleft \langle Message_1, Y_2, K_i, PSID_j\rangle_{D_i}$.

Then according to H_7, S_1, $RULE(1)$, we get S_2: $RA \mid\equiv U_i \mid\sim \langle X_2, G_i, F_i, Z_i, U_i \overset{X_1}{\longleftrightarrow} RA\rangle$.

According to H_3 and $RULE(4)$, we get S_3: $RA \mid\equiv \sharp\langle X_2, G_i, F_i, Z_i, U_i \overset{X_1}{\longleftrightarrow} RA\rangle$.

And according to S_2, S_3 and $RULE(2)$, we get S_4: $RA \mid\equiv U_i \mid\equiv \langle X_2, G_i, F_i, Z_i, U_i \overset{X_1}{\longleftrightarrow} RA\rangle$.

From the Message, we also get S_5: $CS_j \triangleleft \langle Y_2, K_i, PSID_j\rangle_{Skey_j}$.

Then according to H_7, S_1, $RULE(1)$, we get S_6: $CS_j \mid\equiv RA \mid\sim \langle Y_2, K_i, PSID_j\rangle$.

According to H_3 and $RULE(4)$, we get S_7: $CS_j \mid\equiv \sharp\langle Y_2, K_i, PSID_j\rangle$.

And according to S_2, S_3 and $RULE(2)$, we get S_8: $CS_j \mid\equiv RA \mid\equiv \langle Y_2, K_i, PSID_j\rangle$.

From $Message_3$, it is easy to get S_9: $RA \triangleleft \langle X_2, V_{cs}, \langle Q_{cs}\rangle D_i\rangle_{Skey_j}$.

Then according to H_7, S_1, $RULE(1)$, we get S_{10}: $RA \mid\equiv CS_j \mid\sim \langle X_2, V_{cs}, \langle Q_{cs}\rangle D_i\rangle$.

According to H_3 and $RULE(4)$, we get S_{11}: $RA \mid\equiv \sharp\langle X_2, V_{cs}, \langle Q_{cs}\rangle D_i\rangle$.

And according to S_2, S_3 and $RULE(2)$, we get S_{12}: $RA \mid\equiv CS_j \mid\equiv \langle X_2, V_{cs}, \langle Q_{cs}\rangle D_i\rangle$.

From $Message_4$, it is easy to get S_{13}: $U_i \triangleleft \langle Y_2, \langle Q_{cs}\rangle D_i\rangle_{D_i}$.

Then according to H_7, S_1, $RULE(1)$, we get S_{14}: $U_i \mid\equiv RA \mid\sim \langle Y_2, \langle Q_{cs}\rangle D_i\rangle$.

According to H_3 and $RULE(4)$, we get S_{15}: $U_i \mid\equiv \sharp\langle Y_2, \langle Q_{cs}\rangle D_i\rangle$.

And according to S_2, S_3 and $RULE(2)$, we get S_{16}: $U_i \mid\equiv RA \mid\equiv \langle Y_2, \langle Q_{cs}\rangle D_i\rangle$.

As $SK = h(Y_2 \parallel X_2 \parallel N_i \cdot Y_2)$, and combining S_{12}, S_{16}, we get: S_{17}: $U_i \mid\equiv CS_j \mid\equiv U_i \overset{SK}{\longleftrightarrow} CS_j$ (**Goal 1**).

Similarly, as $SK = h(Y_2 \parallel X_2 \parallel N_j \cdot X_2)$, with S_4, S_8, we get S_{18}: $CS_j \mid\equiv U_i \mid\equiv U_i \overset{SK}{\longleftrightarrow} CS_j$ (**Goal 3**).

Finally, according to H_2, S_{17}, and $RULE(3)$, we get S_{19}: $U_i \mid\equiv (U_i \overset{SK}{\longleftrightarrow} CS_j)$ (**Goal 2**).

And according to H_{10}, S_{18}, and $RULE(3)$, we get S_{20}: $CS_j \mid\equiv (U_i \overset{SK}{\longleftrightarrow} CS_j)$ (**Goal 4**).

In conclusion, our scheme achieves Goals 1~4, which promises (1)U_i and CS_j have got authenticated mutually and (2) they negotiate the same session key SK.

5.2. Informal Analysis. Looking at the history of protocol designing, due to its simplicity and effectiveness, the heuristic method "still plays an important role" in cryptanalysis of protocols [20], though it does not have a theoretical form and relies on human experience heavily. Therefore, this section gives the security analysis via the heuristic method.

User Anonymity. As we mentioned in Section 3.2, user anonymity contains two aspects, we prove our user anonymity attribute from two points.

(1) $\mathscr{A}$ has no chance to acquire ID_i. In our scheme, the identity is transmitted in a form of F_i where $F_i = PID_i \oplus h(X_1)$, $PID_i = h(ID_i \parallel b_i)$, it is obvious that $\mathscr{A}$ has two challenges in computing ID_i: firstly, computing PID_i from F_i; then guessing ID_i from PID_i. The one only with N_i or x can compute $h(X_1)$ successfully, while $\mathscr{A}$ has no way to get this two parameters. Furthermore, even with PID_i, $\mathscr{A}$ still cannot conduct a guessing attack to compute ID_i for $\mathscr{A}$ does not know b_i.

Table 2: Notations in BAN logic.

$P \mid\equiv X$	P believes X, i.e., the principal P believes the statement X is true.
$P \triangleleft X$	P sees X, i.e., the principal P receives a message that contains X.
$P \mid\Longrightarrow X$	P has jurisdiction over X, i.e., the principal P can generates or computes X.
$P \mid\sim X$	P said X, i.e., the principal P has sent a message containing X.
$\sharp(X)$	X is fresh, i.e., X is sent in a message only at the current run of the protocol, it is usually a timestamp or a random number.
$P \overset{K}{\longleftrightarrow} Q$	K is the shared key for P and Q.
$P \overset{Y}{\rightleftharpoons} Q$	Y is the secret known only to P and Q or some principals trusted by them.
$\langle X \rangle_Y$	X combined with Y, and Y usually is a secret.
$\{X\}_K$	X encrypted with K.
$\dfrac{P \mid\equiv P \overset{K}{\longleftrightarrow} Q, P \triangleleft \{X\}_K}{P \mid\equiv Q \mid\sim X}$ or $\dfrac{P \mid\equiv P \overset{Y}{\rightleftharpoons} Q, P \triangleleft \langle X \rangle_Y}{P \mid\equiv Q \mid\sim X}$	*RULE*(1): the message-meaning rule. This rule will be used in the proving process.
$\dfrac{P \mid\equiv \sharp(X), P \mid\equiv Q \mid\sim X}{P \mid\equiv Q \mid\equiv X}$	*RULE*(2): the nonce-verification rule. This rule will be used in the proving process.
$\dfrac{P \mid\equiv Q \mid\Longrightarrow X, P \mid\equiv Q \mid\equiv X}{P \mid\equiv X}$	*RULE*(3): the jurisdiction rule. This rule will be used in the proving process.
$\dfrac{P \mid\equiv \sharp(X)}{P \mid\equiv \sharp(X, Y)}$	*RULE*(4): the freshness-conjuncatenation rule. This rule will be used in the proving process.

(2) $\mathscr{A}$ cannot track the users: as we discussed above, the user's related unique identification is concealed in F_i. It consists of a dynamic parameter X_1 whose value depends on N_i; that is, F_i changes with N_i in every session. So $\mathscr{A}$ can neither links the sessions to a specific user nor tells whether the two sessions are sent by a same user.

Therefore our scheme achieves user anonymity.

Forward Secrecy. Forward secrecy requires that even the lon- term secret key was exposed, the previous session is still secure. In our scheme, the session key $SK = h(Y_2 \parallel X_2 \parallel X_3)$ where $X_3 = N_i Y_2 = N_j X_2 = N_i N_j P$. With the help of x, and the intercepted parameters X_2 and Y_2 from the open channel, if $\mathscr{A}$ wants to compute X_3, then he/she has to solve the ECCDH problem which cannot be finished in the polynomial time. As a matter of fact, once getting x, $\mathscr{A}$ almost has the same capacity with RA. If RA cannot compute SK in the scheme, $\mathscr{A}$ is probably not able to. This again confirms our previous view: for the security consideration, we shall not let RA know SK. All in all, our enhanced scheme provides forward security.

Mutual Authentication. Mutual authentication is the most basic requirement of a user authentication scheme. In our scheme, RA firstly authenticates U_i through G_i which contains their preset shared secret parameter D_i and the dynamic parameter X_1 generated by the public-key algorithm in step 2 of Section 4.2. Then RA authenticates CS_j through K_i with their shared secret value $Skey_j$. In a short, RA authenticates U_i and CS_j after this process.

If both of U_i and G_i are authenticated, RA further computes Q_{cs} and V_{cs}. Then in step 3 of Section 4.2, CS_j checks the validity of RA via V_{cs}. If RA is authentic, then CS_j believes RA's judgment on U_i, which means that CS_j also trusts the legitimacy of U_i. In conclusion, CS_j authenticates RA and U_i via this process.

On receiving Q_{cs}, U_i verifies RA with Q_{cs} in step 4 of Section 4.2. If RA is authenticated, then U_i also believes the validity of CS_j. Therefore, U_i authenticates RA and CS_j.

In conclusion, our scheme achieves mutual authentication.

Privileged Insider Attack. To avoid privileged insider attack, when U_i registers to RA, he/she does not submit the identity or password directly, but a transform of them: $\{A_i, PID_i\}$. Thus the identity and password are protected by b_i. Even the administrator of RA cannot get b_i to conduct an offline dictionary attack to guess the value of ID_i and PW_i. Therefore, our scheme is secure against privileged insider attack.

Offline Dictionary Attack. As we mentioned when analyzing Amin et al.'s scheme [3] in Section 3.2, there are two common offline dictionary attacks. So we consider two kinds of adversary here.

Suppose an adversary $\mathscr{A}$ acquires $\{C_i, E_i, DP, n_0\}$ in the smart card, then $\mathscr{A}$ may conduct an offline dictionary attack as follows.

Step 1. Guess ID_i and PW_i to be ID_i^* and PW_i^*, respectively.

Table 3: Performance comparison among relevant schemes in wireless sensor networks.

	Computation overhead		Communication cost		The evaluation criteria in [24]												
	Login(ms)	Auth.(ms)	Login	Auth.	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
Maitra et al. (2016) [25]	$6T_H \approx 0.04$	$5T_S+14T_H \approx 0.12$	640 bits	1280 bits	√	×	√	×	√	×	×	×	√	√	×	√	√
Kumari et al. (2017) [26]	$2T_M+3T_H \approx 2.3$	$6T_M+15T_H \approx 7.0$	2176 bits	9088 bits	√	√	√	√	×	√	√	×	√	√	√	√	√
Amin et al. (2018) [3]	$6T_H \approx 0.004$	$17T_H \approx 0.012$	640 bits	1920 bits	√	×	√	×	×	√	√	×	√	√	×	√	√
Our scheme	$2T_E+7T_H \approx 1.0$	$4T_E+12T_H \approx 2.0$	1408 bits	4096 bits	√	√	√	√	√	√	√	√	√	√	√	√	√

T_M denotes the time of modular exponentiation operation, T_E denotes scalar multiplication on elliptic curve, T_H denotes hash computation, T_S denotes symmetric encryption/decryption, $T_M \gg T_E \gg T_H > T_S$ ($T_M \approx 1.169ms$, $T_E \approx 0.508ms$, $T_H \approx 0.693\mu s$, $T_S \approx 0.541\mu s$ [20]). Let n_0 be 32-bit long; Let ID_i, PW_i, $h(\cdot)$, output of symmetric encryption, timestamp, random numbers be 128-bit long; let P, x, y be 1024-bit long. √ means the property is satisfied; × means the property is not satisfied. Note that the evaluation criteria in [24] are also applied to cloud computing environment when regarding CS_j as the sensor node, while the considerations on designing authentication scheme of this two environment are quite different due to their different network attributes.

Step 2. Compute $b_i^* = DP \oplus h(ID_i^* \parallel PW_i^*)$ Section 4.1.

Step 3. Compute $A_i^* = h(PW_i^* \parallel b_i^*)$.

Step 4. Compute $PID_i^* = h(ID_i^* \parallel b_i^*)$.

Step 5. Compute $D_i^* = E_i \oplus A_i^*$.

Step 6. Compute $C_i^* = h(A_i^* \parallel PID_i^* \parallel D_i^*) \bmod n_0$.

Step 7. Verify the correctness of PW_i and ID_i by checking if $(C_i^* == C_i)$.

Step 8. Repeat steps 1 ~ 7 until the correct valued of PW_i and ID_i are found.

Now we suppose $\mathscr{A}$ has found such a pair of $\{ID_i^*, PW_i^*\}$ after the above steps. While, due to the properties of equation C_i, even the pair satisfies the equations, it is likely that they are not equal to $\{ID_i, PW_i\}$, since there are $|\mathscr{D}_{pw}| * |\mathscr{D}_{id}|\backslash n_0 \approx 2^{32}$ candidates of $\{ID_i, PW_i\}$ pair when $n_0 = 2^8$ and $|\mathscr{D}_{pw}| = |\mathscr{D}_{id}| = 2^6$ [30]. Consequently, $\mathscr{A}$ needs to verify $\{ID_i^*, PW_i^*\}$ online, but it will be restrained by *Honey_List*. Once the failure numbers of user login exceeds the preset value, the smart card will be suspended. Accordingly, our scheme is secure against such an attack scenario.

Suppose an adversary $\mathscr{A}$ not only extracts the message in smart card, but also eavesdrops $\{G_i, F_i, Z_i, X_2\}$ from the open channel, then $\mathscr{A}$ attempts to guess PW_i and ID_i. In this occasion, $\mathscr{A}$ wants to use G_i as the verification parameter to check the correctness of the guessed value of $\{PW_i, ID_i\}$. As G_i consists of PID_i, SID_j, X_1 and D_i, and according to our attack steps 1~5 above, $\mathscr{A}$ can compute the value of PID_i^* and D_i^*. Then $\mathscr{A}$ computes SID_j^* as $Z_i \oplus h(D_i^* \parallel X_2)$. Now $\mathscr{A}$ only needs to compute X_1. While without x or N_i, computing X_1 is equivalent to solving the ECDL problem which cannot be finished in the polynomial time. As a result, our scheme can prevent such an adversary.

In conclusion, the proposed scheme is secure against dictionary attack.

Verifier-Stolen Attack. In our scheme, *RA* only needs to maintain the *User − list* whose elements ($\{PID_i, r_i, Hoeny\_List\}$) are not security-related. Furthermore, even $\mathscr{A}$ steals *User − list*, he/she will learn nothing useful information to conduct an attack. Thus our scheme is resistant to verifier-stolen attack.

Replay Attack. We prevent the replay attack via the random numbers to prevent replay attack. We take one of the message flow $\{G_i, F_i, Z_i, X_2\}$ as an example to explain: suppose $\mathscr{A}$ eavesdrops $\{G_i, F_i, Z_i, X_2\}$, then replays it to CS_j. While $\mathscr{A}$ does not know N_i, he/she cannot compute the correct session key though the replayed message can pass the verification of *RA*. Consequently, $\mathscr{A}$ gain no benefits from such an attack. Equally, it makes no sense for $\mathscr{A}$ to replay other message flows. Accordingly, our scheme is secure against replay attack.

User Impersonation Attack. According to the definition on user impersonation attack in [24], $\mathscr{A}$ does not acquire the smart card (this condition is included in "offline dictionary attack") here. As we analyzed above, $\mathscr{A}$ with smart card can neither guess ID_i and PW_i nor replay $\{G_i, F_i, Z_i, X_2\}$ to impersonate U_i, let alone the adversary without smart card. So there is only one possible method left: constructing $\{G_a, F_a, Z_a, X_{2a}\}$. To construct this message, $\mathscr{A}$ chooses N_a, computes $\{X_{1a}, X_{2a}\}$, forges PID_a and D_a, calculates $\{G_a, F_a, Z_a, X_{2a}\}$, and finally sends it to *RA*. However, after *RA* gets C_{2a}, and PID_a, *RA* may fail to find such a PID_a in *User − list* or computes a D_a' unequal to D_a, both the two conditions lead to the failure in the authentication of U_i. As a result, *RA* finds that $\mathscr{A}$ is not a legitimate user, the attack fails.

Server Impersonation Attack. According to the above analysis, $\mathscr{A}$ can neither conducts the replay attack to impersonate CS_j/*RA* nor finds ways to compute $Skey_j/x$, so $\mathscr{A}$ cannot impersonate CS_j/*RA*.

6. Performance Analysis

Some schemes like [41] which involve only two participants is essentially indistinguishable from the traditional client-server architecture and does not apply to cloud computing environments with multiple servers. Some schemes like [42] which are more concerned with authentication issues between wearable and smart phone belong to entity authentication rather than the user authentication discussed in this article. These schemes are not comparable to ours. Therefore, we only compare those having similar system architectures and application scenarios including [3, 25, 26].

As shown in Table 3, our security performance is obviously superior to other protocols: the proposed scheme

achieves all security requirements, while others have more or less security flaws. More specifically, both the schemes of Maitra et al. [25] and Amin et al. [3] fail to achieve user anonymity and forward secrecy and cannot resist against offline dictionary attack, etc. The best one is Kumari et al.'s scheme [26] which can only provide 11 items of security requirements. In terms of computation or communication performance, as we mentioned in Section 4, the schemes only involving one-way hash operation are inevitable cost less in communication and computation than those deploying the public-key algorithm, but they certainly cannot guarantee the security of authentication. Therefore, among the compared schemes, the schemes of Amin et al. [3] and Maitra et al. [25] certainly cost less communication time and load for they only involve some one-way hash operations. However, sacrificing security to achieve high performance is inadvisable in authentication protocols. As a matter of fact, certain cost is unavoidable for security. Then compared with Kumari et al.'s scheme [26] which is equipped with public-key algorithm, our scheme costs 1ms in login phase and 2ms in authentication phase, while theirs is 2.3ms and 7ms, respectively, our computation overhead is better. Furthermore, our communication cost (1408bits in login phase and 4096 bits in authentication phase) is also lower than theirs (2176 bits and 9088 bits). In conclusion, our scheme with all security attributes is more suitable for cloud computing environment.

7. Conclusion

The rapid development of cloud computing makes people's lives more convenient, but also brings huge security concerns. In order to ensure user's privacy and account security in the cloud environment, a large number of authentication schemes were proposed, but they were subsequently pointed out having one or more flaws. In order to explain the subtleties of designing an authentication protocol in the cloud environment, this paper took Amin et al.'s protocol as a study case to provide ideas for designing secure protocol for cloud environment through elaborating the security weaknesses existing in the protocol and its corresponding solutions. In addition, based on the analysis, we designed a secure authentication protocol, used the BAN logic and heuristic analysis method to prove the security of the protocol. When comparing it with related protocols, we found our scheme has obvious advantages.

Data Availability

There are not any data used in my paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is supported by the National Key Research and Development Program of China (No. 2017YFB0801900); the BUPT Excellent Ph.D. Students Foundation under Grant No. CX2017206; and the National Natural Science Foundation of China under Grant No. 61472016.

References

[1] T. Mell and P. Grance, "Draft nist working definition of cloud computing," *National Institute of Standards and Technology*, vol. 6, 2009.

[2] H. Li, F. Li, C. Song, and Y. Yan, "Towards smart card based mutual authentication schemes in cloud computing," *KSII Transactions on Internet and Information Systems*, vol. 9, no. 7, pp. 2719–2735, 2015.

[3] 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*, 2016.

[4] M. Wazid, A. K. Das, S. Kumari, X. Li, and F. Wu, "Provably secure biometric-based user authentication and key agreement scheme in cloud computing," *Security and Communication Networks*, vol. 9, no. 17, pp. 4103–4119, 2016.

[5] S. Kumari, M. K. Khan, X. Li, and F. Wu, "Design of a user anonymous password authentication scheme without smart card," *International Journal of Communication Systems*, vol. 29, no. 3, pp. 441–458, 2016.

[6] R. Amin, S. H. Islam, G. P. Biswas, M. K. Khan, L. Leng, and N. Kumar, "Design of an anonymity-preserving three-factor authenticated key exchange protocol for wireless sensor networks," *Computer Networks*, vol. 101, pp. 42–62, 2016.

[7] X. Li, J. Niu, S. Kumari, J. Liao, W. Liang, and M. K. Khan, "A new authentication protocol for healthcare applications using wireless medical sensor networks with user anonymity," *Security and Communication Networks*, vol. 9, no. 15, pp. 2643–2655, 2016.

[8] D. He, S. Zeadally, N. Kumar, and W. Wu, "Efficient and Anonymous Mobile User Authentication Protocol Using Self-Certified Public Key Cryptography for Multi-Server Architectures," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2052–2064, 2016.

[9] X. Huang, Y. Xiang, E. Bertino, J. Zhou, and L. Xu, "Robust multi-factor authentication for fragile communications," *IEEE Transactions on Dependable and Secure Computing*, vol. 11, no. 6, pp. 568–581, 2013.

[10] J. H. Yang, Y. F. Chang, and C. C. Huang, "A user authentication scheme on multi-server environments for cloud computing," in *Proceedings of the ICICS 2013*, pp. 1–4, 2013.

[11] J. H. Yang and P. Y. Lin, "An ID-Based User Authentication Scheme for Cloud Computing," in *Proceedings of the IIH-MSP 2014*, pp. 98–101, 2014.

[12] T.-H. Chen, H.-L. Yeh, and W.-K. Shih, "An advanced ECC dynamic ID-Based remote mutual authentication scheme for Cloud Computing," in *Proceedings of the MUE 2011*, pp. 155–159, 2011.

[13] D. Wang, Y. Mei, C. Ma, and Z. Cui, "Comments on an Advanced Dynamic ID-Based Authentication Scheme for Cloud Computing," in *Proceedings of the WISM 2012*, pp. 246–253, 2012.

[14] K.-P. Xue, P.-L. Hong, and C.-S. Ma, "A lightweight dynamic pseudonym identity based authentication and key agreement protocol without verification tables for multi-server architecture," *Journal of Computer and System Sciences*, vol. 80, no. 1, pp. 195–206, 2014.

[15] M.-C. Chuang and M. C. Chen, "An anonymous multi-server authenticated key agreement scheme based on trust computing using smart cards and biometrics," *International Journal of Network Security*, vol. 18, no. 5, pp. 997–1000, 2014.

[16] J.-L. Tsai and N.-W. Lo, "A Privacy-Aware Authentication Scheme for Distributed Mobile Cloud Computing Services," *IEEE Systems Journal*, vol. 9, no. 3, pp. 805–815, 2015.

[17] V. Odelu, A. K. Das, S. Kumari, X. Huang, and M. Wazid, "Provably secure authenticated key agreement scheme for distributed mobile cloud computing services," *Future Generation Computer Systems*, vol. 68, pp. 74–88, 2017.

[18] P. Gope and A. K. Das, "Robust Anonymous Mutual Authentication Scheme for n-Times Ubiquitous Mobile Cloud Computing Services," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1764–1772, 2017.

[19] S. Roy, S. Chatterjee, A. K. Das, S. Chattopadhyay, N. Kumar, and A. V. Vasilakos, "On the Design of Provably Secure Lightweight Remote User Authentication Scheme for Mobile Cloud Computing Services," *IEEE Access*, 2017.

[20] D. Wang, D. He, P. Wang, and C.-H. Chu, "Anonymous Two-Factor Authentication in Distributed Systems: Certain Goals Are Beyond Attainment," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 428–442, 2015.

[21] 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.

[22] C. Wang, G. Xu, and J. Sun, "A secure and anonymous two-factor authentication protocol in multi-server environment," *Security and Communication Networks*, vol. 2018, 15 pages, 2018.

[23] D. Wang, W. Li, and P. Wang, "Measuring Two-Factor Authentication Schemes for Real-Time Data Access in Industrial Wireless Sensor Networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 4081–4092, 2018.

[24] C. Wang, G. Xu, and J. Sun, "An enhanced three-factor user authentication scheme using elliptic curve cryptosystem for wireless sensor networks," *Sensors*, vol. 17, no. 12, article no. 2946, 2017.

[25] T. Maitra, S. H. Islam, R. Amin, D. Giri, M. K. Khan, and N. Kumar, "An enhanced multi-server authentication protocol using password and smart-card: cryptanalysis and design," *Security and Communication Networks*, vol. 9, no. 17, pp. 4615–4638, 2016.

[26] S. Kumari, X. Li, F. Wu, A. K. Das, K.-K. R. Choo, and J. Shen, "Design of a provably secure biometrics-based multi-cloud-server authentication scheme," *Future Generation Computer Systems*, vol. 68, pp. 320–330, 2017.

[27] 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.

[28] 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.

[29] C.-G. Ma, D. Wang, and S.-D. Zhao, "Security flaws in two improved remote user authentication schemes using smart cards," *International Journal of Communication Systems*, vol. 27, no. 10, pp. 2215–2227, 2012.

[30] 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*, 2016.

[31] D. Wang and P. Wang, "On the anonymity of two-factor authentication schemes for wireless sensor networks: Attacks, principle and solutions," *Computer Networks*, vol. 73, pp. 41–57, 2014.

[32] D.-J. He, M.-D. Ma, Y. Zhang, C. Chen, and J.-J. Bu, "A strong user authentication scheme with smart cards for wireless communications," *Computer Communications*, vol. 34, no. 3, pp. 367–374, 2011.

[33] Q. Jiang, J. Ma, Z. Ma, and G. Li, "Security analysis and improvement of bio-hashing based three-factor authentication scheme for telecare medical information systems," *Journal of Medical Systems*, 2017.

[34] Y. Lu, L. Li, H. Peng, and Y. Yang, "A Novel Smart Card Based User Authentication and Key Agreement Scheme for Heterogeneous Wireless Sensor Networks," *Wireless Personal Communications*, vol. 96, no. 1, pp. 813–832, 2017.

[35] C. Wang, D. Wang, G. Xu, and Y. Guo, "A lightweight password-based authentication protocol using smart card," *International Journal of Communication Systems*, vol. 30, no. 16, Article ID e3336, 2017.

[36] Q. Jiang, J. Ma, C. Yang, X. Ma, J. Shen, and S. A. Chaudhry, "Efficient end-to-end authentication protocol for wearable health monitoring systems," *Computers and Electrical Engineering*, 2017.

[37] M. K. Khan and D. He, "A new dynamic identity-based authentication protocol for multi-server environment using elliptic curve cryptography," *Security and Communication Networks*, vol. 5, no. 11, pp. 1260–1266, 2012.

[38] A. K. Das, S. Kumari, V. Odelu, X. Li, F. Wu, and X. Huang, "Provably secure user authentication and key agreement scheme for wireless sensor networks," *Security and Communication Networks*, vol. 9, no. 16, pp. 3670–3687, 2016.

[39] D. Wang, H. Cheng, D. He, and P. Wang, "On the Challenges in Designing Identity-Based Privacy-Preserving Authentication Schemes for Mobile Devices," *IEEE Systems Journal*, vol. 12, no. 1, pp. 916–925, 2018.

[40] M. Burrows, M. Abadi, and R. Needham, "A logic of authentication," *ACM Transactions on Computer Systems*, vol. 8, no. 1, pp. 18–36, 1990.

[41] S. Kumari, M. Karuppiah, A. K. Das, X. Li, F. Wu, and N. Kumar, "A secure authentication scheme based on elliptic curve cryptography for IoT and cloud servers," *The Journal of Supercomputing*, 2017.

[42] F. Wu, X. Li, L. Xu, S. Kumari, M. Karuppiah, and J. Shen, "A lightweight and privacy-preserving mutual authentication scheme for wearable devices assisted by cloud server," *Computers and Electrical Engineering*, vol. 63, pp. 168–181, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 1875489, 12 pages
https://doi.org/10.1155/2018/1875489

Research Article

Efficient Message Authentication Scheme with Conditional Privacy-Preserving and Signature Aggregation for Vehicular Cloud Network

Yong Xie[1], **Fang Xu**[2], **Dong Li**[1] **and Yu Nie**[3]

[1]*Department of Computer Technology and Application, Qinghai University, China*
[2]*School of Computer and Information Science, Hubei Engineering University, Xiaogan, China*
[3]*School of Computer and Information Science, Jingdezhen Ceramic Institute, Jingdezhen, China*

Correspondence should be addressed to Fang Xu; xf2012@whu.edu.cn

Received 11 April 2018; Revised 10 August 2018; Accepted 30 August 2018; Published 23 September 2018

Academic Editor: Kim-Kwang Raymond Choo

Vehicular cloud network (VCN) is deemed as the most promising platform for providing transportation safety, road optimization, and valued-added application services. Because VCN is of distinguishing feature with super-large scale and unstable communication, it is a challenging task to study efficient authentication scheme for VCN without losing security and conditional privacy-preserving. To meet the challenge, a new efficient message authentication scheme is proposed in this paper. A batch message verification and signature aggregation are included in the proposed scheme to improve the authentication efficiency and decrease the communication cost. Compared with the similar conditional privacy-preserving authentication schemes, the proposed scheme has superior performance in computation and communication cost. Simulation analysis further proves that the proposed scheme has better advantages in reducing the verification loss rate and message delay in the application of VCN.

1. Introduction

As the growing demand for transportation safety, driver comfort, and traffic efficiency, it is crucial for vehicles to obtain current traffic-related information accurately and timely. To meet the goal, vehicular ad hoc networks (VANETs) have been raised and caused heated joint researches among researchers, car manufactures, and governments in recent years [1]. Due to the specific features and applications of VANETs, people expect that a vehicle can perform all the tasks of communication, computing, sensing, and storage. On the one hand, a vehicle has some in-car resources, such as sensor, power, CPU, communication units, and actuator, and it should schedule the in-car resource harmoniously to achieve optimal efficiency. On the other hand, a vehicle should cooperate with other units, such as other vehicles and Roadside Units (RSUs), to make use of the unstable external resources in an effective way [2]. Therefore, vehicle will gradually become a complicated integrated intelligent system with computing, mechanical, and communication function in the near future.

Because cloud computing technology has shown many outstanding advantages in practice application, some researchers have proposed vehicular cloud computing, which has been a new paradigm employed by vehicle (driver) to leverage services as a utility and handle a mass of data on demand at any time and anywhere [3]. Thus, to improve efficiency of vehicle-related services to vehicles, some interesting vehicular cloud network (VCN) architectures over VANETs have been proposed recently [4, 5]. A general VCN architecture consists of three tiers: the top tier includes the trusted authority (TA) and cloud servers; the middle tier includes intermediate units including road side units (RSUs), 3/4G base stations (BSs), and other network access units; the bottom tier includes in-car units of vehicles including On-Board Unit (OBU), sensors, 3/4G module, and other modules, as shown in Figure 1. RSUs and BSs are placed on the side of road and can communicate with TA and cloud servers via wired communication. OBU is in charge of communication with other vehicle's OBUs by Vehicle-to-Vehicle (V2V) communication technology, and it also can communicate with RSUs by Vehicle-to-Infrastructure (V2I)

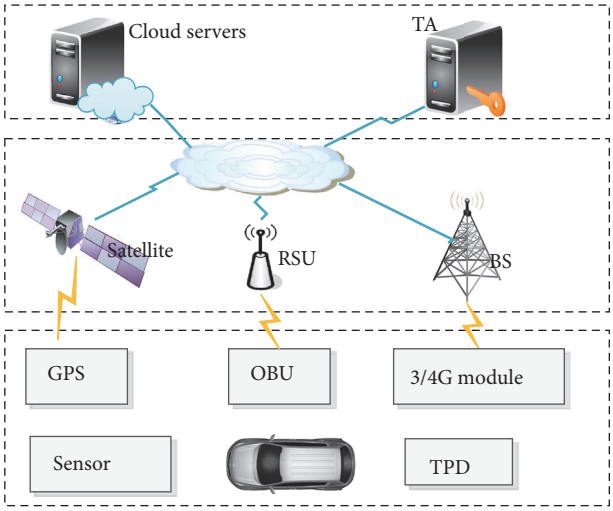


Figure 1: A general VCN architecture.

communication technology. Ranging from transportation safety to valued-added application services, VCN is regarded as one of the most promising platforms for future vehicle-centered applications [6].

Nonetheless, benefits usually come with challenges. Because messages in VCN are usually life-critical, the foremost issue is security that the messages must be authenticated and reliable [7]. Nowadays, privacy protection has become the most urgent requirement that users are most concerned about in the open and insecure wireless communication environment [8, 9]. If an attacker could retrieve the private information of a vehicle by linking the messages, the most promising VCN will be gutted. Therefore, the second important issue is privacy-preserving. However, privacy-preserving is the double-edged sword of VCN: A honest vehicle is willing to broadcast real message to its neighbor vehicles; a malicious vehicle may send wrong messages for personal gain by abusing the privacy protection mechanisms, where wrong message has a valid signature and untrue content. Because a wrong message may cause inestimable damage to the traffic system or people's personal safety, there must be one and only one (usually is the TA) that should have the ability to trace the real identity of wrong message generator. Therefore, conditional privacy-preserving (CPP) should be involved in VCN. It is generally known that a huge volume of messages of VCN may be produced in a short time and the communication instability problems of VCN is particularly serious. In order to improve the quality of VCN service, it should decrease communication cost and computation cost. Therefore, the third key issue is to improve authentication efficiency and decrease communication cost without losing security and cryptographic witnesses. To solve the three challenges, industry and academia have done a lot of research works and put forward a lot of interesting results [10].

1.1. Motivations and Contributions. In VCN, there are usually millions of messages being produced in a very short time, and many messages must be processed timely because they are time sensitive and life-critical. However, it is an arduous task for OBUs or RSUs to verify vast messages timely [11]. Thus, it is a significant challenge to design a practical message authentication scheme for VCN under the precondition of ensuring safety and conditional privacy-preserving.

To meet this challenge, we propose a new message authentication scheme with CPP and signature aggregation. In short, our main contributions can be summarized as follows:

(i) A new efficient message authentication scheme is proposed for VCN using elliptic curves cryptography (ECC). Signature aggregation and batch verification are involved to improve verification efficiency further, where the batch verification allows verifier to verify multiple messages simultaneously and the signature aggregation allows verifiers to aggregate multiple signatures into a single one before forwarding them to its top manager (e.g., cloud servers).

(ii) A rigorous security analysis shows that the proposed scheme could satisfy all security requirements of VCN and provides CPP.

(iii) Performance analysis indicates that our proposed scheme can perform much better in terms of computation cost and the communication cost than most recent schemes proposed in [12–14]. The signature aggregation of the proposed scheme could further decrease communication cost. Simulations show that the proposed scheme also could reduce verification loss rate and message delay in VCN scenario.

1.2. Organization of the Paper. The rest of the paper is organized as follows. Preliminaries and background are introduced in Section 2. Section 3 shows background and Section 4 puts forward a new message authentication scheme for VCN. Section 5 demonstrates security proof and analysis. Section 6 discusses complexity analysis and comparisons. The last section concludes the current and future works.

2. Related Work

To achieve CPP authentication, some researchers have proposed classic authentication schemes by using group signature [15–18]. Before a vehicle communicates with other vehicles, it should join in the group to get signing key from the group manager. After then the vehicle uses signing key to sign messages on behalf of the group. Only the group manager can retrieve the identity of message signer, so this kind of authentication schemes can meet conditional privacy-preserving requirement. But, these authentications have much higher communication and computation cost than traditional signatures and have inextricable problem on member revocation [19].

To decrease communication and computation cost, Raya et al. [20] adopted anonymous certificate based on Public Key Infrastructure (PKI) to construct an anonymous authentication scheme for vehicle network. Later, some similar CPP authentication scheme has been proposed [16, 21, 22]. However, it is extremely difficult for these schemes using PKI to overcome issues related to certificate management.

To overcome certificate issues, researchers introduced identity-based public key cryptosystem (ID-PKC) [23] to design message authentication scheme for vehicle network, where no certificate is needed to bind to public key pairs. Zhang et al. [24] used bilinear pairing to construct message authentication scheme based on IP-PKC. Zhang et al.'s scheme [24] no longer needs any certificates. Unfortunately, relay attack and impersonation attack can be launched easily in their scheme. By using two shared secretes, Chim et al. [25] put forward one identity-based authentication scheme. Under the condition of providing anonymity, Chim et al.'s scheme need less communication cost than Zhang et al.'s [24]. But, Chim et al.'s scheme is demonstrated to suffer from impersonation attack. Lee et al. [4] presented a new message authentication scheme employing bilinear pairing. Unfortunately, their scheme could not provide tracing and nonrepudiation and also suffers from relay attacking. To overcome secure issues, Bayat et al.'s [12] presented an reformative authentication scheme over Lee et al.'s scheme [4]. They demonstrated security analysis to show that their scheme can resist various security attacks. However, the aforementioned schemes based on PKC use complex bilinear pairing operations, which is quit complex cryptographic operation in modern cryptography and not suited for OBUs that is limited in computational capacity. To wipe off bilinear pairing, He et al.'s [14] proposed a new conditional preserving scheme by using ECC. He et al. demonstrated that their scheme takes more lower computation cost and communication cost, which makes their scheme more suited for deployment in VCN. Xie et al. [26] proposed an identity-based message authentication scheme for vehicle network using ECC. Their scheme provides not only single message verification but also batch message verification; it can decrease much authentication costs. Unfortunately, it can not provide aggregate authentication. Kang et al. [27] used homomorphic encryption to allow every vehicle to generate any number of authenticated identities to realize anonymity in vehicle network. Recently, Liu et al. [28] proposed a mutual authentication and key agreement scheme for secure vehicle-to-vehicle communication. But the TA should include each authentication process in their scheme, which brings a very large computational overhead to the TA.

Signature aggregation on cryptographic witnesses has drawn more attention due to its special way to improve system performance. Zhang et al. [19] proposed an aggregate privacy-preserving authentication scheme for VANETs. In their scheme, aggregate signature technique is used as an important way to decrease computation and communication overhead during data transmission and signature authentication. But when a vehicle joins a RSU authentication group, the RSU must forward vehicle's information to the root TA through a secure channel. Wasef et al. [29] proposed aggregation protocols based on PKI in vehicle ad hoc network, respectively. The two protocols can aggregate multiple signatures into a single one but cannot aggregate different certificates, which remains a problem on certificate management. To eliminate problem on certificate management, signature aggregation based on identity-based PKC was proposed in [30]. Zhang et al. [13] proposed a hierarchical aggregation to suit for hierarchical management in VANETs. In their scheme, a secure channel must be preestablished between an RSU and the KGC for vehicle's identity authentication.

All kinds of identity-based schemes for vehicle networks proposed during the last decades can be divided into two major categories. One is using traditional authentication way without using Tamper-proof devises (TPD) [31]; the other more efficient authentication way is by using TPDs. Compared with non-TPD, schemes using TPD are more efficient. Therefore, we construct the proposed scheme using TPD to solve the very arduous message authentication tasks in vehicular cloud network.

3. Background

3.1. System Architecture of VCN. The three-tier architecture proposed in [32] is used in this paper. The top tier consists of the trusted authority (TA) and cloud services, the middle tier consists of intermediate units, the bottom tier consists of in-car units of vehicles, as shown in Figure 1.

(i) Top Tier. The same assumption applies with [13]; the TA is a fully trusted administrator, and it is in charge of generating system parameters and allocating Tamper-proof devises (TPD) to each registered role, such as RSUs, vehicles, and cloud serves. A secure access password will be set according to the rules proposed in [33, 34] for each TPD and can be used when the user inputs the correct password. In the system, only the TA is able to retrieve the real identities from valid messages when necessary. The TA is assumed to be never compromised by any adversaries. The cloud services are provided cloud servers by using cloud computing technique and are usually made up of road traffic monitoring, diver body monitoring, whether information, entertainment service, and other services that can be customized by users.

(ii) Middle Tier. This tier consists of communication entities, such as RSU, Base stations, and satellite (for connecting to Internet), GPS module (for connecting to satellite network), and 3/4G communication module (for connecting 3/4G wireless network). RSUs are a number of substance units placed on the side of roads. A RSU communicates with vehicles' OBUs by using DSRC protocol and with TA and cloud servers using wired channel. A RSU must verify signatures as soon as receiving messages from vehicles and decides whether to process them locally or deliver them to the top server (including cloud service). BS and satellite connect the 3/4G module and GPS module of vehicles, respectively.

(iii) Bottom Tier. This tier consists of On-Board Unit (OBU), TPD, GPS module, 3/4G module, sensors and reactors, and other in-car units. The TA will issue a TPD for each registered vehicle. TPD has high-level ability to withstand any security attacks and no one can extract any data from TPDs, such as secret key and codes [12, 16]. Any message will be signed by TPD before being broadcasted. The OBU collects raw data from other in-car units and then broadcasts messages about traffic status and other service request message. In addition, it is also responsible for communicating with other OBUs and

RSUs under DSRC protocol. The 3/4G module is responsible for communicating with the BS.

3.2. Security Requirements. A lot of attacks threaten the security of VCN, such as privacy disclosure, relay attack, man-in-the-middle attack, and modification attack. To avoid these attacks, the following security requirements should be provided in the authentication scheme.

(1) Message Authentication. In VCN, each verifier can authenticate every message and determines whether the message signer is a registered member and judges whether the message is modified by others.

(2) Conditional Privacy-Preserving (CPP) [35]. As with other scenarios of privacy protection, the true identity of the vehicle should be anonymous, including other vehicles, RSUs, and attackers. But registered vehicles with malicious behavior may abuse anonymous mechanism and broadcast wrong messages. In order to restrict the registered vehicles to use anonymity mechanism in rational way, the TA must extract the signer of valid message (with valid signature). As a consequence, authentication schemes must provide CCP functionality [36].

(3) Resistance to Attacks. To meet the requirements of security, authentication schemes must be able to withstand all possible attacks, e.g., forgery attack and man-in-the-middle attacks.

4. The Proposed Scheme

In this section, we propose a new efficient identity-based authentication scheme for VCN, which achieves CPP functionality. The proposed scheme includes four phases: initialization, pseudonym generation and message signing phase, message verification phase, and identity extraction phase. To improve efficiency, batch message verification and signature aggregation are involved in message verification phase.

In order to understand the phases of the proposed scheme more intuitively, the main phases of proposed scheme are illustrated as in Figure 2. In Figure 2, PMS denotes Pseudonym Generation and Message Signing, which is executed by the messages signer, i.e., vehicles; SMV, BMV, and SA denote single message verification, batch message verification, and signature aggregation, respectively, which are executed by low-lever verifier, such as RSUs or vehicles; AMV denotes aggregated messages verification, which is executed by top manager, such as cloud severs or application servers.

Next, we will show the details of each phase as in the following subsections.

4.1. Initialization. In this phase, the system parameter is initialized by the TA, the detailed steps are as follows:

I1: the TA selects an elliptic curve $E_p(a,b)$, which is defined by $y^2 = x^3 + ax + b \bmod p$, where p is a large prime number, $a, b \in F_p$. Then the TA chooses a generator point P from $E_p(a,b)$, and generates group G by P with order q.

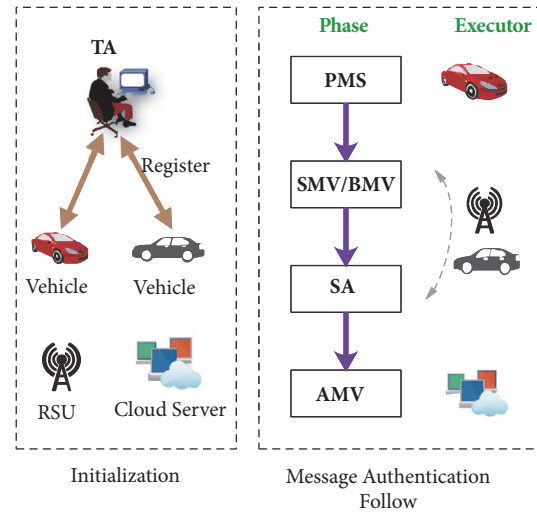


Figure 2: The main phases of proposed scheme.

Next, the TA chooses $s \in {}_RZ_q^*$ as its private key and computes public key $P_{pub} = s \cdot P$.

I2: two hash functions, $h_1 : G \longrightarrow Z_q$, $h_2 : \{0,1\}^* \longrightarrow Z_q$, are chosen as cryptographic hash function. Now, $params = \{E_p(a,b), p, q, P, h_1, h_2, P_{pub}\}$ is set as system public parameter.

I3: when a vehicle V_i registers in the system, the TA assigns a TPD to the vehicle, where the TPD will be preloaded parameters $\{RID_i, PWD_i, s, params\}$. Therefore, each vehicle will obtain unique identifier RID_i and password PWD_i.

I4: at last, the public parameter $params$ is published to each registered vehicle, RSU and cloud server.

4.2. Pseudonym Generation and Message Signing Phase. When a vehicle V_i wants to broadcast or send a message, it generates a pseudonym and sign messages by using its TPD as follows.

S0: the user input the valid RID_i and PWD_i to gain the right to use the TPD. To be practical, the user can employ the TPD to generate pseudonym for a period after he/she has input valid RID_i and PWD_i; i.e., this step will not be run during the next period, while steps **S1**-**S3** will be run in this phase.

S1; when a message m_i is generated by the OBU or sensors, it is transmitted to the TPD.

S2: on receiving m_i, the TPD chooses $r_i \in {}_RZ_q^*$ and current timestamp T_i and then calculates $PID_{i,1} = r_i \cdot P$, $PID_{i,2} = RID_i \oplus h_1(r_i \cdot P_{pub})$. Let PID_i denote $\{PID_{i,1}, PID_{i,2}\}$. Next, the TPD computes $h_i = h_2(m_i, PID_i, T_i)$, $\delta_i = (r_i + h_i)/s$. Finally, the TPD sends $m_i, PID_i, T_i, \delta_i$ to the OBU.

S3: the vehicle V_i broadcasts $m_i, PID_i, T_i, \delta_i$.

The steps of this phase are outlined in Figure 3.

4.3. Message Verification Phase. It is a normal state in VCN that an entity (such as a vehicle or a RSU) receives a mass of messages in a brief period. To improve the efficiency of message verification, there are two ways to verify that the received messages are presented in our scheme. One is traditional

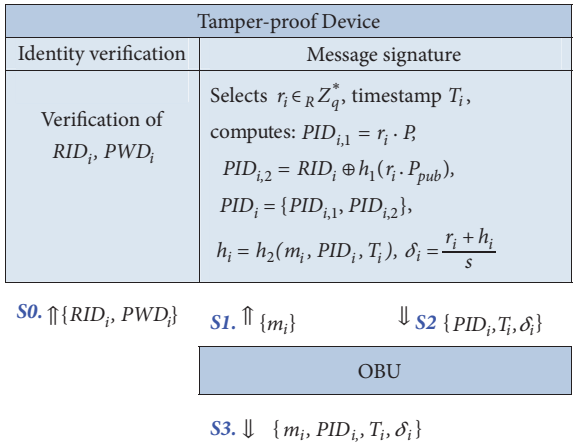


Figure 3: The steps of pseudonym generation and message signing phase.

single message verification for one message. The other is batch verification for multiple messages simultaneously.

(i) Single Message Verification. Assume $m_i, PID_i, T_i, \delta_i$ generated by the vehicle V_i is a message needed to be verified. The T_i of message m_i will be checked firstly. If is not fresh, the verifier discards this message. Otherwise, the verifier computes $h_i = h_2(m_i, PID_i, T_i)$ and then examines if this message satisfies the verification equation as follows:

$$\delta_i \cdot P_{pub} = PID_{i,1} + h_i \cdot P \tag{1}$$

If not, this message will be discarded. Or, it will be accepted.

(ii) Batch Message Verification. After n messages $\{m_1, PID_1, T_1, \delta_1\}, \{m_2, PID_2, T_2, \delta_2\}, \ldots, \{m_n, PID_n, T_n, \delta_n\}$ are received by the verifier, they could be verified simultaneously as the following steps.

B1: the T_i of message m_i $(i = 1, 2, \ldots, n)$ will firstly be checked. If it is not fresh, the verifier discards m_i.

B2: to reduce false acceptation, the small exponent test technology [4] is included in batch verification. A vector including small random integers is used to distinguish any modification on multiple signatures during batch verification. The verifier chooses $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots \lambda_n\}$, where λ_i is randomly chosen in $[1, \gamma]$; γ is a very small integer and only causes little computational overhead [4].

B3: the verifier checks whether (2) holds or not.

$$\left(\sum_{i=1}^{n} \lambda_i \delta_i\right) \cdot P_{pub} = \sum_{i=1}^{n} (\lambda_i \cdot PID_{i,1}) + \left(\sum_{i=1}^{n} \lambda_i h_i\right) \cdot P \tag{2}$$

where $h_i = h_2(m_i, PID_i, T_i)$. If (2) holds, the n messages will be accepted. Or, one or more messages are invalid in the n messages. To detect invalid message, the way proposed in [37] is used in the proposed scheme. For more details, please see [37].

If the n messages are valid, the verifier accepts the n messages and can send n messages as $M^* = \{(m_1, PID_1, T_1, \delta_1), (m_2, PID_2, T_2, \delta_2), \ldots, (m_n, PID_n, T_n, \delta_n)\}$ to its top manager in traditional ways. To improve efficiency and decrease communication cost, a signature aggregation is included in the proposed scheme.

(iii) Signature Aggregation. To decreasing communication cost, a verifier in the lower layer of system can make aggregate signature on the messages that have been verified before forwarding these messages to its top managers.

Firstly, the verifier computes $\delta_A^* = \sum_{i=1}^{n} \delta_i$. Then he/she generates the aggregated message $M_A^* = \{(m_1, PID_1, T_1), (m_2, PID_2, T_2), \ldots, (m_n, PID_n, T_n), \delta_A^*\}$. At last, the verifier forwards the aggregated message to its top manager.

When the top manager receives $\hat{n}$ aggregated messages $\{(m_1^1, PID_1^1, T_1^1), \ldots, (m_n^1, PID_n^1, T_n^1), \delta_A^1\}, \ldots, \{(m_1^{\hat{n}}, PID_1^{\hat{n}}, T_1^{\hat{n}}), \ldots, (m_n^{\hat{n}}, PID_n^{\hat{n}}, m_n^{\hat{n}}, T_n^{\hat{n}}), \delta_A^{\hat{n}}\}$, it can verify single aggregated message by following verification equation (3):

$$\delta_A^* \cdot P_{pub} = \sum_{i=1}^{n} PID_{i,1} + h_A^* \cdot P \tag{3}$$

where $h_A^* = \sum_{i=1}^{n} h_2(m_i, PID_i, T_i)$. If (3) holds, the top manager accepts the aggregated message. To improve efficiency, the top manager also can verify the aggregated messages by following verification equation (4):

$$\left(\sum_{j=1}^{\hat{n}} \delta_A^j\right) \cdot P_{pub} = \sum_{j=1}^{\hat{n}} \sum_{i=1}^{n} PID_{i,1}^j + \left(\sum_{j=1}^{\hat{n}} h_A^j\right) \cdot P \tag{4}$$

If (4) holds, the top manager accepts the $\hat{n}$ aggregated messages.

4.4. Identity Tracing Phase. To obtain profit or disrupt traffic, a registered vehicle V_i perhaps sends false message m_i; that is, m_i has wrong/untrue context with valid signature. Therefore, the functionality of tracing the identity of false messages must be provided in message authentication scheme. Assume the message m_i in $M^* = \{(m_1, PID_1, T_1), (m_2, PID_2, T_2), \ldots, (m_n, PID_n, T_n), \delta^*\}$. Note that the n messages have passed the signature verification. The TA traces the real identity RID_i from m_i by calculating $RID_i = PID_{i,2} \oplus h_1(s \cdot PID_{i,1})$, where s is its private key.

5. Security Proof and Analysis

In this section, we demonstrate that the proposed scheme satisfies the security requirements of VCN described in Section 3.2. In order to prove that the proposed scheme is secure against all types of attacks, we show the nonforgery of the proposed scheme firstly.

5.1. Security Proof. In order to prove the security of the proposed scheme, the security model is defined as a game that is performed by an adversary and a challenger based on the ability of the adversary and the network model.

Theorem 1. *The proposed scheme is existentially unforgeable against an adaptive chosen-message under the random oracle model.*

Proof. Assume an ECDLP instance $(P, Q = xP)$ is given, where P, Q are two points on E/E_p and an adversary $\mathscr{A}$ could forge message $\{m_i, PID_i, T_i, \delta_i\}$. Now, we set up a game between $\mathscr{A}$ and a challenger $\mathscr{C}$, which is able to solve the ECDLP by running $\mathscr{A}$ as a subroutine with a probability that cannot be ignored.

Setup. The challenger $\mathscr{C}$ executes system setup algorithm, lets $P_{pub} = Q = xP$ as system public key, and defines system parameter params=$\{E_p(a,b), p, q, P, h_1, h_2, P_{pub}\}$ and then creates and preserves two lists. One is list L_{h1} formed by $< \alpha, \tau_{h1} >$, which contains the queries and answers of h_1-Oracle and is empty initially. Another is list L_{h2} formed by $< m_i, PID_i, T_i, \tau_{h2} >$, which includes the queries and answers of h_2-Oracle and is empty initially. At last $\mathscr{C}$ sends params to $\mathscr{A}$.

h_1-Oracle. When $\mathscr{A}$ queries message α, $\mathscr{C}$ checks whether the tuple $< \alpha, \tau_{h1} >$ is already in L_{h1} or not. If so, $\mathscr{C}$ sends $\tau_{h1} = h_1(\alpha)$ to $\mathscr{A}$. Otherwise, $\mathscr{C}$ chooses $\tau_{h1} \in Z_q^*$ at random and then adds $< \alpha, \tau_{h1} >$ to L_{h1}. At last, $\mathscr{C}$ sends $\tau_{h1} = h_1(\alpha)$ to $\mathscr{A}$.

h_2-Oracle. When $\mathscr{A}$ queries message $< m_i, PID_i, T_i, \tau_{h2} >$, $\mathscr{C}$ checks if the tuple $< m_i, PID_i, T_i >$ is already in L_{h2}. If so, $\mathscr{C}$ sends $\tau_{h2} = h_2(m_i, PID_i, T_i)$ to $\mathscr{A}$. Or, $\mathscr{C}$ randomly chooses $\tau_{h2} \in Z_q^*$ and then adds $< m_i, PID_i, T_i, \tau_{h2} >$ to L_{h2}. At last, $\mathscr{C}$ sends $\tau_{h2} = h_2(m_i, PID_i, T_i)$ to $\mathscr{A}$.

Sign-Queries. When $\mathscr{A}$ makes sign-query on message m_i, $\mathscr{C}$ randomly chooses $\delta_i, h_i \in Z_p^*$, $PID_i \in G$, and computes $PID_{i,1} = \delta_i \cdot P_{pub} - h_i \cdot P$. Then, $\mathscr{C}$ adds $< m_i, PID_i, T_i, h_i >$ to L_{h2}. At last, $\mathscr{C}$ constructs a message $\{m_i, PID_i, T_i, \delta_i\}$ and sends it to $\mathscr{A}$. According to the rules of the game, each response to the Sign-queries is valid because $\{m_i, PID_i, T_i, \delta_i\}$ answered in the game is able to meet the following equation:

$$\begin{aligned} \delta_i \cdot P_{pub} &= PID_{i,1} + h_i \cdot P = \left(\delta_i \cdot P_{pub} - h_i \cdot P\right) + h_i \cdot P \\ &= \delta_i \cdot P_{pub} \end{aligned} \tag{5}$$

Output. At last, $\mathscr{A}$ outputs $\{m_i, PID_i, T_i, \delta_i\}$ as a valid message with nonnegligible probability. $\mathscr{C}$ can verify the message using

$$\delta_i \cdot P_{pub} = PID_{i,1} + h_i \cdot P \tag{6}$$

If it does not hold, $\mathscr{C}$ terminates this progress.

$\mathscr{A}$ could output $\{m_i, PID_i, T_i, \delta_i^*\}$ as another valid message if $\mathscr{A}$ executes the progress with another h_2-oracle query (let its answer be h_i^*) on the basis of the forgery lemma [38]. Likewise, the message is able to satisfy

$$\delta_i^* \cdot P_{pub} = PID_{i,1} + h_i^* \cdot P \tag{7}$$

According to (6) and (7), we can deduce

$$\begin{aligned} \left(\delta_i - \delta_i^*\right) \cdot x \cdot P &= \left(\delta_i - \delta_i^*\right) \cdot P_{pub} \\ &= \delta_i \cdot P_{pub} - \delta_i^* \cdot P_{pub} \\ &= PID_{i,1} + h_i \cdot P - \left(PID_{i,1} + h_i^* \cdot P\right) \\ &= \left(h_i - h_i^*\right) \cdot P \end{aligned} \tag{8}$$

From (8), we could obtain (9) as follows:

$$\left(\delta_i - \delta_i^*\right) \cdot x = \left(h_i - h_i^*\right) \bmod q \tag{9}$$

Now, $\mathscr{C}$ outputs $\left(\delta_i - \delta_i^*\right)^{-1} \cdot \left(h_i - h_i^*\right)$ as a solution for the given instance of the ECDLP. However, it contradicts with the difficulty of solving the ECDLP. So the proposed scheme can resist forgery attack. □

5.2. Security Analysis. In the subsection, we analyze how the proposed scheme meets the security requirements of VCN.

(1) Message Authentication [39]. In the proposed scheme, an adversary cannot forge a message with nonnegligible probability to meet the verification equation $\delta_i \cdot P_{pub} = PID_{i,1} + h_i \cdot P$ according to Theorem 1. Therefore, a verifier is able to check the validity of message by the verification equation (1). Not that $h_i = h_2(m_i, PID_i, T_i)$ in signature can also be used to check the integrity of message. Therefore, the proposed scheme is able to accomplish signature and integrity verification for VCN.

(2) Conditional Privacy-Preserving (CPP). Vehicle V_i sends message to others with form of $\{m_i, PID_i, T_i, \delta_i\}$, where $PID_{i,1} = r_i \cdot P$, $PID_{i,2} = RID_i \oplus h(r_i \cdot P_{pub})$. The identity of the vehicle is perfectly protected for $PID_{i,2}$ is a pseudoidentity including a random number. To reveal V_i's real identity, an adversary needs to compute $RID_i = PID_{i,2} \oplus h(r_i \cdot P_{pub}) = PID_{i,2} \oplus h(r_i xP)$. However, without knowing r_i and x, the adversary cannot reveal RID_i because it is an instance of CDH problem to compute $w_i xP$. On the contrary, only the TA could reveal the identity from the message by calculating $RID_i = PID_{i,2} \oplus h(x \cdot PID_{i,1})$, if it is necessary. Therefore, the proposed scheme can achieve CPP.

(3) Resistance to Attacks. The proposed scheme can resist the main security attacks of VCN as follows.

(i) Replay Attack. When an attacker launches a replay attack on $\{m_i, PID_i, T_i, \delta_i\}$, it should forge another T_i^* to pass the exam of time freshness. According to Theorem 1, the attacker cannot forge another valid signature δ_i^* to pass message authentication. So this scheme can resist replay attack.

(ii) Modification Attack [40]. As the design of scheme, a valid message consists of its digital signature $\{PID_i, \delta_i\}$. If an attacker makes any modification on the message, the verifier can easily find the modification by verifying (1). Thus, the proposed scheme can resist modification attack.

(iii) Impersonation Attack. An attacker launches an impersonation attack; it should forge a message [41]. However, the probability of the forged message to meet the verification equation can be negligible according to Theorem 1. Therefore, the proposed scheme can resist the impersonation attack.

Table 1: The cryptographic operation and execution time.

Cryptographic operation	Abbr.	Time (ms)
operations related to bilinear pairing	T_{bp}	4.2110
	T_{pm}	1.7090
	T_{ps}	0.0535
	T_{pa}	0.0071
operations related to ECC	T_{em}	0.4420
	T_{es}	0.0138
	T_{ea}	0.0018
Map-To-Point hash function	T_H	4.4060
One-way Hash function	T_h	0.0001

(iv) Verifier Table Attack. As attacks on verifier table become a more and more serious security attack, authentication scheme should focus more attention on these attacks. In the proposed scheme, there is no need for a verifier table in the TA, vehicles, or RSUs. Therefore, an attacker cannot launch any attack on verifier table. Therefore, the proposed scheme can resist the verifier table attack.

6. Performance Analysis and Comparison

In this section, we analyze the performance of the proposed scheme in terms of computation cost and communication cost. The performance comparisons are demonstrated between the proposed scheme and several newly proposed CPP authentication schemes for vehicle network, which are Bayat et al.'s scheme [12] (BAS-CPP, for short), Zhang et al.'s scheme [13] (ZAS-CPP, for short), and He et al.'s scheme [14] (HAS-CPP, for short). Then, the impact on system performance posed by signature aggregation is analyzed. At last, detailed simulations and analysis are shown to evaluate the performance of the proposed scheme according to verification loss rate and message delay.

6.1. Computation Cost Analysis and Comparison. Due to the difference in design, BAS-CPP [12] and ZAS-CPP's [13] cryptographic operations are built on bilinear pairings, while HAS-CPP [14] and our proposed scheme's cryptographic operations are built on ECC. We construct a bilinear pairing cryptography system and an ECC system at 80-bit security level. Table 1 lists the cryptographic operations and corresponding abbreviations and execution times in the four schemes.

Column Abbr. lists the abbreviation of cryptographic operations. Bilinear pairing operation is abbreviated as T_{bp}. Three operations related to bilinear pairing, i.e., scale multiplication, small scale multiplication, and point addition, are abbreviated as T_{pm}, T_{ps}, and T_{pa}, respectively. Three operations related to ECC, i.e., normal scale multiplication, small scale multiplication, and point addition, are abbreviated as T_{em}, T_{es}, and T_{ea}, respectively.

Pseudonym-generating and message signing phase, single message verification phase, and batch message verification phase are called PMS, SMV, and BMV for short.

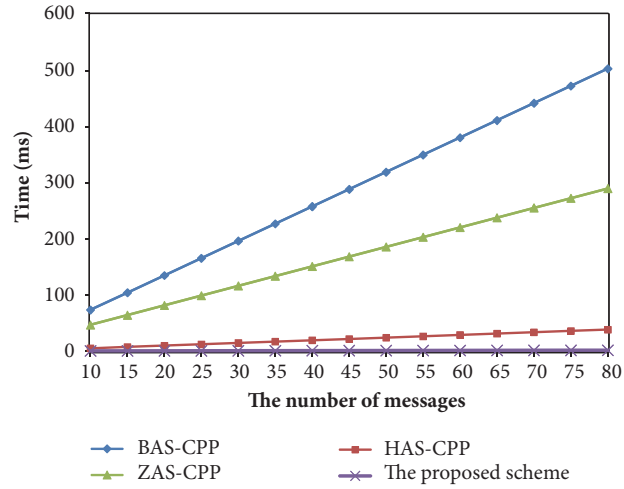


Figure 4: The computation cost comparison of batch verification.

In BAS-CPP [12], the PMS includes five scalar-multiplication operations, one point-addition operation, one Map-To-Point function operation, and two one-way hash operations. The total execution time of BAS-CCP's PMS is $5T_{pm} + 1T_{pa} + 1T_H + 2T_h \approx 12.9583$ ms. The SMV includes three bilinear pairing operations, one point-addition operation, one operation of Map-To-Point function, and one operation of one-way hash function. So the total execution time of BAS-CCP's SMV is $3T_{bp} + 1T_{pa} + 1T_H + 1T_h \approx 18.7481$ ms. The BMV includes three bilinear pairings, $(n + 1)$ operations of scalar multiplication, n small scalar-multiplication operations, $3n - 3$ point-addition operations, and n one-way hash function operations. So the total execution time of BAS-CPP's BMV is $3T_{bp} + (n)T_{pm} + 2nT_{ps} + (3n - 3)T_{pa} + nT_h \approx 6.1364n + 12.6117$ ms. We also can compute ZAS-CPP's [13] computation cost in the same way. For simplicity, the detailed analysis of its computation cost is not presented here.

The PMS of the proposed scheme includes two scalar-multiplication operations and two one-way hash function operations. So the total execution time of PMS in the proposed scheme is $2T_{em} + 2T_h = 0.8842$ ms. The SMV of the proposed scheme includes two scalar-multiplication operations, one point-addition operation, and one one-way hash function operation. So the total execution time of SMV in the proposed scheme is $2T_{em} + 1T_{ea} + 1T_h \approx 0.8859$ ms. The BMV of the proposed scheme includes two scalar-multiplication operations, n small-scalar-multiplication operations, n point-addition operations, and n one-way hash function operations. So the total execution time of BMV in the proposed scheme is $2T_{em} + nT_{es} + nT_{ea} + nT_h \approx 0.0157n + 0.8840$ ms. The cryptographic construction of the HAS-CPP [14] is same as the proposed scheme. For simplicity, the detail analysis of its computation cost is not presented here.

Therefore, we can compute the computation cost of each phase of the four schemes according to Table 1, as shown in Table 2. The result indicates that the proposed scheme has the higher superiority in the computation cost.

Figure 4 illustrates the computation costs of BMV for the different number of messages. As shown in Figure 4, the

Table 2: The computation cost of the four authentication schemes.

	PMS	SMV	BMV
BAS-CPP [12]	12.9583 ms	18.7481 ms	6.1364 n + 12.6117 ms
ZAS-CPP [13]	9.5473 ms	16.0584 ms	3.4715n + 12.6330 ms
HAS-CPP [14]	1.3263 ms	1.3298 ms	0.4752 n + 0.8822 ms
The proposed scheme	0.8842 ms	0.8859 ms	0.0157n + 0.8840 ms
The improvement of the proposed scheme over other schemes			
BAS-CPP [12]	93.17%	95.27%	99.48% (n =50)
ZAS-CPP [13]	90.73%	94.48%	99.10% (n =50)
HAS-CPP [14]	33.33%	33.38%	93.23% (n =50)

Table 3: The comparison of communication cost.

	The component of single message	size
BAS-CPP [12]	$\{m_i, AID_i^1, AID_i^2, T_i, \delta_i\}$	280 bytes
ZAS-CPP [13]	$\{m_i, S_{i,1}, S_{i,2}, STP_{i,j}\}$	276 bytes
HAS-CPP [14]	$\{m_i, AID_{i,1}, AID_{i,2}, T_i, R_i, \delta_i\}$	144 bytes
The proposed scheme	$\{m_i, PID_{i,1}, PID_{i,2}, T_i, \delta_i\}$	84 bytes

proposed scheme is more efficient than the three others in BMV phase regardless of the number of messages

6.2. Communication Cost Analysis and Comparison. In this subsection, the proposed scheme is compared with BAS-CPP [12], ZAS-CPP [13], and HAS-CPP [14] in communication cost. According to the definition in previous section, the size of a bilinear pairing group element is 128 bytes, and the size of an ECC system group element is 40 bytes. Let the sizes of a timestamp and a one-way hash output be 4 and 20 bytes. Here we do not consider original content in message for it is the same to all schemes. According to the component of single message of the four schemes, Table 3 shows their communication costs. Obviously, compared with BAS-CPP, ZAS-CPP, and HAS-CPP, the proposed scheme requires less communication cost.

6.3. Signature Aggregation Analysis. In this subsection, we show the performance improvement of signature aggregation over traditional ways, i.e., forwarding message one by one.

BAS-CPP [12] and HAS-CPP [14] do not offer signature aggregation. Different from them, the proposed scheme and ZAS-CPP [13] provide signature aggregation. As shown in message verification phase in Section 4, after the verifier has checked n messages, the verifier forwards the n messages to top managers one by one. To decrease communication and computation cost, the verifier can aggregate multiple signatures into a single one, i.e., the verifier could make n messages into an aggregated signature $M_A^* = \{(m_1, PID_1, T_1), (m_2, PID_2, T_2), \ldots, (m_n, PID_n, T_n), \delta_A^*\}$, where the size of δ_A^* in M_A^* is identical to the size of δ_i in a single message $\{m_i, PID_i, T_i, \delta_i\}$, regardless of the number of messages. During forwarding 50 messages to top managers, the verifier in our scheme can decrease communication cost by 1000 bytes using signature aggregation compared to using traditional way, details shown in Figure 5. As far as signature aggregation is concerned, ZAS-CPP [13] can decrease more communication cost, though it needs more sign and verification cost. Therefore, our scheme and ZAS-CPP [13] can further decrease communication cost by signature aggregation.

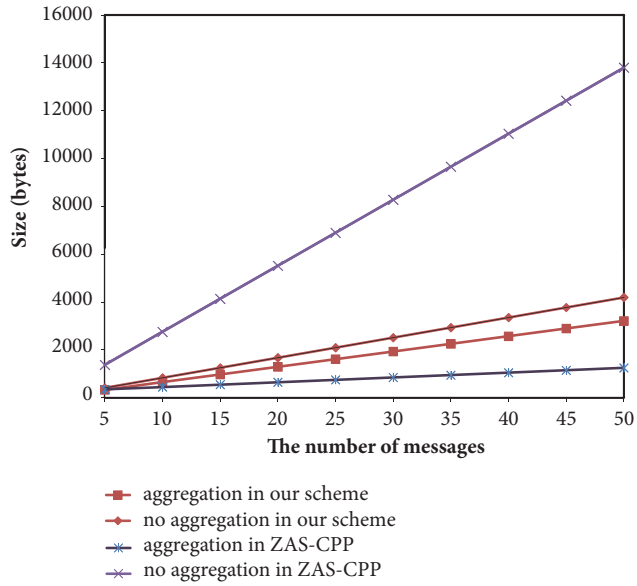


Figure 5: The communication cost comparison of signature aggregation.

From the above performance analysis and comparison, it is easy to draw a conclusion that the proposed scheme has more advantages. Compared with BAS-CPP and HAS-CPP, the proposed scheme not only has less computation and communication cost in message signing phase, single message verification phase, and batch message verification phase, but also decreases communications cost by signature aggregation. Compared with ZAS-CPP, although the proposed scheme is insufficient in signature aggregation, it has a great advantage in computation and communication cost in signing phase and verification phase. Table 4 shows the comprehensive comparison results of the four schemes in terms of the computation costs of PMS, SMV, and BMV, the communication cost (C-cost for short), and the signature aggregation functionality (SA-func for short). It obviously shows that the proposed scheme has most advantages. Therefore, the proposed scheme can further satisfy the requirements of VCN.

Table 4: The comprehensive comparison results of the four schemes.

scheme name	PMS	SMV	BMV	C-Cost	SA-func
BAS-CPP [12]	High	High	High	High	No
ZAS-CPP [13]	High	High	High	High	Yes
HAS-CPP [14]	Middle	Middle	Middle	Middle	No
The proposed scheme	Low	Low	Low	low	Yes

Figure 6: The simulation scenario.

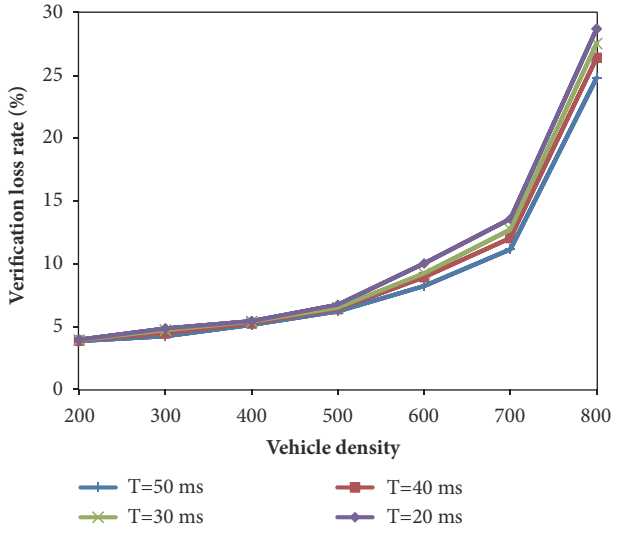


Figure 7: Verification loss rate related to vehicle density and interval.

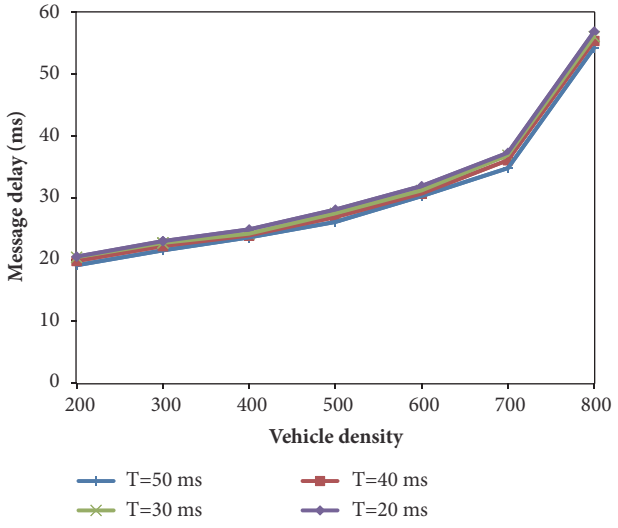


Figure 8: Message delay related to vehicle density and interval.

6.4. Simulation and Analysis. In this section, we evaluate the performance of the proposed scheme by several simulations. The simulation scenarios are constructed in the Veins framework [42] and the OMNeT++ simulation platform [43] with the surrounding roads of Wuhan University, as shown in Figure 6, where all roads are two-way multilane. The main goal of this simulation is to test the advantages and disadvantages of the proposed scheme in terms of loss rate and message delay.

In the simulation, one RSU is deployed every 2 km along the roads, and it can send messages to vehicles within 800 m; vehicles run along roads and communicate with others within 250 m. Let each vehicle generate a traffic message every 300 ms and send it to RSUs and other vehicles; then RSUs verify and aggregate the messages to cloud sever. Let the size of a message be 200 bytes, the wired communication bandwidth between RSUs and cloud server is 10 mb/s, and the wireless communication bandwidth between vehicles is 200 kb/s. The vehicle density (the number of vehicles in the scenario) in the scene is set between 200 and 800. Let 2% vehicles be malicious ones that have invalid signature messages. The speed of vehicles is randomly generated by the system in a normal distribution between 40 and 90 km/h.

In order to test the impact of batch authentication time interval setting on the proposed scheme, four batch verification simulations with different intervals are designed, where the intervals T are 20 ms, 30 ms, 40 ms, and 50 ms. The verification loss rate and message delay during the simulations are shown in Figures 7 and 8.

The T in Figure 7 denotes the interval for batch verification, and the verification loss rate has a certain function with vehicle density under different T. It shows that the greater the vehicle density, the greater the communication overhead of the whole system. Meanwhile, the verification loss rate is rising as communication overhead is rising under any T. Of course, as T decreases, the verification loss rate of the proposed scheme increases, but its increase is in a smaller range.

Figure 8 shows the relationship between message delay and vehicle density in the proposed scheme. It shows that the greater the vehicle density is, the greater the communication overhead is, which results in adding the instability of the

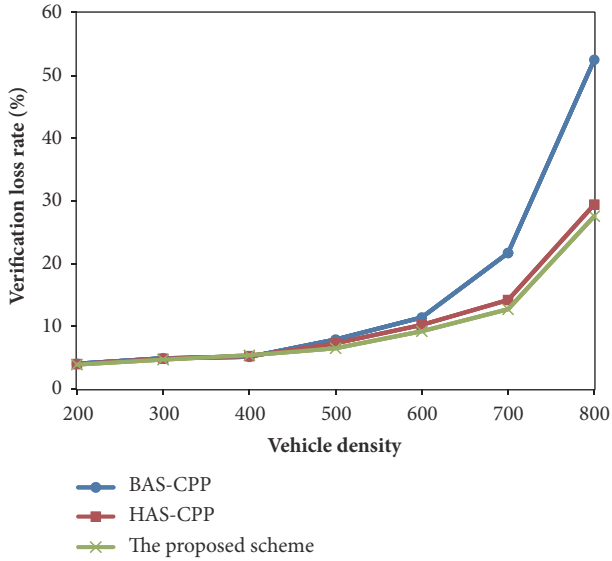


Figure 9: The comparison of verification loss rate among three schemes.

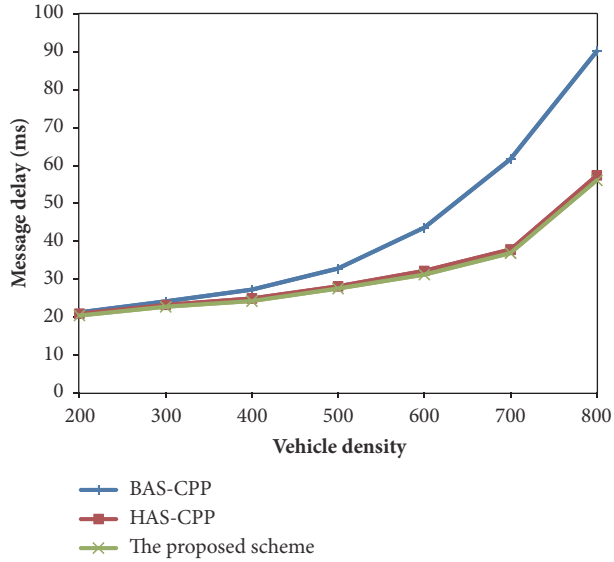


Figure 10: The comparison of message delay among three schemes.

communication system. Therefore, message delay is rising as vehicle density is rising under any T. However, the message delay increases slightly as T decreases.

Next, the comparison simulations are executed among the proposed scheme, BAS-CPP [12], and HAS-CPP [14] in terms of verification loss rate and message delay. In these simulations, $T = 30$ ms. Figure 9 shows the comparison of verification loss rate among three schemes in the simulations. As can be seen from Figure 9, as the vehicle density increases, the message loss rate of the three schemes increases. The verification loss rate of BAS-CPP is increasing rapidly, and the rates of HAS-CPP and the proposed scheme are relatively slow, which could prove that the improved message verification efficiency can improve the speed of receiving and processing messages and reduce the loss rate.

Figure 10 shows the comparison of message delay among three schemes. As the vehicle density increases, the message delay of the proposed scheme and HAS-CPP increases, but the delay growth rate is smaller than BAS-CPP. The simulation results further prove that the proposed scheme can reduce the message delay and improve the performance of the VCN system.

7. Conclusion

A new efficient message authentication scheme for VCN is presented in this paper, and it achieves conditional privacy-preserving. In order to solve urgent authentication issue for life-critical message in VCN, batch message verification and signature aggregation are included in the proposed scheme, which is suitable for VCN because verifiers are limited in computation capacity and communication channel is very strained in VCN. The security proof and analysis show that the proposed scheme could satisfy the security requirements of VCN. The performance analyses show that the proposed scheme has obvious advantages in decreasing communication and computation cost when compared with recent proposed identity-based authentication schemes. A detailed simulations and analysis are shown to evaluate the performance of the proposed scheme according to verification loss rate and message delay, which prove that the proposed scheme can reduce verification loss rate and message delay, and improve the performance of the VCN system.

Our next research will focus on improving the signature aggregation to decrease more communication cost while keeping the efficiency of signature and verification.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported in part by the National Natural Science Foundation of China under Grant 61862052, the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (17YJCZH203), and the Hubei Provincial Department of Education research projects (D20182702).

References

[1] M. S. Kakkasageri and S. S. Manvi, "Information management in vehicular ad hoc networks: a review," *Journal of Network and Computer Applications*, vol. 39, no. 1, pp. 334–350, 2014.

[2] H. Tan, D. Choi, P. Kim, S. Pan, and I. Chung, "Comments on 'dual authentication and key management techniques for

secure data transmission in vehicular ad hoc networks,'" *IEEE Transactions on Intelligent Transportation Systems*, 2017.

[3] R. Yu, Y. Zhang, S. Gjessing, W. Xia, and K. Yang, "Toward Cloud-based vehicular networks with efficient resource management," *IEEE Network*, vol. 27, no. 5, pp. 48–55, 2013.

[4] C.-C. Lee and Y.-M. Lai, "Toward a secure batch verification with group testing for VANET," *Wireless Networks*, vol. 19, no. 6, pp. 1441–1449, 2013.

[5] J. Wang, J. Cho, S. Lee, and T. Ma, "Real time services for future cloud computing enabled vehicle networks," in *Proceedings of the International Conference on Wireless Communications and Signal Processing (WCSP '11)*, pp. 1–5, November 2011.

[6] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya, "A survey on vehicular cloud computing," *Journal of Network and Computer Applications*, vol. 40, no. 1, pp. 325–344, 2014.

[7] J. Cui, J. Zhang, H. Zhong, and Y. Xu, "SPACF: A secure privacy-preserving authentication scheme for VANET with cuckoo filter," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10283–10295, 2017.

[8] D. Hughes and V. Shmatikov, "Information hiding, anonymity and privacy: A modular approach," *Journal of Computer Security*, vol. 12, no. 1, pp. 3–36, 2004.

[9] D. Wang, H. Cheng, H. Debiao, and P. Wang, "On the challenges in designing identity-based privacy-preserving authentication schemes for mobile devices," *IEEE Systems Journal*, vol. 12, no. 1, pp. 916–925, 2018.

[10] S. S. Manvi and S. Tangade, "A survey on authentication schemes in VANETs for secured communication," *Vehicular Communications*, vol. 9, pp. 19–30, 2017.

[11] T. Gao, X. Deng, N. Guo, and X. Wang, "An anonymous authentication scheme based on PMIPv6 for VANETs," *IEEE Access*, vol. 6, pp. 14686–14698, 2018.

[12] M. Bayat, M. Barmshoory, M. Rahimi, and M. R. Aref, "A secure authentication scheme for VANETs with batch verification," *Wireless Networks*, vol. 21, no. 5, pp. 1733–1743, 2015.

[13] L. Zhang, C. Hu, Q. Wu, J. Domingo-Ferrer, and B. Qin, "Privacy-preserving vehicular communication authentication with hierarchical aggregation and fast response," *IEEE Transactions on Computers*, vol. 65, no. 8, pp. 2562–2574, 2016.

[14] 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.

[15] X. Lin, X. Sun, P.-H. Ho, and X. Shen, "GSIS: a secure and privacy-preserving protocol for vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3442–3456, 2007.

[16] R. Lu, X. Lin, H. Zhu, P.-H. Ho, and X. Shen, "ECPP: efficient conditional privacy preservation protocol for secure vehicular communications," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 1229–1237, April 2008.

[17] L. Zhang, Q. Wu, A. Solanas, and J. Domingo-Ferrer, "A scalable robust authentication protocol for secure vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1606–1617, 2010.

[18] Y. Liu, W. Guo, Q. Zhong, and G. Yao, "LVAP: Lightweight V2I authentication protocol using group communication in VANETs," *International Journal of Communication Systems*, vol. 30, no. 16, 2017.

[19] L. Zhang, Q. Wu, J. Domingo-Ferrer, B. Qin, and C. Hu, "Distributed aggregate privacy-preserving authentication in VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 516–526, 2017.

[20] M. Raya and J.-P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[21] J. Freudiger, R. Maxim, M. Félegyházi, P. Papadimitratos, and H. Jean-Pierre, "Mix-zones for location privacy in vehicular networks," in *Proceedings of the ACM Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS '07)*, number LCA-CONF-2007-016, 2007.

[22] C. Zhang, X. Lin, R. Lu, and P.-H. Ho, "RAISE: an efficient RSU-aided message authentication scheme in vehicular communication networks," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 1451–1457, May 2008.

[23] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Proceedings of the Workshop on the Theory and Application of Cryptographic Techniques (CRYPTO '84)*, vol. 196, pp. 47–53, Springer, 1984.

[24] C. Zhang, R. Lu, X. Lin, P.-H. Ho, and X. Shen, "An efficient identity-based batch verification scheme for vehicular sensor networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 246–250, April 2008.

[25] T. W. Chim, S. M. Yiu, L. C. K. Hui, and V. O. K. Li, "SPECS: Secure and privacy enhancing communications schemes for VANETs," *Ad Hoc Networks*, vol. 9, no. 2, pp. 189–203, 2011.

[26] Y. Xie, L. Wu, J. Shen, and A. Alelaiwi, "EIAS-CP: new efficient identity-based authentication scheme with conditional privacy-preserving for VANETs," *Telecommunication Systems*, vol. 65, no. 2, pp. 229–240, 2017.

[27] J. Kang, D. Lin, W. Jiang, and E. Bertino, "Highly efficient randomized authentication in VANETs," *Pervasive and Mobile Computing*, vol. 44, pp. 31–44, 2018.

[28] Y. Liu, Y. Wang, and G. Chang, "Efficient privacy-preserving dual authentication and key agreement scheme for secure V2V communications in an IoV paradigm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2740–2749, 2017.

[29] A. Wasef and X. Shen, "ASIC: Aggregate signatures and certificates verification scheme for vehicular networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '09)*, pp. 1–6, December 2009.

[30] R. W. Van Der Heijden, S. Dietzel, and F. Kargl, "SeDyA: Secure dynamic aggregation in VANETs," in *Proceedings of the 6th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '13)*, pp. 131–142, April 2013.

[31] L. Wu, J. Fan, Y. Xie, J. Wang, and Q. Liu, "Efficient location-based conditional privacy-preserving authentication scheme for vehicle ad hoc networks," *International Journal of Distributed Sensor Networks*, vol. 13, no. 3, 2017.

[32] E. Lee, E.-K. Lee, M. Gerla, and S. Y. Oh, "Vehicular cloud networking: architecture and design principles," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 148–155, 2014.

[33] 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.

[34] 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*, 2016.

[35] Q. Jiang, J. Ni, J. Ma, L. Yang, and X. Shen, "Integrated authentication and key agreement framework for vehicular cloud computing," *IEEE Network*, vol. 32, no. 3, pp. 28–35, 2018.

[36] X. Hu, J. Zhang, Z. Zhang, and F. Liu, "Anonymous password authenticated key exchange protocol in the standard model," *Wireless Personal Communications*, vol. 96, no. 1, pp. 1451–1474, 2017.

[37] J.-L. Huang, L.-Y. Yeh, and H.-Y. Chien, "ABAKA: an anonymous batch authenticated and key agreement scheme for value-added services in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 248–262, 2011.

[38] D. Hankerson, S. Vanstone, and A. J. Menezes, *Guide to Elliptic Curve Cryptography*, Springer, New York, NY, USA, 2004.

[39] Q. Jiang, C. Zhiren, L. Bingyan, J. Shen, L. Yang, and M. Jianfeng, "Security analysis and improvement of bio-hashing based three-factor authentication scheme for telecare medical information systems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 4, pp. 1061–1073, 2018.

[40] X. Hu, J. Zhang, Z. Zhang, and J. Xu, "Universally composable anonymous password authenticated key exchange," *Science China Information Sciences*, vol. 60, no. 5, 2017.

[41] Q. Jiang, J. Ma, C. Yang, X. Ma, J. Shen, and S. A. Chaudhry, "Efficient end-to-end authentication protocol for wearable health monitoring systems," *Computers and Electrical Engineering*, vol. 63, pp. 182–195, 2017.

[42] M. Segata, S. Joerer, B. Bloessl, C. Sommer, F. Dressler, and R. L. Cigno, "Plexe: A platooning extension for Veins," in *Proceedings of the IEEE Vehicular Networking Conference (VNC '14)*, pp. 53–60, Paderborn, Germany, December 2014.

[43] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops (ICST '08)*, pp. 60–69, March 2008.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 3284324, 14 pages
https://doi.org/10.1155/2018/3284324

Research Article

A Secure Three-Factor Multiserver Authentication Protocol against the Honest-But-Curious Servers

Hua Guo,[1,2] Chen Chen,[3] Ya Gao,[4] Xiong Li,[5,6] and Jiongchao Jin[7]

[1]*School of Cyber Science and Technology, Beihang University, Beijing 100191, China*
[2]*Hefei Innovation Institute, Beihang University, Anhui 230012, China*
[3]*Informatization Office of Beihang University, Beijing 100191, China*
[4]*Beijing Key Laboratory of Network Technology, Beihang University, Beijing 100191, China*
[5]*School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China*
[6]*Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China*
[7]*School of Computer Science and Engineering, Beihang University, Beijing 100191, China*

Correspondence should be addressed to Xiong Li; lixiongzhq@163.com

Received 13 April 2018; Revised 26 July 2018; Accepted 26 August 2018; Published 13 September 2018

Academic Editor: Ding Wang

Three-factor multiserver authentication protocols become a prevalence in recent years. Among these protocols, almost all of them do not involve the registration center into the authentication process. To improve the protocol's efficiency, a common secret key is shared among all severs, which leads to a serious weakness; i.e., we find that these protocols cannot resist the passive attack from the honest-but-curious servers. This paper takes Wang *et al.*'s protocol as an example, to exhibit how an honest-but-curious server attacks their protocol. To remedy this weakness, a novel three-factor multiserver authentication protocol is presented. By introducing the registration center into the authentication process, the new protocol can resist the passive attack from the honest-but-curious servers. Security analyses including formal and informal analyses are given, demonstrating the correctness and validity of the new protocol. Compared with related protocols, the new protocol possesses more secure properties and more practical functionalities than others at a relatively low computation cost and communication cost.

1. Introduction

Nowadays, with the rapid development of networks, remote communication becomes increasingly prevalent and provides highly useful services in many aspects. Consequently, communication security significantly attracts public's attention. Cryptographic authentication allows users to submit their credentials and acquire authorization to access the various online services from remote networks [1–5]. Since Lamport [6] firstly proposed a password-based remote authentication protocol, great quantities of authentication protocols were proposed to make up continued emerging problems and provide authorized communication between remote entities. However, the traditional protocols gradually cannot catch up with the pace of increasing demand for more users and servers in communication. Multiserver authentication schemes became the mainstream, because most of the practical communication environments are based on several servers to alleviate the pressure of the increasing number of users.

Lots of authentication protocols for multiserver environments were proposed to satisfy the security requirements and provide versatile functionalities to make the scheme more convenient and practical to utilize in real occasions [7–20]. In 2001, Li *et al.* [7] proposed a remote multiserver authentication protocol with no verification table, which was found insecure by Lin *et al.* [8]. They also presented an improved protocol, while it was vulnerable to impersonation attack [9]. Juang *et al.* [10] adopted symmetric-key cryptosystem to propose a multiserver authentication protocol but it was cracked soon. In 2004, a novel protocol was presented by Chang and Lee [11]. However, all of them ignored user

anonymity [12]. In 2009, a remote multiserver authentication scheme which satisfies anonymity property was proposed [13], but it does not have forward security [14]. Besides, Hsiang and Shih [15] presented a new protocol to resist various attacks; however some drawbacks on mutual authentication are pointed out [16]. Recently, a big breakthrough, i.e., the inner relationships of evaluation criteria for anonymous two-factor authentication protocol, is explored by Wang *et al.* [17]. To improve the security of remote communication, smart card gradually came into use in authentication, which made it possible for more convenient authentication and communication. Some remote authentication protocols for multiserver environment with a smart card were proposed but proved to be insecure in the end [21–23].

Relying on smart card and password as the authentication method already cannot meet today's needs. In the latest few years, more and more authentication protocols adopt biometrics messages in mutual authentication to strengthen the security and enhance the efficiency of the existing protocols. In 2010, Yang *et al.* [24] introduced a three-factor multiserver authentication protocol. Unfortunately, the protocol has low computation efficiency and can not resist the insider attack. Li *et al.* proposed an efficient protocol [25] which allows users to change the password and the calculation cost is low. However, their scheme cannot provide appropriate certification and failed to resist man-in-the-middle attack [26]. Adopting elliptic curve cryptography, Yoon *et al.* [27] in 2011 designed a novel protocol; unfortunately Kim *et al.* [28] showed that Yoon *et al.* [27] protocol is insecure. In 2014, Chuang *et al.* [29] put forward an anonymous protocol, but Mishra *et al.* [30] broke their protocol. Later, Lu *et al.* [31] found that there are several weaknesses in Mishra *et al.*'s improved protocol, and they also presented an improved protocol which is broken by Reddy *et al.* [32]. Meanwhile, Wang *et al.* [33] also found that Mishra *et al.*'s improved protocol is insecure. Regrettably, some weaknesses in Wang *et al.*'s protocol [33] were shown by Yang *et al.* [24] and Reddy *et al.* [34] separately. Recently, Jiang *et al.* [35] and He *et al.* put forward multiserver authenticated protocols using elliptic curve cryptography(ECC), separately. Unfortunately, Odelu *et al.* found that there are flaws in He *et al.*s protocol in login and password change phases and can not resist the impersonation attack.

All of the above three-factor multiserver authentication protocols can be categorized into two classes, i.e., the protocols which implement the authentication independent of the registration center and the protocols which need the help of the registration center in the authentication phase. After carefully examining the known three-factor multiserver authentication protocols, we find that almost all of the first kind of protocols cannot resist the passive attack from an honest-but-curious server since all servers share a common secret key. More precisely, an honest-but-curious server can compute session keys which are shared between a user and other servers by eavesdropping messages transmitting between the users and other servers. In this paper, we take Wang *et al.*'s protocol as an example, to show how an honest-but-curious server obtains a session key which should be kept secret from him. Moreover, we find some other drawbacks in their protocol. For example, in the reregister or revocation phase, a user can still use his original password to login and send message even if he is revoked.

To resist the passive attack from the honest-but-curious servers, a trivial solution is to distribute different secret keys to different severs, which would aggravate the user's storage burden. Another method is introducing the registration center into the authentication phase to deal with secret messages. As we mentioned above, such protocols are based on either ECC or symmetric encryption cryptosystems, which heavily affect the computation efficiency. To balance the security problem brought by the honest-but-curious servers and the efficiency problem brought by involving the registration center into the authentication, we propose a novel multiserver authentication protocol. In authentication phase of new protocol, the involved registration center only adopts hash and XOR operations for the computation, instead of ECC and symmetric cryptosystem, thus greatly improving the protocol's computation efficiency. As far as we known, this is the first time to consider the passive attack from honest-but-curious servers for multiserver authentication protocols. Moreover, the new protocol is the first protocol which only adopts hash and XOR operations for computation when involving the registration center into the authentication.

The remaining of the paper is organized as follows. Section 2 reviews and analyzes the security of Wang *et al.*'s protocol. In Section 3, we present the new three-factor multiserver authentication scheme in detail. Section 4 provides the formal and informal secure analysis of the new protocol. In Section 5, comparisons including security, functionalities, computation cost, and communication cost are conducted. The last section gives a conclusion.

2. Some Weakness of Wang *et al.*'s Scheme

We firstly give the details about Wang *et al.*'s protocol and then show how an honest-but-curious server attacks their protocol step by step.

2.1. Review of Wang et al.'s Protocol. Wang *et al.*'s protocol involves five phases, i.e., registration phase, login phase, authentication phase, password changing phase, and revocation/reregistration phase which are executed by the user U_i, the server S_j, and the registration center RC. The symbols and notations are listed in Table 1. Assume RC is a trusted third party which is able to register for users and servers.

Registration Phase

(i) Registration phase of server

(a) S_j sends a request message to RC.
(b) RC authorizes S_j once it receives the message and returns PSK (preshared key) to S_j securely.
(c) S_j uses PSK to check U_i's legitimacy in authentication phase.

(ii) Registration phase of user

(a) The new user U_i inserts SC_i into the card reader, inputs ID_i, PW_i, and imprints B_i at the

Table 1: Symbols and notations in Wang *et al.*'s scheme.

U_i	the ith user
S_j	the jth server
RC	registration centre
ID_i	U_i's identity
AID_i	U_i's dynamic identity
SID_j	S_j's identity
SC_i	U_i's smart card
PW_i	U_i's password
BIO_i	U_i's biometrics
R_i	U_i's nearly random binary string
PSK	pre-shared key between server and registration centre
x	master secret key between user and registration centre
$\oplus$	exclusive OR
$h(\cdot)$	one-way hash function
$\|\|$	concatenating operation

sensor. After that, (R_i, P_i) is extracted from BIO_i through $Gen(BIO_i) \longleftarrow (R_i, P_i)$. Finally U_i calculates $RPW_i = h(PW_i \| R_i)$ and sends $\{ID_i RPW_i\}$ to RC securely.

(b) RC generates $< ID_i, N_i = 1 >$ and stores it to the database. Note that N_i indicates the state of U_i's account. When U_i revokes his account, RC sets $N_i = 0$. When U_i reregisters his account, RC sets $N_i = N_i + 1$. After that RC calculates $A_i = h(ID_i \| x \| T_r)$, $B_i = RPW_i \oplus h(A_i)$, $C_i = B_i \oplus h(PSK)$, $D_i = PSK \oplus A_i \oplus h(PSK)$, and $V_i = h(ID_i \| RPW_i)$ where T_r is the time of registration. Finally, RC sends $\{B_i, C_i, D_i, V_i\}$ to U_i securely.

(c) U_i receives $\{B_i, C_i, D_i, V_i\}$ from RC, stores $\{B_i, C_i, D_i, V_i, P_i\}$ into SC_i, and stores P_i in SC_i.

Login Phase

(i) U_i inputs ID_i and PW_i with his/her smart card SC_i and imprints BIO_i^* at the sensor.

(ii) SC_i calculates $R_i = Rep(BIO_i^*, P_i)$ and $RPW_i = h(PW_i \| R_i)$. After that, SC_i checks whether $h(ID_i \| RPW_i) = V_i$ holds or not. If it is right, SC_i computes $h(PSK) = B_i \oplus C_i$.

(iii) SC_i calculates $AID_i = h(N_1) \oplus ID_i$ after choosing a random number N_1. After that, SC_i computes $M_1 = RPW_i \oplus h(PSK) \oplus N_1$ and $M_2 = h(AID_i \| N_1 \| RPW_i \| SID_j \| T_i)$ where T_i is a timestamp.

(iv) SC_i sends $\{B_i, D_i, M_1, M_2, AID_i, T_i\}$ to S_j.

Authentication Phase

(i) S_j checks whether $T_i - T_j \leq \Delta T$. Note that T_j is the time that S_j receives the login message, and ΔT means the time interval.

(ii) If the verification is valid, S_j calculates $A_i = PSK \oplus D_i \oplus h(PSK)$, $RPW_i = h(A_i) \oplus B_i$, and $N_1 = M_1 \oplus RPW_i \oplus h(PSK)$ and checks whether $h(AID_i \| N_1 \| RPW_i \| SID_j \| T_i) = M_2$.

(iii) If this verification is valid, S_j chooses a random number N_2 and calculates $SK_{ij} = h(AID_i \| SID_j \| N_1 \| N_2)$ as the session secret key. Then, S_j computes $M_3 = h(PSK) \oplus h(AID_i \| N_1) \oplus N_2$ and $M_4 = h(SID_j \| N_2 \| AID_i)$. Finally S_j returns $\{M_3 M_4, SID_j\}$ to U_i.

(iv) SC_i calculates $N_2 = h(PSK) \oplus h(AID_i \| N_1) \oplus M_3$ and $SK_{ij} = h(AID_i \| SID_j \| N_1 \| N_2)$ and checks whether $h(SID_j \| N_2 \| AID_i) = M_4$. If it is valid, SC_i calculates $M_5 = h(SK_{ij} \| N_1 \| N_2)$ and sends M_5 to S_j.

(v) S_j checks whether the condition $h(SK_{ij} \| N_1 \| N_2)$ matches with M_5. If holds, S_j confirms the session key SK_{ij}. Otherwise, S_j terminates the session immediately.

Password Change Phase

(i) U_i inserts his/her smart card, inputs ID_i and PW_i, and imprints BIO_i^*.

(ii) SC_i retires R_i from $Rep(BIO_i^*, P_i)$ and computes $RPW_i = h(PW_i \| R_i)$. Then SC_i checks whether $h(ID_i \| RPW_i)$ matches with V_i. If it holds, U_i can input the new password.

(iii) U_i inputs the new password PW_i^n and SC_i calculates $RPW_i^n = h(PW_i^n \| R_i)$, $B_i^n = B_i \oplus RPW_i \oplus RPW_i^n$, $C_i^n = C_i \oplus RPW_i \oplus RPW_i^n$, and $V_i^n = h(ID_i \| RPW_i^n)$.

(iv) SC_i displaces B_i with B_i^n, C_i with C_i^n, and V_i with V_i^n, respectively.

Revocation and Reregistration Phase

(i) If U_i is revoked, he needs to send verification message RPW_i to RC securely.

(ii) RC checks the validity of U_i. If U_i is a valid user, RC sets $< ID_i, N_i = 0 >$.

(iii) *RC* follows the user registration phase and uses $< ID_i, N_i = N_i + 1 >$ to replace $< ID_i, N_i >$.

2.2. Analysis of Wang et al.'s Protocol. This subsection analyzes Wang *et al.*'s protocol and shows how to mount a passive attack by an honest-but-curious server.

2.2.1. Passive Attack from an Honest-But-Curious Server. In this attack, an honest-but-curious server (say S_j) only passively eavesdrops messages between the user U_i and other servers, so that he can obtain the session keys shared by the user U_i and other servers which should be kept secret from S_j using his secret key and eavesdropping messages. More precisely, suppose a user U_i has finished the protocol with a server S_1 and is running the protocol with the other server S_2. Now we will show how the server S_1 obtains the session key between U_i and S_2 step by step.

(i) Step 1. S_1 finished the protocol with U_i successfully. Thus the server has knowledge of *PSK*.

(ii) Step 2. During the protocol process between U_i and S_2, S_1 firstly intercepts $\{B_i, D_i, M_1, M_2, AID_i, T_i\}$ sent by U_i to S_2. From these messages, S_1 obtains N_1 by calculating $A_i = D_i \oplus PSK \oplus h(PSK)$, $RPW_i = B_i \oplus h(A_i)$, and $N_1 = RPW_i \oplus M_1 \oplus h(PSK)$.

(iii) Step 3. After that, S_1 intercepts the messages $\{SID_j, M_3 M_4\}$ which are sent from S_2 to U_i. Then the server S_1 obtains N_2 by calculating $N_2 = M_3 \oplus h(AID_i \parallel N_1) \oplus h(PSK)$. In this case, the server S_1 acquired N_1 and N_2 generated in this session.

(iv) Step 4. With the intercepted AID_i and SID_j, S_1 can obtain the session key SK_{ij} by calculating $SK_{ij} = h(AID_i \parallel SID_j \parallel N_1 \parallel N_2)$ successfully which should be kept secret from him.

2.2.2. User's Anonymity. User's anonymity means that user's ID_i and other urgent information indicating user's identity directly should be protected carefully. In Wang *et al.*'s scheme, an honest-but-curious server S_j can compute N_1 by $N_1 = RPW_i \oplus M_1 \oplus h(PSK)$. At the same time, S_j receives AID_i from U_i. S_j can obtain U_i's identity ID_i by computing $ID_i = AID_i \oplus h(N_1)$. As a consequence, the server can obtain ID_i of the user U_i. This does not guarantee the anonymity of the user's identity.

2.2.3. User Impersonation Attack. The honest-but-curious server S_j can collect B_i and D_i which are sent by U_i and thus can calculate RPW_i using *PSK*. After that, S_j can pretend to be U_i and apply authentication from other servers. Specifically, S_j randomly choose a number N_1 and calculates $AID_i = h(N_1) \oplus ID_i$, $M_1 = RPW_i \oplus N_1 \oplus h(PSK)$, and $M_2 = h(AID_i \parallel N_1 \parallel RPW_i \parallel SID_j \parallel T_i)$. Then S_j sends $\{B_i, D_i, M_1, M_2, AID_i, T_i\}$ to other servers through a public channel. In this way, S_j can disguise as U_i.

2.2.4. Wrong Revocation and Reregistration. In this phase, users are allowed to revoke or reregister when he confronts the situation about losing the smart card or his account. In Wang *et al.*'s scheme, when a user wants to revoke his account, he has to pass the authentication. After that, *RC* changes N_i to 0, indicating that ID_i is not available any more. In the reregistration phase, *RC* also changes N_i to $N_i + 1$. However, *RC* is not involved in login phase and authentication phase. As a result, there is no access for *RC* to check whether the user's account is revoked or not. Thus the user can access the legal servers only using his former password and his biometrics; even he already has been revoked.

2.3. Reasons for the Weakness. In Wang *et al.*'s protocol, two important temporary secret values N_1 and N_2 are protected by *PSK*. Unfortunately, all servers keep the same private key *PSK*. As a result, an honest-but-curious server S_j with *PSK* can obtain all session keys which should be kept secret from him. This attack usually exists in the multiserver environment. In the most cases, a legitimate server after registration is assumed to be completely trustworthy, without taking into account the possibility that a particular server can act as an honest-but-curious adversary.

To resist this attack, it is bound to distribute different secret keys to different servers, which can be implemented by involving the registration center in the authentication process. Unfortunately, in most of this kind of protocol, authentication process is excused between the user and the server independent of the registration center. Therefore, to design a secure three-factor authentication protocol against the passive attack from an honest-but-curious server, the registration center should be introduced into the authentication process to protect the important temporary secret values N_1 and N_2.

3. The New Protocol

In this section, we first discuss the threat model used in our protocol. We then give the list of notations used in our proposed scheme. Finally, we describe the different phases relate to our scheme.

3.1. Threat Model. In this subsection, we introduce a threat model following the definition of [36–39].

(i) The adversary $\mathcal{A}$ is able to control the open communication channel completely; that is, he can intercept, modify, delete, block, and resend the messages over the open channel.

(ii) The adversary $\mathcal{A}$ can list all pairs of ID_i from the space of identities and PW_i from the space of passwords in a polynomial time.

(iii) The adversary $\mathcal{A}$ can either intercept the password of the user via the malicious device or extract the parameters from the smart card, but both methods cannot be used together. An honest-but-curious server does not have this ability.

(iv) When $\mathcal{A}$ acts as an honest-but-curious server, he can just listen the messages via the open channel.

Table 2: Notations in the new protocol.

U_i, S_j	ith user and jth server
RC	the registration centre
ID_i, AID_i, SID_j	U_i's identity, dynamic identity and S_j's identity
SC_i, PW_i, BIO_i	U_i's smart card, password and biometrics
R_i, P_i	valid biometric characteristic extracted from BIO_i
PSK	master secret key between server and registration centre
X	master secret key between user and registration centre
$h(\cdot), \oplus, \|\|$	hash function, XOR operation and concatenating operation
x_i	pre-shared key between U_i and registration centre
y_j	pre-shared key between S_j and registration centre

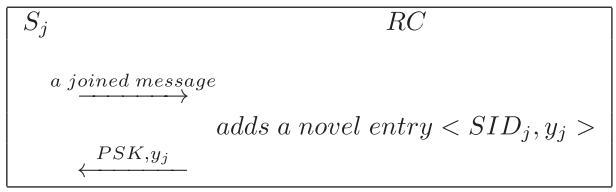


Figure 1: Registration phase of server.

3.2. The Proposed Protocol. As Wang *et al.*'s protocol, the new protocol also involves five phases. Table 2 lists the notations used in the new protocol.

3.2.1. Registration Phase

(i) Server registration phase

During server registration, S_j communicates with RC to authenticate his validity and become a legislative server after receiving the preshared key sent from RC. The whole process of the server registration phase is shown in Figure 1.

(a) S_j sends a request message to RC.

(b) RC authorizes S_j and adds a novel entry $< SID_j, y_j >$ to the database where y_j is a random number. Then $h(PSK), y_j$ is sent to S_j by applying IKEv2 securely.

(c) S_j adopts PSK, y_j to protect the urgent messages and generates the session key SK_{ij}.

(ii) User registration phase

A user sends his personal information to RC and gets his own smart card by executing the process listed in Figure 2.

(a) U_i inputs BIO_i at the sensor and can obtain $\{R_i, P_i\}$ using $Gen(BIO_i) \longrightarrow (R_i, P_i)$. Then U_i selects ID_i and PW_i and calculates $RPW_i = h(PW_i \parallel R_i)$. U_i finally sends $\{ID_i RPW_i\}$ to RC securely.

(b) RC generates a novel entry $< ID_i, x_i >$ to the database where x_i is a random number that records the validity of U_i. If U_i has revoked its account or the account is not available at present, RC generates a negative random number x_i; otherwise, x_i is a positive random number. At the same time, x_i is a preshared key. After that, RC computes $A_i = h(ID_i \parallel X \parallel T_r)$, $B_i = RPW_i \oplus h(A_i)$, $C_i = B_i \oplus h(PSK)$, $D_i = PSK \oplus A_i \oplus h(PSK)$, $V_i = h(ID_i \parallel RPW_i)$, $E_i = RPW_i \oplus x_i$, and $F_i = RPW_i \oplus h(X)$, where T_r is the registration time and X is the masker secret key between the user and the registration center.

(c) RC puts $\{B_i, C_i, D_i, V_i, E_i, F_i\}$ into SC_i. After that, RC issues it to U_i securely.

(d) With SC_i, U_i keeps P_i into SC_i and initials the authentication.

3.2.2. Login Phase. A user U_i tries to login to a server S_j by executing the steps shown in Figure 3.

(i) U_i inputs ID_i, PW_i, and BIO_i^*, then his smart card can recover R_i^* using $Rep(BIO_i^*, P_i) \longrightarrow R_i^*$.

(ii) SC_i computes $RPW_i^* = h(PW_i \parallel R_i^*)$ and then checks whether $h(ID_i \parallel RPW_i^*) = V_i$ or not. If it is true, SC_i calculates $h(PSK) = B_i \oplus C_i$, $x_i = RPW_i \oplus E_i$, and $h(X) = RPW_i \oplus F_i$. Otherwise, U_i does not pass the identity authentication.

(iii) SC_i generates a number N_1 randomly for each session and calculates $h(A_i) = RPW_i \oplus B_i$, $AID_i = ID_i \oplus h(X) \oplus h(A_i)$, $M_1 = RPW_i \oplus N_1 \oplus h(PSK) \oplus x_i$, and $M_2 = h(AID_i \parallel RPW_i \parallel SID_j \parallel T_i)$ where T_i is a timestamp.

(iv) SC_i sends $\{AID_i, M_1, M_2, B_i, D_i, T_i\}$ to S_j.

3.2.3. Authentication Phase. This phase offers the details of mutually authentication which are indicated in Figure 3.

(i) S_j receives the information from U_i and verifies whether $T_i - T_j' \leq \Delta T$ holds or not. If it holds, S_j calculates $A_i = PSK \oplus D_i \oplus h(PSK)$ and $RPW_i = B_i \oplus h(A_i)$. Otherwise, S_j will reject the login request. Then S_j checks whether $M_2 = h(AID_i \parallel RPW_i \parallel SID_j \parallel T_i)$. If it fails, the protocol would be stopped.

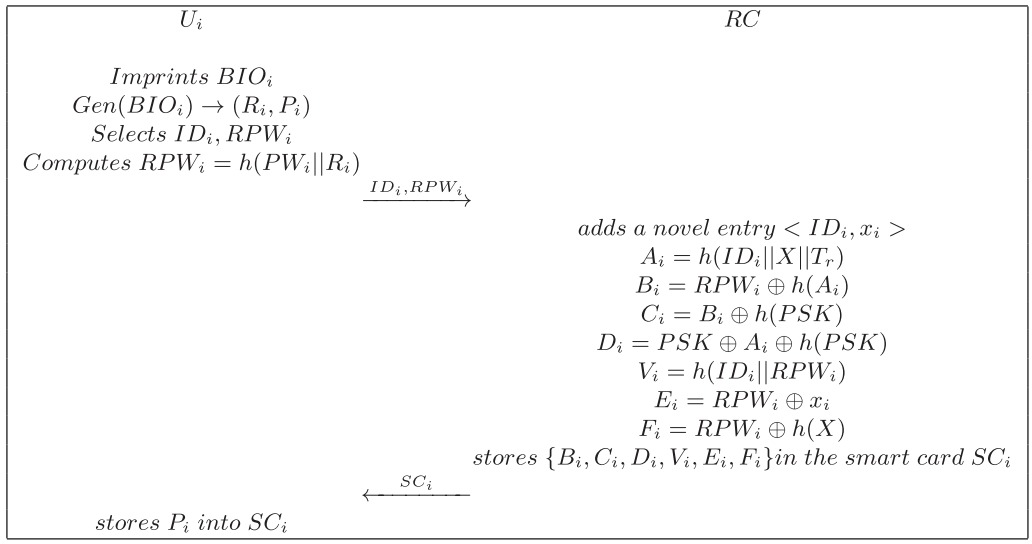


Figure 2: Registration phase of user.

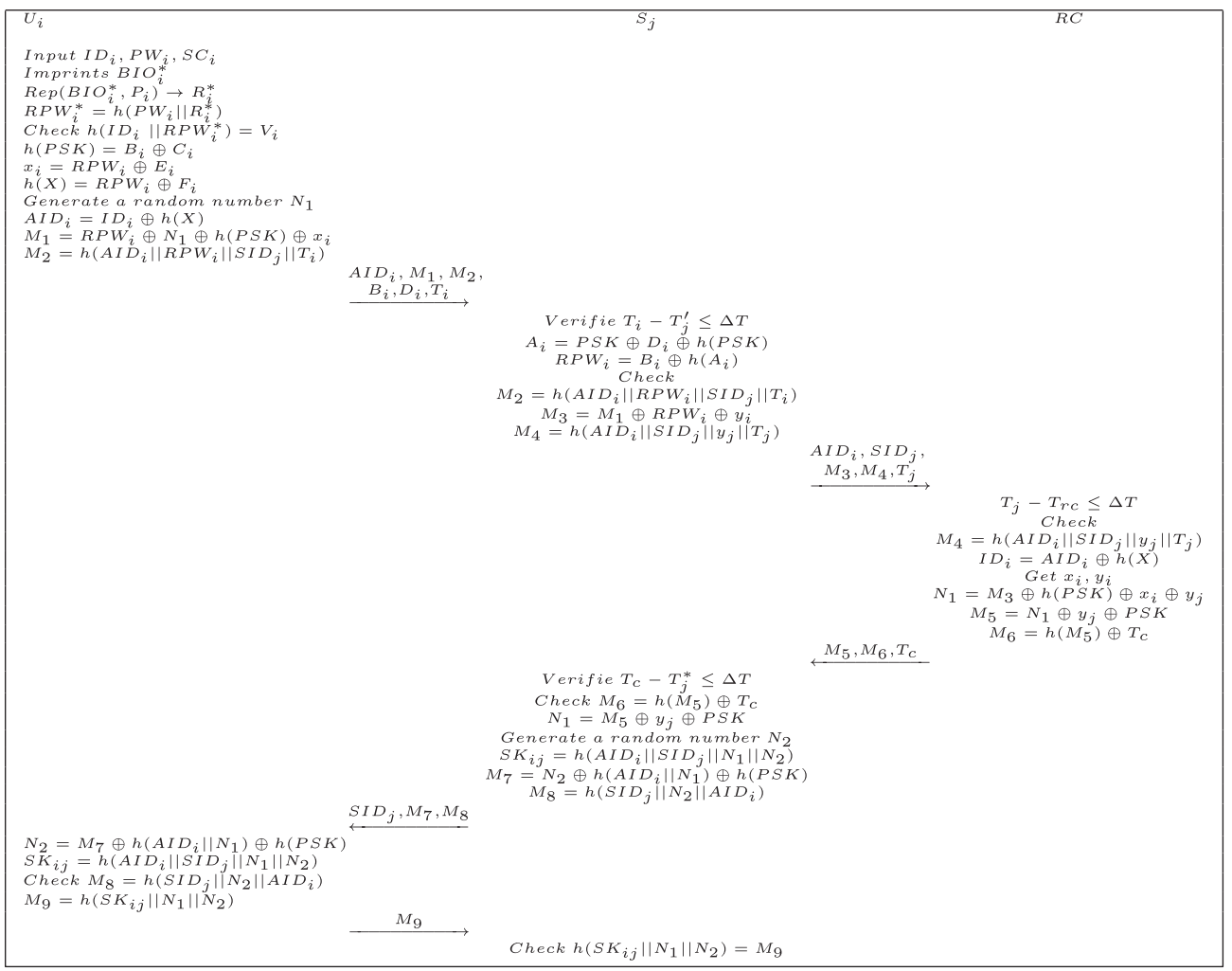


Figure 3: Login and authentication phase.

Otherwise, the server computes $AID_i' = AID_i \oplus h(A_i)$, $M_3 = M_1 \oplus RPW_i \oplus y_i$, and $M_4 = h(AID_i' \parallel SID_j \parallel y_j \parallel T_j)$ and sends messages $\{AID_i', SID_j, M_3, M_4, T_j\}$ to RC.

(ii) RC receives the messages and verifies whether $T_j - T_{rc} \leq \Delta T$ holds or not. If it holds, RC checks whether $M_4 = h(AID_i' \parallel SID_j \parallel y_j \parallel T_j)$. If it fails, the request would be stopped. Otherwise, RC computes $ID_i = AID_i' \oplus h(X)$. Then RC goes through the database $< ID_i, x_i >, < SID_j, y_j >$ stored in RC to get x_i and y_i. If x_i is a negative number, the request would be stopped. After that, RC computes $N_1 = M_3 \oplus h(PSK) \oplus x_i \oplus y_j$, $M_5 = N_1 \oplus y_j \oplus PSK$, and $M_6 = h(M_5) \oplus T_c$ where T_c is an additional timestamp. Finally, RC returns $\{M_5, M_6, T_c\}$ to S_j.

(iii) Once S_j receives the message from RC, it verifies whether $T_c - T_j^* \leq \Delta T$ holds or not. If it holds, S_j checks whether $M_6 = h(M_5) \oplus T_c$. If it fails, the request will be stopped. Otherwise, S_j calculates $N_1 = M_5 \oplus y_j \oplus PSK$. After selecting a number N_2 randomly, S_j computes $SK_{ij} = h(AID_i \parallel SID_j \parallel N_1 \parallel N_2)$, $M_7 = N_2 \oplus h(AID_i \parallel N_1) \oplus h(PSK)$, and $M_8 = h(SID_j \parallel N_2 \parallel AID_i)$. Then, the server sends $\{SID_j, M_7, M_8\}$ to U_i.

(iv) U_i retrieves N_2 and calculates SK_{ij} by computing $N_2 = M_7 \oplus h(AID_i \parallel N_1) \oplus h(PSK)$ and $SK_{ij} = h(AID_i \parallel SID_j \parallel N_1 \parallel N_2)$. After that, U_i checks whether $M_8 = h(SID_j \parallel N_2 \parallel AID_i)$ holds or not. If it is valid, the user calculates $M_9 = h(SK_{ij} \parallel N_1 \parallel N_2)$ and sends M_9 to S_j.

(v) Finally, S_j receives M_9 and checks whether the equation $h(SK_{ij} \parallel N_1 \parallel N_2) = M_9$ holds or not. If so, a secret session key is generated successfully and can be used in the following communication. Otherwise, S_j would reject the authentication.

3.2.4. Password Change Phase. Using this phase, U_i's password can be changed without any exchanging message from both RC and S_j.

(i) U_i inputs ID_i, the old password PW_i, and imprints BIO_i as well and computes $RPW_i = h(PW_i \parallel R_i)$ which is used to pass the authentication.

(ii) U_i inputs a new password PW_i^{new}. After that, U_i computes $RPW_i^{new} = h(PW_i^{new} \parallel R_i)$, $B_i^{new} = B_i \oplus RPW_i \oplus RPW_i^{new}$, $C_i^{new} = B_i^{new} \oplus h(PSK)$, $V_i^{new} = h(ID_i \parallel RPW_i^{new})$, $E_i^{new} = E_i \oplus RPW_i \oplus RPW_i^{new}$, and $F_i^{new} = F_i \oplus RPW_i \oplus RPW_i^{new}$.

(iii) SC_i replaces B_i, C_i, E_i, F_i, and V_i with $B_i^{new}, C_i^{new}, E_i^{new}$, F_i^{new}, and V_i^{new}.

3.2.5. User Revocation or Reregistration Phase. This phase is used for revocation and reregistration when U_i's smart card SC_i is stolen or lost.

(i) In revocation phase, U_i sends revocation requests to RC. RC chooses a negative random number and modifies the value of x_i corresponding to U_i as that random number.

(ii) In reregistration phase, U_i sends reregistration requests to RC. RC selects a positive random number and sets it as x_i of U_i.

4. Security Analysis of the New Protocol

4.1. Verifying the New Protocol with BAN Logic. Burrows-Abadi-Needham (BAN) logic is introduced by Burrows *et al.* [40] and widely used to analyze the security protocol. In this subsection, BAN logic is used to prove that mutual authentication can be obtained after running the new protocol successfully. The notations and postulates in BAN logic are listed in Table 3.

We first define the test goals which the new protocol should achieve using BAN logic:

(g1) $S_j| \equiv U_i| \equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$

(g2) $S_j| \equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$

(g3) $U_i| \equiv S_j| \equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$

(g4) $U_i| \equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$

Secondly, we give the idealized form of the new protocol as follows:

(m1) $U_i \longrightarrow S_j : (ID_i, RPW_i, U_i \overset{N_1}{\longleftrightarrow} S_j)_{U_i \overset{x_i}{\longleftrightarrow} RC}$

(m2) $S_j \longrightarrow RC : (AID_i', SID_j, RPW_i, U_i \overset{N_1}{\longleftrightarrow} S_j)_{S_j \overset{y_j}{\longleftrightarrow} RC}$

(m3) $RC \longrightarrow S_j : (T_c, U_i \overset{N_1}{\longleftrightarrow} S_j)_{S_j \overset{y_j}{\longleftrightarrow} RC}$

(m4) $S_j \longrightarrow U_i : (SID_j, N_1, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)_{U_i \overset{N_1}{\longleftrightarrow} S_j}$

(m5) $U_i \longrightarrow S_j : (N_1, N_2, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)_{U_i \overset{N_1, N_2}{\longleftrightarrow} S_j}$

Next, we list the following initiative premises of the new protocol:

(p1) $U_i| \equiv \#N_1$.

(p2) $S_j| \equiv \#N_2$.

(p3) $U_i| \equiv U_i \overset{x_i}{\longleftrightarrow} RC$.

(p4) $RC| \equiv U_i \overset{x_i}{\longleftrightarrow} RC$.

(p5) $S_j| \equiv S_j \overset{y_j}{\longleftrightarrow} RC$.

(p6) $RC| \equiv S_j \overset{y_j}{\longleftrightarrow} RC$.

(p7) $S_j| \equiv RC \Longrightarrow U_i \overset{N_1}{\longleftrightarrow} S_j$

(p8) $U_i| \equiv S_j \Longrightarrow U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$

Table 3: BAN logic notations and postulates.

Notations and postulates	Description
$P\mid\equiv X$	P believes the statement X is true
$P \lhd X$	P sees X
$P\mid\sim X$	P once said that X or has sent a message containing X
$P \Longrightarrow X$	P has control over X
$\#X$	X is fresh
$P \overset{K}{\longleftrightarrow} Q$	P and Q can communicate using the shared key K, only P, Q or a trusted third party know K
$(X)_k$	The formula X is combined with the formula K
$\dfrac{P\mid\equiv P \overset{K}{\longleftrightarrow} Q, P \lhd \{X\}_K}{P\mid\equiv Q\mid\sim X}$	Rule(a): The message-meaning rule
$\dfrac{P\mid\equiv Q \Longrightarrow X, P\mid\equiv Q\mid\equiv X}{P\mid\equiv X}$	Rule(b): The jurisdiction rule
$\dfrac{P\mid\equiv \#(X), P\mid\equiv Q\mid\sim X}{P\mid\equiv Q\mid\equiv X}$	Rule(c): The nonce-verification rule
$\dfrac{P\mid\equiv \#(X)}{P\mid\equiv \#(X,Y)}$	Rule(d): The freshness-conjuncatenation rule
$\dfrac{P\mid\equiv X, P\mid\equiv Y}{P\mid\equiv (X,Y)}, \dfrac{P\mid\equiv Q\mid\equiv (X,Y)}{P\mid\equiv Q\mid\equiv X}$	Rule(e): The belief rule

(p9) $S_j\mid\equiv U_i \Longrightarrow U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$.

(p10) $U_i\mid\equiv U_i \overset{N_1}{\longleftrightarrow} S_j$

(p11) $S_j\mid\equiv U_i \overset{N_2}{\longleftrightarrow} S_j$

(p12) $S_j\mid\equiv \#T_c$.

(p13) $RC\mid\equiv \#T_j$.

Finally, we analyze the new protocol using the BAN logic rules and the assumptions.

From message m_3, we obtain

(S1) $S_j \lhd (T_c, U_i \overset{N_1}{\longleftrightarrow} S_j)_{S_j \overset{y_j}{\longleftrightarrow} RC}$.

From (p5), (S1), and Rule(a), we get

(S2) $S_j\mid\equiv RC\mid\sim (T_c, U_i \overset{N_1}{\longleftrightarrow} S_j)$.

From (p12) and Rule(d), we get

(S3) $S_j\mid\equiv \#(T_c, U_i \overset{N_1}{\longleftrightarrow} S_j)$.

From (S2), (S3), and Rule(c), we get

(S4) $S_j\mid\equiv RC\mid\equiv (T_c, U_i \overset{N_1}{\longleftrightarrow} S_j)$.

From (S4) and Rule(e), we get

(S5) $S_j\mid\equiv RC\mid\equiv U_i \overset{N_1}{\longleftrightarrow} S_j$.

From (p7), (S5), and Rule(b), we get

(S6) $S_j\mid\equiv U_i \overset{N_1}{\longleftrightarrow} S_j$.

From message (m4), we have

(S7) $U_i \lhd (SID_j, N_1, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)_{U_i \overset{N_1}{\longleftrightarrow} S_j}$.

From (p10), (S7), and Rule(a), we get

(S8) $U_i\mid\equiv S_j\mid\sim (SID_j, N_1, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)$.

From (p1) and Rule(d), we get

(S9) $U_i\mid\equiv \#(SID_j, N_1, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)$.

From (S8), (S9), and Rule(c), we get

(S10) $U_i\mid\equiv S_j\mid\equiv (SID_j, N_1, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)$.

From (S10) and Rule(e), we get

(S11) $U_i\mid\equiv S_j\mid\equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$(g3).

From (P8), (S11), and Rule(b), we get

(S12) $U_i\mid\equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j$(g4).

From message (m5), we have

(S13) $S_j \lhd (N_1, N_2, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)_{U_i \overset{N_1,N_2}{\longleftrightarrow} S_j}$.

From (p11), (S6), (S13), and Rule(a), we get

(S14) $S_j\mid\equiv U_i\mid\sim (N_1, N_2, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j)$.

From (p2) and Rule(d), we get

$$(\text{S15})\ S_j| \equiv \#(N_1, N_2, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j).$$

From (S14), (S15), and Rule(c), we get

$$(\text{S16})\ S_j| \equiv U_i| \equiv (N_1, N_2, U_i \overset{SK_{ij}}{\longleftrightarrow} S_j).$$

From (S16) and Rule(e), we get

$$(\text{S17})\ S_j| \equiv U_i| \equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j(\text{g1}).$$

Finally, From (P9), (S17), and Rule(b), we get

$$(\text{S18})\ S_j| \equiv U_i \overset{SK_{ij}}{\longleftrightarrow} S_j(\text{g2}).$$

According to (g1), (g2), (g3), and (g4), we conclude that the new protocol provides the mutual authentication and a shared secret key between the user and the server after a successful running of the protocol.

4.2. Formal Security Analysis. Recent research has shown that user-chosen passwords follow the Zipf's law [41], a vastly different distribution from the uniform distribution. In this subsection, we provide a formal security analysis of the new protocol with the Zipf's law.

Theorem 1. *Let l be the length of the biometric key BIO_i, let $|Hash|$ be the range space of hash function $h(\cdot)$, and both C and s are the Zipf's parameters [41]. Let q_{send} be Send queries and q_h be Hash oracle queries. For any adversary $\mathscr{A}$ in polynomial time t against the new protocol P in the random oracle, the advantage of $\mathscr{A}$ breaking the $SK-$security of P is*

$$Adv_p^{ake} \leq \frac{q_h^2}{|Hash|} + \frac{C \cdot q_{send}^s}{2^{l-1}}. \tag{1}$$

Proof. Let E_i be the event that $\mathscr{A}$ guesses bit b for G_i in the test session successfully. According to the new protocol, $\mathscr{A}$ does not need to guess or compute the user's identity since there is only one user. The games G_0 to G_3 are listed as follows.

Game G_0. This game corresponds to the real attack in the random oracle model. Hence

$$Adv_p^{ake} = \left|2\text{pr}\left[E_0\right] - 1\right|. \tag{2}$$

Game G_1. We simulate $\mathscr{A}$'s eavesdropping attack by querying *Execute* oracles. Then, $\mathscr{A}$ sends the *Test* query and decides whether the outcome of *Test* query matches with SK which can be calculated as $SK_{ij} = h(AID_i \parallel SID_j \parallel N_1 \parallel N_2)$. $\mathscr{A}$ cannot get the message about PSK, $h(X)$. and AID_i, due to the security of S_j's PSK and U_i's $h(X)$. Thus $\mathscr{A}$ cannot increase the chance of winning game G_1. Hence we have

$$\text{pr}\left[E_0\right] = \text{pr}\left[E_1\right]. \tag{3}$$

Game G_2. We simulate $\mathscr{A}$'s active attack by querying *Send* and *Hash* oracles. $\mathscr{A}$ will manage to find the collisions of *Hash* in the way of make queries, but it is impossible for him to know the message of both $\{AID_i, M_1, M_2, B_i, D_i, T_i\}$ and $\{SID_j, M_7, M_8\}$ without the knowledge of T_i, N_1, and N_2. Hence there is no collision when querying *Send* oracles. Using the birthday paradox, we obtain

$$\left|\text{pr}\left[E_2\right] - \text{pr}\left[E_1\right]\right| = \frac{q_h^2}{2.\,|Hash|}. \tag{4}$$

Game G_3. This game simulates the smart card lost attack by querying *CorruptSC* oracle. If $\mathscr{A}$ wants to obtain the secret information of users, he tries online dictionary attack due to the low entropy of password or other computing modes to get P_i which is used as the biometrics key with the message from SC_i. Unfortunately, $\mathscr{A}$ has to know $R_i \in \{0,1\}^l$ with the probability approximated as $1/2^l$, because we use fuzzy extractor function to extract at most l nearly random bits of R_i. Even if using the Zipf's law on passwords, we still have

$$\left|\text{pr}\left[E_3\right] - \text{pr}\left[E_2\right]\right| \leq \frac{C \cdot q_{send}^s}{2^l}. \tag{5}$$

Moreover, $\mathscr{A}$ cannot get any useful messages about the value of c because of the independence and randomness of each session key. Thus, we have

$$\text{pr}\left[E_3\right] = \frac{1}{2}. \tag{6}$$

Combined the above steps, we can get the result as follows:

$$Adv_p^{ake} \leq \frac{q_h^2}{|Hash|} + \frac{C \cdot q_{send}^s}{2^{l-1}} \tag{7}$$

□

4.3. Informal Security Analysis. In this subsection, informal security analysis is conducted to show that the new protocol can withstand various attacks.

Replay Attack. If $\mathscr{A}$ replays a former piece of user's messages to server, he will not success since a timestamp is used in each session to guarantee the freshness of time. If the information in a previous session is replayed, the interval between T_j and T_i will not be in an endurable range. Therefore, in the authentication phase $\mathscr{A}$ cannot pass the authentication in the first step. Hence, the new protocol can resist the replay attack.

Modification Attack. It is assumed that an adversary $\mathscr{A}$ intercepts the information transmitted on the public channel and intends to modify the information to pass the authentication. Unfortunately, the integrity of the transmitted messages in the new scheme is protected by using one-way hash function. Moreover, $\mathscr{A}$ cannot retrieve N_1 and N_2 from the intercepted messages, thus he cannot generate a legitimate authentication message. Therefore the new protocol can resist the modification attack.

Server Session Key Attack. In our proposed scheme, on one hand, session key SK_{ij} contains N_1, N_2, AID_i, SID_j, N_1, and N_2 which are different in every session and thus cannot

be retrieved directly by a malicious adversary $\mathscr{A}$. On the other hand, our scheme provides mutual authentication in the authentication phase and makes an improvement, i.e., both of the user and the server know whether SK_{ij} has already been generated by each other. If the server S_1 wants to obtain the session key SK_{ij} by calculation, he has to obtain y_2 since $N_1 = M_5 \oplus y_2 \oplus PSK$. Unfortunately, the specific value of y_2 is known only to S_2 and RC. After receiving the messages transmitted from the user, the server calculates $M_9 = h(SK_{ij} \parallel N_1 \parallel N_2)$ which means that the authentication is passed and valid session key has already be generated by each other. Therefore our scheme holds the security of session Key.

User Impersonation Attack. If $\mathscr{A}$ is going to impersonate a valid user, $\mathscr{A}$ has to retrieve BIO_i, PW_i, and ID_i of U_i to pass the authentication in calculating $h(ID_i' \parallel RPW_i') = V_i$ in login phase. It is impossible for him to make it as a result of our perfect user anonymity and the uniqueness of biometric message. If the adversary wants to get access to S_j as a valid user with the messages AID_i, T_i, B_i, D_i, M_1, and M_2, he cannot pass the check $M_2 = h(AID_i \parallel RPW_i \parallel SID_j \parallel T_i)$ and form a session key with the server he communicates with.

Forgery Attack. The forgery attack refers to the existence of a legitimate but malicious user $\mathscr{A}$ who attempts to falsify the identity information of another legitimate user to login and authenticate. In the communication between the legal server S_j and the user U_i, the real identity ID_i of U_i is protected by AID_i, i.e., $AID_i = ID_i \oplus h(X) \oplus h(A_i)$. In addition, the identity A_i is different for each user. Therefore, the malicious user $\mathscr{A}$ cannot obtain the real identity ID_i of another legitimate user. Therefore, our scheme can prevent forgery attack.

Masquerade Attack. Under this attack, $\mathscr{A}$ can authenticate with the server S_j as a legal user and attempt to acquire the session key SK using the information transmitted at the authentication phase. In order to resist this attack, all messages transmitted in the public channel contain the destination or source information, such as $M_2 = h(AID_i \parallel RPW_i \parallel SID_j \parallel T_i)$ and $M_4 = h(AID_i' \parallel SID_j \parallel y_j \parallel T_j)$ with AID or SID. So that U_i and S_j verify whether one wants to be authenticated by the other. Therefore our protocol can resist the masquerade attack.

Smart Card Attack. If the user's smart card is stolen or lost and all the messages stored have been divulged by the adversary, there still no way for him to pass the authentication. At first, after acquiring B_i, D_i, V_i, E_i, and F_i, $\mathscr{A}$ still cannot get ID_i and RPW_i. So $\mathscr{A}$ is not capable of forging a valid user U_i. Also, $\mathscr{A}$ cannot get any useful messages such as RPW_i, A_i, and PSK using the messages stored in a smart card. Therefore, the new protocol is resistant to the stolen or lost smart card attack.

Offline Guessing Attack. $\mathscr{A}$ may get B_i, C_i, D_i, E_i, F_i, and V_i by side channel attack such as SPA and DPA. However, he cannot change the user's password without BIO_i, $h(PSK)$, X, or x_i during the offline environment. In addition, one-way hash function is adopted to protect user's password. Since it is impossible for different user to own the same biometric template, offline guessing attack can be avoided in the new protocol.

DoS Attack. DoS attack can seriously affect the efficiency of the server, causing the server to lose availability. However, all messages transmitted to the server and RC would be time stamped. With the help of the timestamp, the server and RC would verify the freshness and legitimacy of the message by checking M_2, M_4, and M_6. In addition, login operations require a fuzzy extractor to meet the biometric requirements. Therefore, our scheme can resist DoS attack.

Server Spoofing Attack. If $\mathscr{A}$ attempts to imitate a valid server, he is supposed to have the preshared key, a long-term secret key shared between RC and S_j. In the new protocol, $h(PSK)$ and y_j function as the preshared key which are transmitted through a secure channel and is unavailable to anyone other than RC and servers. Without y_j, it is impossible for the adversary to calculate N_1 in the authentication phase since $N_1 = M_5 \oplus y_j$. And also, without $h(PSK)$ and N_1, the adversary cannot get N_2 since $N_2 = M_7 \oplus h(AID_i \parallel N_1) \parallel h(PSK)$. Thus the adversary cannot imitate a valid server.

User Anonymity. The users real identity is protected by replacing ID_i with AID_i where $AID_i = ID_i \oplus h(X) \oplus h(A_i)$. Also, due to the hash function and the secret key, either an outside adversary $\mathscr{A}$ or an honest-but-curious server cannot figure out ID_i through AID_i. Thus the weak anonymity of the user is guaranteed.

Regrettably, the anonymity of the new scheme is not perfect. For example, assuming that the server cooperates with a malicious user, the malicious user provides $h(X)$ by calculating $h(X) = RPW_i \oplus F_i$, and the server calculates A_i through $A_i = PSK \oplus D_i \oplus h(PSK)$, then the server can calculate ID_i by calculating $ID_i = AID_i \oplus h(X) \oplus h(A_i)$. Moreover, an adversary $\mathscr{A}$ with U_i's lost smart card can also compute U_i's identity. Therefore, our scheme just provides the weak anonymity.

5. Efficiency Analysis

Efficiency analysis is conducted in this section to evaluate the new protocol. The comparisons including the resistance, functionality, and performance are summarized. In Table 5, let (S1) denote Chuang *et al.*'s protocol [29], (S2) denote Wang *et al.*'s protocol [33], (S3) denote Yang *et al.*'s protocol [24], (S4) denote Reddy *et al.*'s protocol [34], (S5) denote HE-WANG's protocol [42], and (S6) denote Odelu *et al.*'s protocol [43]. The following notations are defined in Table 4.

Security comparison is offered by Table 5. In Table 5, "/" denotes that the security has not been analyzed until now. From Table 5, it is easy to see that protocols of (S1), (S2), (S3), and (S4), which do not include the registration center into the authentication phase, can not resist the passive attack from an honest-but-curious server. Although (S5) is resistant to above attack, it can not resist the user impersonate attack and smart card attack. The new protocol, together with (S6), achieves all resistance requirements, since they implement the authentication with the help of the authentication center. Thus they are more secure than the first five protocols.

Table 4: Notations in security comparison table.

$R1$	resistance to replay attack
$R2$	resistance to modification attack
$R3$	resistance to Server session key attack
$R4$	resistance to user impersonate attack
$R5$	resistance to forgery attack
$R6$	resistance to masquerade attack
$R7$	resistance to smart card attack
$R8$	resistance to off-line guessing attack
$R9$	resistance to Dos attack
$R10$	resistance to server spoofing attack

Table 5: The security comparison.

	S1	S2	S3	S4	S5	S6	our scheme
$R1$	No	Yes	Yes	Yes	No	Yes	Yes
$R2$	Yes	Yes	Yes	/	Yes	Yes	Yes
$R3$	No	No	/	/	No	Yes	Yes
$R4$	No	No	Yes	Yes	No	Yes	Yes
$R5$	Yes	Yes	/	/	Yes	Yes	Yes
$R6$	No	No	/	/	Yes	Yes	Yes
$R7$	No	No	Yes	Yes	Yes	Yes	Yes
$R8$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R9$	No	Yes	Yes	Yes	Yes	Yes	Yes
$R10$	No	Yes	Yes	/	Yes	Yes	Yes

Table 6: Notations in functionality comparison table.

$F1$	anonymity
$F2$	mutual authentication
$F3$	session key agreement
$F4$	perfect forward secrecy
$F5$	user revocation/re-registration

Functionalities comparison is listed in Table 7. The notations that appear in Table 7 are lists in Table 6. It can be seen that (S1), (S2), (S3), and (S4) do not provide user revocation/reregistration functionality, and (S5) does not offer anonymity property. Only our new protocol and (S6) provide all five basic functionality requirements.

Now we conduct the efficiency analysis including computation overhead and communication overhead. To compare with other related works, only login and authentication phase are considered.

Tables 9 and 10 list the computation cost comparisons from different aspects. The notations that appear in Table 9 are listed in Table 8. For the computation efficiency, we only calculate the number of hash functions, while ignore Exclusive OR operation and concatenating operation since they require little computational cost. Let T_ω denote the computation time for symmetric-key encryption/decryption which is known as about 0.005ms, T_h denote the computation time for one-way hash function which is known as about 0.002ms, and T_m denote the computation time for elliptic curve point multiplication which is known as about 2.226ms.

Table 9 compares the computation time according to protocol's different phase. From Table 9, we can find that the new protocols, together with (S1), (S2), (S3), and (S4), spend almost the same time since only hash function contributes to computation cost. On the other hand, (S5) and (S6) take more time for computation due to the expensive elliptic curve point multiplication operations.

Table 10 compares the computation time according to different participants. The user's executing time in the new protocol only needs 0.014ms, which proves that the new protocol provides the most efficient use's computation. In terms of server's executing time, the new protocol spends almost the same time as that of the most efficient protocols, i.e., (S1), (S2), and (S3). To resist the passive attack from the honest-but-curious servers, (S5), (S6), and the new protocol introduce the registration center into the authentication phase, which would bring extra burden for the trusty registration center. As shown in Table 10, *RC* needs extra 4.47ms for (S5) and extra 2.263ms for (S6). In the new protocol, *RC* is only used to transmit the secret information instead of authenticating user and server. As a result, the extra executing time for *RC* in the new protocol is only 0.008ms, which is much less than that of (S5) and (S6). Therefore, the new protocol is the most efficient one among the second kind of multiserver authentication protocols. In conclusion, among all of the multiserver protocols against the passive attack from

Table 7: The functionality comparison.

	S1	S2	S3	S4	S5	S6	our scheme
$F1$	Yes	Yes	Yes	Yes	No	Yes	Yes
$F2$	No	Yes	Yes	Yes	Yes	Yes	Yes
$F3$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$F4$	No	Yes	Yes	Yes	Yes	Yes	Yes
$F5$	No	No	Yes	Yes	Yes	Yes	Yes

Table 8: Notations in computation comparison table.

$C1$	computation overhead in the login phase
$C2$	execution overhead in the login phase
$C3$	computation overhead in the authentication phase
$C4$	execution overhead in the authentication phase
$C5$	total execution overhead

Table 9: Computation cost comparison in different phase.

	S1	S2	S3	S4	S5	S6	our scheme
$C1$	$4T_h$	$4T_h$	$5T_h$	$6T_h+1T_m$	$3T_h+2T_m$	$5T_h+2T_m+1T_\omega$	$3T_h$
$C2$	0.008ms	0.008ms	0.01ms	2.238ms	4.458ms	4.467ms	0.006ms
$C3$	$13T_h$	$11T_h$	$13T_h$	$9T_h+3T_m$	$18T_h+6T_m$	$19T_h+4T_m+5T_\omega$	$17T_h$
$C4$	0.026ms	0.022ms	0.026ms	6.696ms	13.392ms	8.967ms	0.034ms
$C5$	0.034ms	0.03ms	0.036ms	8.934ms	17.85ms	13.434ms	0.04ms

Table 10: Computation cost comparison in different participants.

	S1	S2	S3	S4	S5	S6	our scheme
User cost	$9T_h$	$8T_h$	$9T_h$	$9T_h+2T_m$	$7T_h+3T_m$	$7T_h+3T_m+1T_\omega$	$7T_h$
User time	0.018ms	0.016ms	0.018ms	4.47ms	6.692ms	6.697ms	0.014ms
Server cost	$8T_h$	$7T_h$	$9T_h$	$6T_h+2T_m$	$5T_h+3T_m$	$6T_h+2T_m+2T_\omega$	$9T_h$
Server time	0.016ms	0.014ms	0.018ms	4.464ms	6.688ms	4.474ms	0.018ms
RC cost	/	/	/	/	$9T_h+2T_m$	$11T_h+1T_m+3T_\omega$	$4T_h$
RC time	/	/	/	/	4.47ms	2.263ms	0.008ms
Total cost	$17T_h$	$15T_h$	$18T_h$	$15T_h+4T_m$	$21T_h+8T_m$	$24T_h+6T_m+6T_\omega$	$20T_h$
Total time	0.034ms	0.03ms	0.036ms	8.934ms	17.85ms	13.434ms	0.04ms

an honest-but-curious attack, the new protocol is the most computational efficient one.

Table 12 lists the new protocol's communication cost together with the other related protocols. Suppose the random number x_i is 160 bits, the length of the user identity is 160 bits, the length of the timestamp is 16 bits, and the output length of one-way hash function is 160 bits if SHA-1 is adopted. Table 11 shows the notations that appear in Table 12. In the new protocol, when U_i logs in, he has to transmit M_1, AID_i, M_2, T_i, B_i, and D_i; thus the length of these messages is $(160*5+16)/8 = 102$ bytes. In the authentication phase, we introduce the registration center, so the communication cost is a little more than (S1), (S3), and (S4), about 180 bytes or so. Among all of the multiserver protocols against the passive attack from an honest-but-curious attack, the new protocol has the high communication efficient.

Combined with the security properties and the functionalities, we conclude that the new protocol and (S6) achieve all basic security properties and satisfy all functionalities. In terms of efficiency, (S6) spends much more computation time, bandwidth, and storage space compared with the new protocol. In conclusion, the new protocol is the most efficient multiserver authentication protocol which satisfies all basic security properties and functionalities.

6. Conclusion

In this paper, we found that a kind of multifactor multiserver authentication protocols can not resist the passive attack from an honest-but-curious servers. We took Wang *et al.*'s protocol as an example, to exhibit how an honest-but-curious server step by step obtained a session key which should

Table 11: Notations in communication comparison table.

$COM1$	communication cost in login phase
$COM2$	communication cost in authentication and key agreement phase
$COM3$	total communication cost

Table 12: Communication cost comparison table.

	S1	S2	S3	S4	S5	S6	our scheme
$COM1$	80 Bytes	102 Bytes	102 Bytes	80 Bytes	64 Bytes	108 Bytes	102 Bytes
$COM2$	80 Bytes	80 Bytes	60 Bytes	80 Bytes	376 Bytes	260 Bytes	180 Bytes
$COM3$	160 Bytes	182 Bytes	162 Bytes	160 Bytes	440 Bytes	368 Bytes	282 Bytes

be kept secret from him. Moreover, we observed that the revocation and reregistration process in their protocol is incorrect. To remedy these weaknesses, this paper proposed a novel multiserver authentication protocol. The new protocol satisfies comprehensive demands of security and provides versatile and practical functionalities. Compared with the related protocols in computation cost and communication cost, the new protocol is the most efficient multiserver authentication protocol which satisfies all basic security properties and functionalities. Therefore, the new protocol is secure and relatively efficient in the remote distributed authentication networks. We have noticed that this kind of attack may also exist in other likewise environment, such as the multifactor multigateway authentication protocol in the wireless sensor networks. As a future work, we would apply the passive attack from an honest-but-curious gateway to the multifactor multigateway authentication protocol in the wireless sensor network and try to design secure protocols for multigateway wireless sensor network.

Data Availability

The paper does not use any data set.

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 (nos. 61572027, 61772194, and U1636208), special foundation for coconstruction project of Beijing, the Hunan Provincial Natural Science Foundation of China under Grant no. 2018JJ3191, and the Guangxi Key Laboratory of Trusted Software (no. KX201707).

References

[1] X. Yang, X. Huang, and J. K. Liu, "Efficient handover authentication with user anonymity and untraceability for Mobile Cloud Computing," *Future Generation Computer Systems*, vol. 62, pp. 190–195, 2016.

[2] S. A. Chaudhry, "A secure biometric based multi-server authentication scheme for social multimedia networks," *Multimedia Tools and Applications*, vol. 75, no. 20, pp. 12705–12725, 2016.

[3] J. Shen, S. Chang, J. Shen, Q. Liu, and X. Sun, "A lightweight multi-layer authentication protocol for wireless body area networks," *Future Generation Computer Systems*, 2016.

[4] D. He, N. Kumar, M. K. Khan, and J.-H. Lee, "Anonymous two-factor authentication for consumer roaming service in global mobility networks," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 4, pp. 811–817, 2013.

[5] C. Jin, C. Xu, X. Zhang, and J. Zhao, "A secure RFID mutual authentication protocol for healthcare environments using elliptic curve cryptography," *Journal of Medical Systems*, vol. 39, no. 3, pp. 1–8, 2015.

[6] L. Lamport, "Password authentication with insecure communication," *Communications of the ACM*, vol. 24, no. 11, pp. 770–772, 1981.

[7] L. Li, I. Lin, and M. Hwang, "A remote password authentication scheme for multiserver architecture using neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 12, no. 6, pp. 1498–1504, 2001.

[8] I. C. Lin, M. S. Hwang, and L. H. Li, "A new remote user authentication scheme for multi-server architecture," *Future Generation Computer Systems*, vol. 19, no. 1, pp. 13–22, 2003.

[9] X. Cao and S. Zhong, "Breaking a remote user authentication scheme for multi-server architecture," *IEEE Communications Letters*, vol. 10, no. 8, pp. 580-581, 2006.

[10] W. S. Juang, "Efficient multi-server password authenticated key agreement using smart cards," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 251–255, 2004.

[11] C.-C. Chang and J.-S. Lee, "An efficient and secure multi-server password authentication scheme using smart cards," in *Proceedings of the Proceedings - 2004 International Conference on Cyberworlds, CW 2004*, pp. 417–422, Japan, November 2004.

[12] J.-L. Tsai, "Efficient multi-server authentication scheme based on one-way hash function without verification table," *Computers & Security*, vol. 27, no. 3-4, pp. 115–121, 2008.

[13] Y. P. Liao and S. S. Wang, "A secure dynamic ID based remote user authentication scheme for multi-server environment," *Computer Standards & Interfaces*, vol. 31, no. 1, pp. 24–29, 2009.

[14] T. Chen Y, M. Hwang S, C. Lee et al., "Cryptanalysis of a Secure Dynamic ID Based Remote User Authentication Scheme for Multi-Server Environment," *Innovative Computing, Information and Control (ICICIC)*, pp. 725–728, 2009.

[15] H. Hsiang C and K. Shih W, "Improvement of the secure dynamic ID based remote user authentication scheme for multi-server environment," *Computer Standards & Interfaces*, vol. 31, no. 6, pp. 1118–1123, 2009.

[16] C. Lee, T. Lin, and R. Chang, "A secure dynamic ID based remote user authentication scheme for multi-server environment using smart cards," *Expert Systems with Applications*, vol. 38, no. 11, pp. 13863–13870, 2011.

[17] D. Wang, D. He, P. Wang, and C.-H. Chu, "Anonymous Two-Factor Authentication in Distributed Systems: Certain Goals Are Beyond Attainment," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 428–442, 2015.

[18] X. Li, M. H. Ibrahim, S. Kumari, A. K. Sangaiah, V. Gupta, and 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.

[19] X. Li, J. Niu, M. Z. A. Bhuiyan, F. Wu, M. Karuppiah, and S. Kumari, "A Robust ECC based Provable Secure Authentication Protocol with Privacy Protection for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3599–3609, 2018.

[20] X. Li, J. Peng, J. Niu, F. Wu, J. Liao, and K.-K. R. Choo, "A Robust and Energy Efficient Authentication Protocol for Industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1606–1615, 2018.

[21] R. Amin, S. K. H. Islam, N. Kumar, and K.-K. R. Choo, "An untraceable and anonymous password authentication protocol for heterogeneous wireless sensor networks," *Journal of Network and Computer Applications*, vol. 104, pp. 133–144, 2018.

[22] D. He and S. Wu, "Security flaws in a smart card based authentication scheme for multi-server environment," *Wireless Personal Communications*, vol. 70, no. 1, pp. 323–329, 2013.

[23] R. S. Pippal, C. D. Jaidhar, and S. Tapaswi, "Robust smart card authentication scheme for multi-server architecture," *Wireless Personal Communications*, vol. 72, no. 1, pp. 729–745, 2013.

[24] L. Yang and Z. Zheng, "Cryptanalysis and improvement of a biometrics-based authentication and key agreement scheme for multi-server environments," *PLoS ONE*, vol. 13, no. 3, 2018.

[25] C.-T. Li and M.-S. Hwang, "An efficient biometrics-based remote user authentication scheme using smart cards," *Journal of Network and Computer Applications*, vol. 33, no. 1, pp. 1–5, 2010.

[26] X. Li, J. W. Niu, J. Ma, W. D. Wang, and C. L. Liu, "Cryptanalysis and improvement of a biometrics-based remote user authentication scheme using smart cards," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 73–79, 2011.

[27] E.-J. Yoon and K.-Y. Yoo, "Robust biometrics-based multi-server authentication with key agreement scheme for smart cards on elliptic curve cryptosystem," *The Journal of Supercomputing*, vol. 63, no. 1, pp. 235–255, 2013.

[28] H. Kim, W. Jeon, K. Lee, Y. Lee, and D. Won, "Cryptanalysis and improvement of a biometrics-based multi-server authentication with key agreement scheme," in *Proceedings of International Conference on Computational Science and Its Applications*, pp. 391–406, 2012.

[29] M.-C. Chuang and M. C. Chen, "An anonymous multi-server authenticated key agreement scheme based on trust computing using smart cards and biometrics," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1411–1418, 2014.

[30] D. Mishra, A. Das, and S. Mukhopadhyay, "A secure user anonymity-preserving biometric-based multi-server authenticated key agreement scheme using smart cards," *Expert Systems with Applications*, vol. 41, no. 18, pp. 8129–8143, 2014.

[31] Y. Lu, L. Li, X. Yang, and Y. Yang, "Robust biometrics based authentication and key agreement scheme for multi-server environments using smart cards," *PLoS ONE*, vol. 10, no. 5, Article ID 0126323, 2015.

[32] A. G. Reddy, A. K. Das, V. Odelu, and K.-Y. Yoo, "An enhanced biometric based authentication with key-agreement protocol for multi-server architecture based on elliptic curve cryptography," *PLoS ONE*, vol. 11, no. 5, 2016.

[33] C. Wang, X. Zhang, and Z. Zheng, "Cryptanalysis and improvement of a biometric-based multi-server authentication and key agreement scheme," *PLoS ONE*, vol. 11, no. 2, 2016.

[34] A. G. Reddy, E.-J. Yoon, A. K. Das, V. Odelu, and K.-Y. Yoo, "Design of Mutually Authenticated Key Agreement Protocol Resistant to Impersonation Attacks for Multi-Server Environment," *IEEE Access*, vol. 5, pp. 3622–3639, 2017.

[35] P. Jiang, Q. Wen, W. Li, Z. Jin, and H. Zhang, "An anonymous and efficient remote biometrics user authentication scheme in a multi server environment," *Frontiers of Computer Science*, vol. 9, no. 1, pp. 142–156, 2015.

[36] A. K. Das, P. Sharma, S. Chatterjee, and J. K. Sing, "A dynamic password-based user authentication scheme for hierarchical wireless sensor networks," *Journal of Network and Computer Applications*, vol. 35, no. 5, pp. 1646–1656, 2012.

[37] D. Wang, Q. Gu, H. Cheng, and P. Wang, "The request for better measurement: A comparative evaluation of two-factor authentication schemes," in *Proceedings of 11th ACM Asia Conference on Computer and Communications Security, ASIA CCS*, pp. 475–486, China, June 2016.

[38] S. Qiu, G. Xu, H. Ahmad, and L. Wang, "A Robust Mutual Authentication Scheme Based on Elliptic Curve Cryptography for Telecare Medical Information Systems," *IEEE Access*, vol. 6, pp. 7452–7463, 2017.

[39] S. Qiu, G. Xu, H. Ahmad, and Y. Guo, "An enhanced password authentication scheme for session initiation protocol with perfect forward secrecy," *PLoS ONE*, vol. 13, no. 3, 2018.

[40] M. Burrows, M. Abadi, and R. Needham, "Logic of authentication," *ACM Transactions on Computer Systems*, vol. 8, no. 1, pp. 18–36, 1990.

[41] D. Wang and P. Wang, "On the anonymity of two-factor authentication schemes for wireless sensor networks," *Computer Networks*, pp. 73-41, 2014.

[42] D. He and D. Wang, "Robust Biometrics-Based Authentication Scheme for Multiserver Environment," *IEEE Systems Journal*, vol. 9, no. 3, pp. 816–823, 2015.

[43] V. Odelu, A. K. Das, and A. Goswami, "A Secure Biometrics-Based Multi-Server Authentication Protocol Using Smart Cards," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1953–1966, 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 9149730, 14 pages
https://doi.org/10.1155/2018/9149730

Research Article

Efficient Multifactor Two-Server Authenticated Scheme under Mobile Cloud Computing

Ziyi Han[1], **Li Yang**[1], **Shen Wang**[2], **Sen Mu**[2] **and Qiang Liu**[3]

[1]*School of Computer Science and Technology, Xidian University, Xi'an 710071, China*
[2]*Aisino Corporation, Beijing 100195, China*
[3]*Science and Technology on Communication Networks Laboratory, Shijiazhuang 050081, China*

Correspondence should be addressed to Ziyi Han; nnthzy@163.com

Received 28 April 2018; Accepted 9 August 2018; Published 12 September 2018

Academic Editor: Jian Shen

Because the authentication method based on username-password has the disadvantage of easy disclosure and low reliability and the excess password management degrades the user experience tremendously, the user is eager to get rid of the bond of the password in order to seek a new way of authentication. Therefore, the multifactor biometrics-based user authentication wins the favor of people with advantages of simplicity, convenience, and high reliability. Now the biometrics-based (especially the fingerprint information) authentication technology has been extremely mature, and it is universally applied in the scenario of the mobile payment. Unfortunately, in the existing scheme, biometric information is stored on the server side. As thus, once the server is hacked by attackers to cause the leakage of the fingerprint information, it will take a deadly threat to the user privacy. Aiming at the security problem due to the fingerprint information in the mobile payment environment, we propose a novel multifactor two-server authenticated scheme under mobile cloud computing (MTSAS). In the MTSAS, it divides the authentication method and authentication means; in the meanwhile, the user's biometric characteristics cannot leave the user device. Thus, MTSAS avoids the fingerprint information disclosure, protects user privacy, and improves the security of the user data. In the same time, considering user actual requirements, different authentication factors depending on the privacy level of authentication are chosen. Security analysis proves that MTSAS has achieved the authentication purpose and met security requirements by the BAN logic. In comparison with other schemes, the result shows that MTSAS not only has the reasonable computational efficiency, but also keeps the superior communication cost.

1. Introduction

With the vigorous development of the mobile Internet, the cloud computing service based on the mobile terminal has emerged. The mobile cloud computing is the deep fusion of the mobile Internet and the cloud computing, which represents the future trend in the development of the cloud computing [1]. Compared with the traditional cloud services, the mobile cloud service has the characteristics of mobile interconnection, flexible terminal application, and convenient data access [2]. However, the abundant mobile cloud service applications also bring more problems of security and privacy [3]. In the meanwhile, the mobile computing environment also puts forward the new requirements of security, convenience, and privacy protection.

Standing on the user's point of view, the traditional approach of the user authentication is based on the username and the password. In the past, the username and the password are selected to log in their account for the simplicity and the facility. But in recent years, owing to the massive popularity of the mobile terminal, people would prefer to put more and more works in the mobile terminal. In consequence, there exist more and more accounts to be managed. Research result shows that, on average, each person has 25 accounts and 6.5 passwords and logs in eight times one day. The complex password is difficult to remember again, so it is inevitable for people to use a simple and weak password and even share the same password in different network services. In case that the user's password is deceived by the phishing site or intercepted by the virus and Trojan horse, users' personal

information will be compromised. Thereby, it will threat the user account and reveal the user privacy. And also, the password itself has a lot of insurmountable defects. In the static password way, the password leakage (offline dictionary attack, etc.) and the overlapping library will greatly threaten the security of the user data. In the dynamic password way, it is low reliable for short message to authenticate users. In order to reduce the dependence on the password, there appears a mass of alternative strong authentication method [4], such as the authentication method based on the USB key [5] or the security chips [6] and the biometric-based authentication method. These strong authentication methods win the favor of the user for its convenience and strong security. In particular, the unique feature of the biometric-based authentication reinvests it with higher security [7].

In fact, in some critical systems which need the high security level of the authentication [8], such as the mobile financial environment which is related to the user's property safety, a single factor authentication method is insufficient to guarantee the security and the reliability of the authentication. In the circumstances, these factors can be combined together to establish a multifactor authentication method to ensure the strong security. The authentication factors can include the password, the token, and the biological characteristics. The mulitifactor authentication has been widely used in practice. For example, Alibaba, JDcom, and PayPal have fully supported the payment based fingerprint. The FIDO alliance, established in February 2013, also focuses on the multifactor market and devotes to establish a unified multifactor mobile authentication standard.

From the server's side, generally, peoples employ a single server to store user data and authenticate users previously. The user's passwords or the verification data of the passwords are stored in a server [9]. In this way, once the server is captured by an attacker, all passwords and the verification data stored in the server will be stolen by the attacker. Hence, this would be a serious threat to the security of the user data and leads to the leakage of users' privacy [10]. In order to overcome the inherent defect of the single server, the multiserver authentication scheme is proposed [11]. As a result, the risks will be distributed to multiple servers. In this way, the attacker must capture multiple servers in the same time to acquire user data. Consequently, the multifactor can greatly improve the security of the user data. But on the other hand, in the existing multiserver authentication scheme, when the cloud server verifies the user identity, they need to collect user's personal privacy information in the register phase, such as a password and biometric information. Then this personal privacy information are transferred via the link transmission and stored in the server side to verify the user identity. So user's personal privacy information is able to be stolen both in the process of transmission and storage. And some cloud service providers may leak even sell user's privacy information on account of business interests. As a result, the user increasingly mistrusts the cloud. Although the nonrepeatability and the uniqueness of the biometric information such as the fingerprint bring unique safety for the user, it also means that the leakage of the biometric information will have the disastrous consequence for the user to threaten the user privacy seriously. As is known to all, fortune and misfortune are neighbors. The biometric information can uniquely mark the user identity; unfortunately at the same time, it also brings the higher level security risks [12]. Once the server is hacked by attackers to cause the leakage of the biometric information, it will take a deadly threat to the user privacy. It is in urgent need of solving this problem.

Therefore, the user prefers to store the biometric information in the local device rather than in the cloud server to ensure the security of the biometric information. And also, in the current mobile device, it is widely equipped with the security mechanism, such as the trusted execution environment TEE, security chips (TPM chip, SE chips, etc.). Thus the user authentication information such as the biometric information can be stored in the trusted zone to ensure the security of the user privacy. Moreover, fingerprint information has become the most widely used biological feature [13]. And fingerprint authentication function has become the standard practice of mobile phone with one thousand yuan. So the fingerprint is chosen as the authentication factor to strengthen the authentication level in the existing biometric-based authentication schemes [7, 12, 14–16]. And the fingerprint is stored in the server. Absolutely the leakage of the fingerprint information in the server will bring serious consequences.

In allusion to the serious security problem of the fingerprint leakage in the mobile payment environment, we propose a novel multifactor two-server authenticated scheme under mobile cloud computing, shorted as MTSAS [17]. In the MTSAS, the server authenticates the device, and the device verifies the user. In the meanwhile, the user's biometric characteristics are stored locally in the user device and cannot be stolen by attackers [18]. The server side never stores the user's fingerprint information. Specially, the authentication server is deployed by the private cloud [19]. This way can download the security threat that the authentication server may face. Thus, MTSAS avoids the fingerprint information disclosure, protects user privacy, and improves the security of the user data. Moreover, the user requirements are given the full consideration. The different authentication factors depending on the different security levels are selected to make the reasonable use of the server's resources. For the more, MTSAS without introducing a third party is a lightweight protocol. Therefore, MTSAS is more suitable for the mobile payment environment.

Last but not least, the authentication server is applied in the private cloud environment. As a result, the public key of the authentication will not be broadcasted in the whole Internet but just be broadcasted in the private cloud. Consequently, the material's costs will be reduced and MTSAS is easier to be industrialized.

2. Related Work

In order to improve the security of the key exchange protocol, Pointcheval and Zimmer [20] proposed a multifactor authentication protocol with three factors, the password, the security device, and the biological characteristic. They also established

a security model and proved that it is secure in the random oracle security model. Tiwari [8] put forward a multifactor authentication system based on the Transaction Identification Code (*TIC*) and the Short Message Service (*SMS*) for the wireless payment scenario. The protocol is divided into layers so as to provide a highly secure environment, which is easy to use and deploy, and does not need to change any infrastructure or the protocol of the wireless network. Layeghian [21] focused on issues of the customer privacy in the wireless payment protocol. The identity of the customer is hidden through a blind pseudorandom signature certificate and an anonymous bank account. This scheme achieves the anonymity of the customer identity.

With the emergence of the cloud environment, Khan [22] targeted on the multifactor authentication problem in the cloud environment. For the sake of the user's security and privacy, they implemented a verification system. This system is combined with the built-in human factors (handwritten signature biometrics) and the standard knowledge factors (user's specific password) to achieve a high level of security.

A remote login two-factor scheme [23] based on the smart card was proposed in the multiserver architecture. This scheme takes advantages of the smart card as the second factor with the password to verify the user identity together. Li [24] found Chang's scheme [23] cannot resist the smart card lost attack, the check value reveal attack and the session key reveal attack. Specific to this a few security threats, an improved two-factor authentication scheme [24] is further proposed in the multiserver network.

Shen [15] pointed at the multiserver environment in the critical system; they put forward a multifactor authentication scheme with the authentication factors of the password, biological characteristics, and the random numbers. This scheme combines the multiple factors and multiple servers. Later, Li [16] discovered that Shen's scheme in the [15] was vulnerable to the denial of service attacks. The biological templates are directly stored in the smart card with no anonymity. Therefore, the loss of the biological templates will cause the direct threats to the security of the user data. Regarding this defects, an obfuscator is presented to enhance the security of the biological templates for the multiserver environment in the critical system. It is difficult to guarantee the security of the biological templates. Anyway, the biological templates are still stored in the server.

3. BAN

3.1. Basic Principle. The BAN [25] is a type of formal logic analysis method based on knowledge and belief. BAN starts with the initial basic beliefs of the protocol executive, according to every participant's issuing and receiving messages. Then it concludes the participants' eventual beliefs through the formal axioms and the logic reasoning.

When BAN is selected to verify a specific protocol, first of all, we need to idealize protocol messages and transfer them into the formulae in the BAN. Then we carry on the reasonable assumptions on the basis of the specific situation and infer the idealized messages based on the inference rules. Finally, we deduce whether the protocol can achieve the expected goal. If, at the end of the protocol, we are able to build the trust such as sharing communication key with the other identity, it can prove that the protocol is secure; otherwise, the protocol may suffer from security vulnerabilities.

Table 1: Basic concepts of BAN.

Symbols	Meaning
A	The specific main body
k_a	The specific public key of a
$k_{a^{-1}}$	The corresponding private key of a
P, Q, R	Any main body
X, Y	Any sentence
k	Any key
P *believes* X	P believes X
P *sees* X	P has received X
P *said* X	P has sent X
P *controls* X	P has jurisdiction of X
(X, Y)	X links to Y
fresh(X)	X is fresh
$\{X\}_k$	The result of X encrypted by K
$P \overset{k}{\leftrightarrow} Q$	P and Q communicate with each other by shared key K
$H(X)$	P and Q X is one-way hash function
$\overset{k}{\rightarrow} P$	K is the pubic key of P
$P \overset{X}{\rightleftharpoons} Q$	X is the secret value between P and Q

3.2. Basic Concept. First of all, Table 1 shows several basic concepts of BAN.

3.3. Logic Axioms. BAN has 17 axioms in total. We list some important axioms as follows:

(1) Message meaning rule logic axiom:

$$\frac{P \text{ believes } Q \overset{k}{\leftrightarrow} P,\ P \text{ sees } \{X\}_k}{P \text{ believes } Q \text{ said } X} \tag{1}$$

This logic axiom means that if P believes that K is the shared key between P and Q and P has received the result of X encrypted by K, P believes that Q has sent X.

(2) To public key, there is similar axiom:

$$\frac{P \text{ believes } \overset{k}{\rightarrow} Q,\ P \text{ sees} \{X\}_k{}^{-1}}{P \text{ believes } Q \text{ said } X} \tag{2}$$

(3) Temporary value verification rule logic axiom:

$$\frac{P \text{ believes } fresh(X),\ P \text{ believes } Q \text{ said } X}{P \text{ believes } Q \text{ believes } X} \tag{3}$$

This logic axiom means that if P believes that X is fresh and P believes that Q has sent X, P believes that Q believes X.

(4) Jurisdiction rule logic axiom:

$$\frac{P \text{ believes } Q \text{ controls } X,\ P \text{ believes } Q \text{ believes } X}{P \text{ believes } X} \quad (4)$$

This logic axiom means that if P believes that Q has the jurisdiction of X and P believes that Q believes X, P believes X.

(5) Logic axiom:

$$\frac{P \text{ believes } fresh(\mathrm{X})}{P \text{ believes } fresh(\mathrm{X}, \mathrm{Y})} \quad (5)$$

This logic axiom means that if the part of the formula is fresh, the whole formula is fresh.

(6) Logic axiom:

$$\frac{P \text{ believes } \overset{k}{\rightarrow} P,\ P \text{ sees } \{\mathrm{X}\}_{\mathrm{k}}}{P \text{ sees } \mathrm{X}} \quad (6)$$

This logic axiom means that if P receives messages which are encrypted by the public key itself, P is able to decrypt the received messages.

The security of MTSAS is proved from some basic beliefs to final beliefs using above axioms. If final beliefs can meet the security requirement, MTSAS is secure in theory. Unless, there may exit some security loopholes in MTSAS.

4. MTSAS

4.1. Scenario. Mobile payment scenario is a high security level authentication environment involving user's property security. Therefore, the authentication protocol in this scenario should provide high efficiency and strong authentication. The user has registered for an account in the bank by his/her mobile terminal before. When the user wants to login the system through the mobile terminal, there exists the difference of the user's behavior by means of analyzing the user's behavior. The user's behavior can be divided into two groups. One group is that the user logs in the system only to check the account information, not to process the transaction operation. Another group is that the user logs in the system to perform the financial transaction exactly. Therefore, in order to improve the user experience and the user's efficiency, it is essential to provide the authentication with different security levels aiming at this two kinds of user behaviors in the mobile payment scenario.

Because the former user's behavior does not involve the transaction information, this situation has low security requirements of the user authentication. So only providing the weak authentication, namely, the basic authentication is enough. Therefore, the traditional username-password authentication method is suitable to verify the user's identity in this situation.

For the latter type of the user's behavior, the user conducts the financial transaction operation indeed, so the strong authentication must be provided to ensure the security of the user's property and privacy information. Therefore, on the basis of the basic identification, by adding the authentication factors of the Dynamic Verification Code (*DVC*) and the fingerprint (*fp*), the security of the mobile payment process can be strongly enhanced in the way of the multifactor authentication way. At the same time so as to avoid the too concentrated risk problem of the single server and the privacy leakage problem once the single server is compromised, and there is high security requirement in the mobile payment scenario. As a result, it is essential to adopt the two-server way to share risks and ensure the robustness and the stability of the service.

Finally, the user's fingerprint information is stored locally (such as the trusted execution environment based on TPM, SE, and so on), thereby to prevent the potential security hazard on account of the fingerprint information leakage in the server side. The scenario is shown in Figure 1.

4.2. Model Description. In our system, there are two main participating entities, the mobile terminal and the cloud. The user, the fingerprint module and the user agent are located in the mobile terminal. And there are the web server and the authentication server in the cloud. The fingerprint module is embedded in the mobile terminal's chips (TPM chip, SE chip, and so on) and is protected by the trusted execution environment (the ARM TrustZone and so on). As a result, the security issues in the fingerprint module can be guaranteed.

The security model we assumed is as follows. On the one hand, the web server and the user are exposed in the open network environment. They not only should ensure the security of the interaction, but also achieve the purpose of authentication. On the other hand, the authentication server is deployed in the private cloud environment. We consider that the web server and the authentication server have been authenticated before; the authentication problem between them is not considered. And especially, they will not collude. Consequently, we only focus on the security of data transmission between the web server and the authentication server in the MTSAS.

4.3. The Proposed Scheme. The proposed authentication scheme is divided into two phases, the registration phase and the authentication phase. Before we describe the MTSAS in detail, the meanings of the symbols are presented in Table 2. There are some details needed to be described. In the basic authentication phase, *AuthLevel* is low; and *AuthLevel* is high in the transaction authentication phase.

We assume that the mobile terminal has a fingerprint module, and the fingerprint can be stored securely. In other words, there is the secure area in the mobile terminal, such as the trusted execution environment TEE based on the security chip TPM, SE, TrustZone, etc. The MTSAS includes five participants: the user, fingerprint module, the user agent, the web server, and the authentication server. The users, the fingerprint module, and the user agent are located in the mobile terminal. The fingerprint module is responsible for the fingerprint's collection, comparison, and secure storage. The user agent can be the APP or the browser. The web server is in charge of the communication with the user agent, and

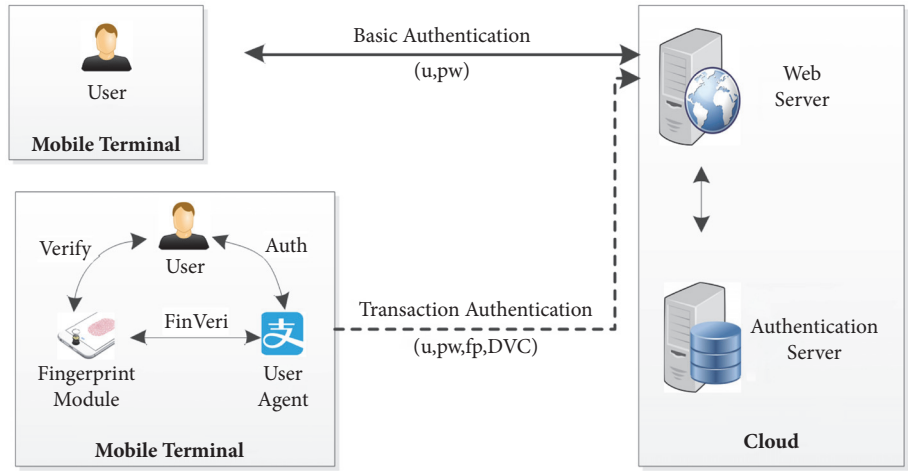


Figure 1: Scenario.

Table 2: Symbols meanings.

Symbols	Meaning
U	User
WS	Web server
AS	Authentication server
u	Username
pw	Password
phn	Phone number
fp	Fingerprint
DVC	Dynamic verification code
n	Temporary number
hpw	The hash of the password
(PK_x, SK_x)	The public/ private key pair of X
K	Session key
$E_K(M)$	Encrypt the message M by the key K
$Sig(M)$	Signature of M
$h()$	Hash function
$AuthLevel$	Authentication Level, can be High/ Low
$BaResult$	Basic authentication result, can be Success/ Failure
$TranResult$	Transaction authentication result, can be Success/ Failure
$LocResult$	Local authentication result, can be Success/ Failure

the forwarding of the authentication data. The authentication server is deployed in the private cloud environment. The authentication is performed by the authentication server. During this process, the user information is stored in the database of the authentication server.

4.3.1. Registration. In the registration phase, the user delivers $\{u, pw, phn\}$ to the web server. The web server forwards the message to the authentication server. The authentication server preserves these data and makes the username u as the index. On the other hand, the user registers locally. That is to say, the fingerprint module registers the user. The user enters the fingerprint to the fingerprint module. If there exists the fingerprint in the fingerprint module, the fingerprint only needs to associate with the user account and stores the hash value of the fingerprint in the trusted execution environment TEE.

For ensuring the security of the communication environment, we assume that the two pairs, the mobile terminal and the web server and the web server and the authentication server both have their own authentication key pair and have obtained the other's public key. The key of the mobile terminal is $\{PK_u, SK_u\}$. The key pair of the web server is $\{PK_{WS}, SK_{WS}\}$. The authentication server also has the key pair $\{PK_{AS}, SK_{AS}\}$. K is the session key between the mobile terminal and the web server.

4.3.2. Authentication. In the authentication phase, we provide the authentication scheme based on the privacy level, which is separated into the basic authentication and the transaction authentication.

When the user just logs in their account to check the account information, we choose the basic authentication scheme of the username-password to provide the authentication of the weak security level. For another, when the user needs to be trade online, we select the transaction authentication scheme by adding the authentication factors-the fingerprint and DVC to support the strong authentication.

(A) Data Transmission. To avoid giving unnecessary details, the data transmission of the user in the mobile terminal, the web server, and the authentication server is abided by the following method.

(1) User ⟷ Web Server. First, the sender encrypts the delivering data by the session key K and signs them with its own private key. Then the sender encrypts the session key K by receiver's public key.

After the receiver accepts, the receiver verifies the signature with the sender's public key. If the verification fails, the authentication fails. Otherwise, the sender passes the verification. Next, the receiver decrypts the session key K by its own private key and then decrypts the delivering data by the session key. As in the following, the receiver obtains the delivering data.

(2) Web Server ⟷ Authentication Server. The sender encrypts the delivering data by the receiver's public key and signs them with its own private key.

After the receiver accepts, the receiver verifies the signature with the sender's public key. If the signature does not pass the verification, the authentication fails. Otherwise, the sender passes the verification. Next, the receiver decrypts the delivering data by its own private key to obtain the delivering data.

(B) Basic Authentication. In this phase, the user log in his account with the method of the basic authentication using the username and his/her password only. The specific process is represented in Figure 2.

A user U enters the username u and the password pw in the user agent of the mobile terminal. The mobile terminal sets the authentication level "*AuthLevel*" to low and generates the temporary number n_u and n_w. And then the mobile terminal sends u, pw, the identifier of the web server *WS*, *AuthLevel*, and n_u, n_w to *WS*. *WS* forwards these data to *AS*. After *AS* receives, *AS* retrieves stored *hpw*, and computes the hash *h(pw)* of the received pw. Then *AS* checks whether $h(pw)$ and hpw are equal. If so, *AS* sets the basic authentication result *BaResult* to "success". Otherwise, *AS* sets *BaResult* to "Failure". Afterwards, *AS* sends BaResult, n_u, and n_w to *WS*. *WS* forwards the data to *U* to feedback. After the mobile terminal of U receives, the mobile terminal checks the temporary number firstly. The mobile terminal compares the received temporary number with the stored temporary number. If they are not equal, it proves that the data has been expired. Otherwise, the mobile terminal reads *BaResult*. If *BaResult* is "success", the user agent displays "authentication success" to the user. Otherwise, the user agent displays "authentication fail".

(C) Transaction Authentication. When the user has passed the basic authentication and requires the online trading. The strong authentication is needed to verify the user identity. In consequence, the fingerprint and *DVC* are added as the authentication factors to ensure the strong security. When user requires the online trading, the fingerprint module triggers the local verification process and then triggers the *DVC* verification process. Figure 3 shows the specific process.

The mobile terminal sets *AuthLevel* to high and generates the temporary number n_{u1}. Then the mobile terminal triggers the local verification process; that is to say, the user agent delivers the fingerprint verification request to the fingerprint module. After the fingerprint receives, it verifies the user by alerting the user to enter the fingerprint and obtain the fingerprint information $fp*$. Then the fingerprint module retrieves the stored hash *hfp* of the fingerprint from the trusted execution environment TEE. After that the fingerprint module checks whether $h(fp*)$ and *hfp* are equal. If so, the fingerprint module sets the local verification result *LocResult* to "success". Otherwise, the fingerprint sets *LocResult* to "failure". If *LocResult* is "Failure", the fingerprint module repeats the local verification process. On the condition that the local verification fails three times consecutively; the authentication fails and stops the transaction authentication. Or else, the mobile terminal transmits *U*, *AuthLevel*, *LocResult*, and n_{u1} to *WS*. *WS* forwards the data to *AS*.

After *AS* receives, *AS* retrieves the stored *phn* in the database by the index of u. And then *AS* generates *DVC* and stores it in the database by the index of u. After that, *AS* sends *DVC* and to the mobile terminal by out-of-band way. After *U* receives, the user agent of the mobile terminal reads n_{u1} and checks whether the received temporary number and the stored temporary number are equal. If they are not equal, it proves that the data has been expired and the transaction authentication fails. Otherwise, the mobile terminal generates the temporary number n_{u2} and allows the user filling the received dynamic verification code $DVC*$ in the user agent. Afterwards, the mobile terminal sends $DVC*$ and n_{u2} to *WS*. *WS* forwards the data to *AS*. After *AS* receives, *AS* retrieves the stored dynamic verification code *DVC* and checks whether *DVC* and $DVC*$ are equal. If so, AS sets the transaction result *TranResult* to "Success". Otherwise, *AS* sets *TranResult* to "failure". At last, *AS* shows the transaction result *TranResult* to the mobile terminal like the basic authentication phase and returns the authentication result to the user.

5. Security Analysis

In our proposed scheme MTSAS, the security of the fingerprint is guaranteed by the TEE in the mobile terminal. And the reliability of the DVC ensured the reliability of the out-of-band transmission. Thus, the security of MTSAS depends on the data transmission way in the authentication phase. In MTSAS, the data transmission the user, the web server and the authentication server are followed by the data transmission way in the part 3. Therefore, we prove the security of the data transmission way by the BAN.

The security model shows that MTSAS has different security requirements for the data transmission between the user and the web server and the data transmission between the web server and the authentication server. For one thing, the data transmission between the user and the web server not only needs to guarantee the security of the data transmission but also meets the authentication needs. For another, the data transmission between the web server and the authentication server just needs to ensure the security of the data transmission. The meanings of symbols are shown in Table 3.

5.1. Modeling. Modeling the above process, we can get the following process. The temporary number and the key are related to the security. Therefore, we idealize them in the following message:

$$U \longrightarrow WS : \{n_u, n_w\}_{K_{uw}}, \{K_{uw}\}_{K_w}, \{n_u, n_w, K_{uw}\}_{Ku^{-1}} \quad (7)$$

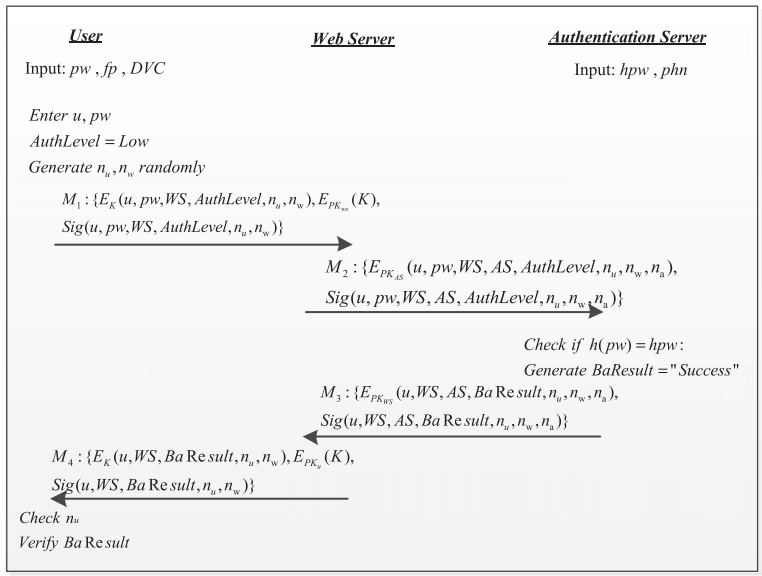


Figure 2: Basic authentication.

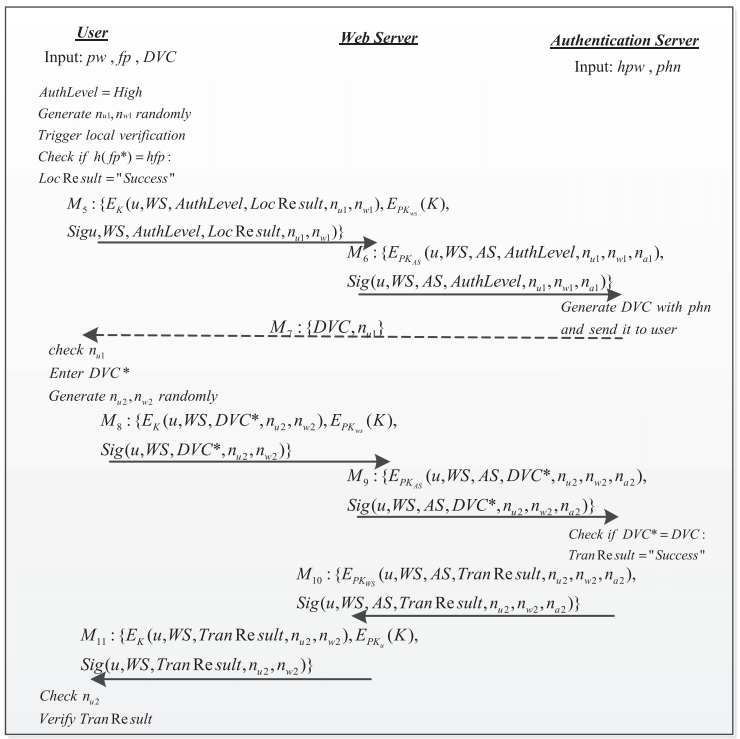


Figure 3: Transaction authentication.

$$WS \longrightarrow AS : \{n_u, n_w, n_a\}_{K_A}, \{n_u, n_w, n_a\}_{K_W{}^{-1}} \quad (8)$$

$$AS \longrightarrow WS : \{n_u, n_w, n_a\}_{K_W}, \{n_u, n_w, n_a\}_{K_A{}^{-1}} \quad (9)$$

$$WS \longrightarrow U : \{n_u, n_w\}_{K_{uw}}, \{K_{uw}\}_{K_u}, \{n_u, n_w, K_{uw}\}_{Kw^{-1}} \quad (10)$$

5.2. Protocol Analysis. The basic beliefs of MTSAS are as follows:

$$U \text{ believes } \xrightarrow{K_W} WS \quad (11)$$

$$WS \text{ believes } \xrightarrow{K_u} U \quad (12)$$

$$WS \text{ believes } \xrightarrow{K_A} AS \quad (13)$$

$$AS \text{ believes } \xrightarrow{K_W} WS \quad (14)$$

$$WS \text{ believes } U \text{ controls } U \xleftrightarrow{K_{uw}} WS \quad (15)$$

$$U \text{ believes } U \xleftrightarrow{K_{uw}} WS \quad (16)$$

Table 3: The meanings of symbols.

Symbol	Meaning
K_i	the public key of i
K_i^{-1}	the private key of i
$\{X\}K$	encrypts X by the key K
$\{X\}_{K_i^{-1}}$	signs X with the public key of i
K_{ij}	the session key between i and j

$$U\ believes\ fresh(n_u) \tag{17}$$

$$WS\ believes\ fresh(n_w) \tag{18}$$

$$AS\ believes\ fresh(n_a) \tag{19}$$

$$WS\ believes\ AS\ controls\{n_a\} \tag{20}$$

$$AS\ believes\ WS\ controls\{n_w\} \tag{21}$$

From beliefs (11), (12), (13), (14), U believes that K_w is the public key of WS, WS believes that K_u is the public key of U, WS believes that K_A is the public key of AS, and AS believes that K_w is the public key of WS. From belief (15), the session is sponsored by U, so WS believes that U has the jurisdiction of the session key K_{uw} between U and WS. From belief (16), U believes that K_{uw} is the session key between U and WS. From belief (17), U believes that n_u is fresh. From belief (18), WS believes that n_w is fresh. From belief (19), AS believes that n_a is fresh.

First, we analyze the data transmission between U and WS. By message (7) and the BAN logic axiom (6), we can obtain the following formula:

$$WS\ sees\{n_u, n_w, K_{uw}\} \tag{22}$$

By message (7), the basic belief (12), and the BAN logic axiom (2), we can receive the following formula:

$$WS\ believes\ U\ said\{n_u, n_w, K_{uw}\} \tag{23}$$

By the basic belief (18) and BAN logic axiom (6), we can gain the following formula:

$$WS\ believes\ fresh(n_u, n_w, K_{uw}) \tag{24}$$

By formula (23), formula (24), and the BAN logic axiom (3), we can get the following formula:

$$WS\ believes\ U\ believes\{n_u, n_w, K_{uw}\} \tag{25}$$

That is the formula

$$WS\ believes\ U\ believes\ U \overset{K_{uw}}{\longleftrightarrow} WS \tag{26}$$

By the basic belief (15), formula (26), and the BAN logic axiom (4), we can obtain the following formula:

$$WS\ believes\ U \overset{K_{uw}}{\longleftrightarrow} WS \tag{27}$$

By the message (10) and the BAN logic axiom (6), formula (28) can be received:

$$U\ sees\{n_u, n_w, K_{uw}\} \tag{28}$$

By message (10), the basic belief (11), and the BAN logic axiom (2), we can gain the following formula:

$$U\ believes\ WS\ said\{n_u, n_w, K_{uw}\} \tag{29}$$

By the basic belief (17) and the BAN logic axiom (5), formula (30) can be gained:

$$U\ believes\ fresh(n_u, n_w, K_{uw}) \tag{30}$$

By formula (28), formula (29), and BAN logic axiom (3), we can get the following formula:

$$U\ believes\ WS\ believes\{n_u, n_w, K_{uw}\} \tag{31}$$

That is formula (32):

$$U\ believes\ WS\ believes\ U \overset{K_{uw}}{\longleftrightarrow} WS \tag{32}$$

By the basic belief (16), we can obtain the following formula:

$$U\ believes\ U \overset{K_{uw}}{\longleftrightarrow} WS \tag{33}$$

Therefore, we can obtain the final beliefs: formulae (26), (27), (32), and (33). The final beliefs (27) and (33) are the beliefs in level 1. The final beliefs (26) and (32) are the beliefs in the level 2. Thus, the communication between the user and the web server not only achieves the purpose of the authentication, but also guarantees the security of the data transmission.

Next, we analyze the communication between WS and AS. By message (8) and the BAN logic axiom (6), we can receive the following formula:

$$AS\ sees\{n_u, n_w, n_a\} \tag{34}$$

By message (8), the basic belief (14), and the BAN logic axiom (2), formula (35) can be obtained:

$$AS\ believes\ WS\ said\{n_u, n_w, n_a\} \tag{35}$$

By the basic belief (19) and the BAN logic axiom (5), we can gain the following formula:

$$AS\ believes\ fresh(n_u, n_w, n_a) \tag{36}$$

By formula (35), formula (36), and the BAN logic axiom (3), formula (37) can be gotten:

$$AS\ believes\ WS\ believes\{n_u, n_w, n_a\} \tag{37}$$

That is formula (38) and formula (39):

$$AS\ believes\ WS\ believes\{n_w\} \tag{38}$$

$$AS\ believes\ WS\ believes\{n_a\} \tag{39}$$

By the basic belief (21), formula (38), and the BAN logic axiom (4), we can receive formula (40):

$$AS\ believes\ \{n_w\} \tag{40}$$

Similar to the analysis process to message (8), we can analyze message (9). Then, we can obtain formula (41) and formula (42):

$$WS\ believes\ AS\ believes\ \{n_w\} \tag{41}$$

$$WS\ believes\ \{n_a\} \tag{42}$$

As a result, we can obtain the final beliefs: formulae (39), (40), (41), and (42).

The final beliefs (40) and (42) are beliefs in level 1. The belief (40) means that *AS* believes that the random number n_w is really from *WS*. In other words, *AS* believes *WS*. And also, the result shows that *WS* believes that *AS* can be obtained similarly in the final belief (42).

The final beliefs (39) and (41) are beliefs in level 2. The belief (39) represents that *AS* believes that *WS* believes that the random number n_a is really from *AS* itself. That is to say, *AS* belief itself is believed by *WS*. And also, Belief (41) explains that *WS* belief itself is believed by *AS* similarly.

Therefore, the security of the data transmission between WS and AS can be ensured; the security of the proposed scheme MTSAS is proved.

6. Performance and Security Comparison

In this part, on the one hand, we compare the performance of the proposed scheme MTSAS with the Shen's scheme in [15] and the Li's scheme in [16] in the authentication stage. On the other hand, the security of three schemes is also compared. Because MTSAS has two types of authentication, there we compare the performance of our proposed scheme in the two aspects of MTSAS1 and MTSAS2. MTSAS1 is the basic authentication and MTSAS2 is the transaction authentication.

6.1. Performance Comparison. The performance of the MTSAS is measured in three aspects, the symmetric encryption/decryption computation time t_{sym}, the public key encryption/ decryption computation time t_{asym}, and the hash computation time t_{hash}.

The public key encryption/decryption computation time involves the exponential computation, which costs the most computational resources. Thus the number of the public key encryption/decryption computation will have crucial influence on the performance of a scheme. Because the symmetric encryption/decryption computation does not involve the exponential computation, typically it is related to addition, subtraction, multiplication, division, or some low order operations and costs less computational resources. The number of the symmetric encryption/decryption computation has less effect on the scheme's performance. The hash computation only refers to the one-way hash function operation and consumes the least resources. Therefore the number of the hash computation has the least impact on the performance. From [26], when the public key encryption/decryption computation adopts the ECC algorithm, the symmetric encryption/decryption computation adopts AES-128, and the hash algorithm adopts the SHA-1 algorithm, $t_{asym} \approx 7.3592ms$, $t_{sym} \approx 0.005ms$, $t_{hash} \approx 0.0004ms$. The detailed performance comparison is shown in Table 4.

Table 4: Performance comparison.

	Computational Overhead (ms)
Shen in [16]	$6t_{asym}+17t_{hash}=44.162$
Li in [25]	$6t_{asym}+22t_{hash}=44.164$
MTSAS1	$4t_{sym}+4t_{asym}+10t_{hash}=29.460$
MTSAS2	$6t_{sym}+6t_{asym}+16t_{hash}=44.284$

Shen's scheme [16] consists of public key encryption/decryption computations with 6 times and hash computations with 17 times in the authentication phase. Li's scheme [25] includes public key encryption/decryption computations with 6 times and hash computations with 22 times in the authentication phase. Our proposed scheme of the basic authentication MTSAS1 consists of public key encryption/decryption computations with 4 times, symmetric encryption/decryption computations with 4 times, and hash computations with 10 times in the authentication phase. Significantly, the number of the public key encryption/decryption computation in MTSAS1 is lower than those schemes of Shen and Li. Therefore, MTSAS1 has lower computational overhead, outstanding performance, and superior user experience. Last but not least, the transaction authentication MTSAS2 concludes public key encryption/decryption computations with 6 times, symmetric encryption/decryption computations with 6 times and hash computations with 16 times in the authentication phase. The MTSAS2 has the same number of the public key encryption/decryption computation as Shen's scheme and Li's scheme, but due to inserting the symmetric encryption/decryption computation. Therefore, the amount of computational time of the MTSAS2 is basically equal and to Shen's scheme and Li's scheme, just slightly higher. But the MTSAS2 mainly pays attention to the high security demand in the mobile payment scenario and enhances the security of the user's biometric characteristic. The MTSAS2 solves the security hidden danger due to the fingerprint information leakage and provides users with the strong authentication. As a result, the MTSAS2 improves the security when the user authenticates the identity. Besides, the computational resources that the symmetric encryption/decryption computation takes are limited. Therefore, the minor performance loss that the MTSAS2 takes can be acceptable

6.2. Security Comparison. In this part, we mainly focus on the security of three schemes. The detailed security comparison is presented in Table 5.

As is shown in Table 5, because these three schemes all adopt the signature to resist forgery; they all satisfy the forward security. So they can guarantee the message's confidentiality and integrity and is able to resist replay attack and key guessing attack. Since the biometrics template is

Table 5: Security comparison.

	Shen in [16]	Li in [25]	MTSAS
Message confidentiality	√	√	√
Message integrity	√	√	√
Resist key guessing attack	√	√	√
Resist replay attack	√	√	√
Forward Security	√	√	√
User anonymity	✗	√	√
Resist biometrics template attack	✗	√	√
Resist data leakage of the server	✗	✗	√
Two servers	✗	✗	√
Divide the security level	✗	✗	√

stored in the user's smart card in Shen's scheme. Therefore, it is vulnerable to biometrics template lost attack that the user's biometrics template can be retrieved if the smart card is stolen by an adversary. Of course, Shen's scheme also cannot meet the requirement of the user anonymity. On the contrary, on account that Li's scheme makes anonymity of the biological templates, it can meet user anonymity and resist biometrics template attack. The biological characteristics are in the trusted execution environment TEE in our proposed scheme. The biological characteristics never leave the local device. And when the server authenticates the user, what the server authenticates is the user device, not the user. The user biometric data never stored any data in the server side. Therefore, MTSAS can also satisfy the requirements of the user anonymity and is able to resist biometrics template attack.

But when the data in the server leaks, the biological characteristics of the user might be captured by the attackers to greatly threat user privacy. Unfortunately, the server stores the data related to the user biometric in the schemes of the Shen and the Li. Therefore, they both cannot resist this attack. Whereas any data related to the biological characteristics is never stored in the server side in MTSAS. So MTSAS is more superior to other two schemes in improving the security of the user data and protecting the user privacy. And compared with the former two schemes, MTSAS adopts the two-server authentication method to disperse security risks. So the stability and the robustness of the MTSAS are greatly improved. At the same time, the MTSAS analyzes user behavior in detail. According to the authentication scenario that the user is in, the MTSAS provides the authentication method with different security level. MTSAS provides the weak authentication in the basic authentication scenario and the strong authentication in the transaction scenario. Our proposed scheme MTSAS stands on the user's point to perfect the user experience and improves the authentication efficiency. Thus, the practicability and the feasibility of our proposed scheme MTSAS are both stronger

7. Experiment

In the experiment part, we realize FREDP based on the FIDO UAF framework. And then we test the performance of basic authentication and transaction authentication, respectively.

FIDO (Fast Identity Online) is an online authentication alliance sponsored by PayPal, Nok Nok, validity, Infineon, AGNITIO, and Lenovo in July 2012. And till September 2016, alliance members have amounted to 252. Specific to the problem of "isolated island" resulting from multiple authentication standards in the mobile identity authentication, FIDO dedicates to unified mobile authentication standard in order to reduce the user dependence to password. FIDO aims to solve the problem of security, convenience, and privacy and provides users with strong authentication. The FIDO architecture has been widely applied in the mobile payment area. Several companies such as Alipay, JD finance, and ICBC have fully supported FIDO. The FIDO framework includes standards of UAF and U2F. UAF (Universal Authentication Framework Protocol) is able to realize without password by binding users' biological characteristics with mobile device. UAF makes the authentication mode, server authenticates device and device authenticates user, come true, which greatly improves the convenience and the reliability of mobile identity authentication. U2F (2nd Universal Framework Protocol) provides the two-factor experience to users by adding the authentication information in hardware devices (such as U shield and mobile hardware information) as the second factor. So FIDO can make the authentication strong.

We implement FREDP based on the FIDO UAF framework. In the mobile terminal, in consideration of current authentication status and user requirements, our proposed scheme provides the authentication way of distinguishing security levels. The basic authentication adopts one-factor authentication way based on password. And the transaction authentication way adopts multifactor authentication way based on password, DVC, and fingerprint. In the server side, in order to guarantee the stability of the server and the security of authentication server, FREDP adopts the two-server mode. The web server is deployed prior to FIDO authentication server. This two servers are both deployed in the cloud environment.

7.1. Experiment Environment. Our experimental environment is as follows.

The mobile terminal adopts ZUK Z2125; the hardware is configured as CPU 2.35HZ ROM 64GB, RAM 6GB, and android version: 6.0.1. The cloud environment is the public cloud service, and the web server and authentication server are deployed on two cloud hosts with the same configuration. The cloud host is configured as Ubuntu 14.04 32-bit server version, with 1GB kernel, 1Mbps bandwidth, and 50G hard disk.

7.2. Experiment Process. In the concrete implementation process, we write code in JAVA language based on FIDO UAF framework. FREDP scheme is realized by FIDO server as authentication server and web server writing by ourselves.

Figure 4: Basic authentication schematic.

During the FREDP operation, in the mobileterminal, we conduct the basic authentication with the user test 03. The basic authentication schematic is shown in Figure 4. The transaction authentication schematic is shown in Figure 5.

In the public cloud, the log of the web server is shown in Figure 6, and the console log of the authentication server is shown in Figure 7.

After the realization of FREDP, we make experiment in four aspects. In order to ensure the reliability and the accuracy of the data, we test three times and average them for each class of data.

7.2.1. Experimental Target: Comparing the Total Authentication Time T_{sum} of Single Factor/Multifactor Authentication Method. We analyze the impact of increasing identity authentication factors under the two-server architecture to the total authentication time. In the FREDP, actually, the basic authentication is the method of the single factor authentication, namely, the password. The transaction authentication is the method of the multifactor authentication, namely, the password, the dynamic verification code, and the fingerprint. Thus, it is necessary to compare the authentication time between the basic identification and the transaction authentication. The test results are shown in Table 6. The timing interval of the authentication time in Table 6 is from the time that the mobile terminal collects all the authentication factors to the time that the mobile terminal receives the authentication result.

Table 6: Comparison of single factor/multifactor authentication.

	Basic Authentication (ms)	Transaction Authentication (ms)
T_{sum1}	252	364
T_{sum2}	277	381
T_{sum3}	283	406
T_{sum}	270.7	383.7

As shown in Table 6, due to the addition of dynamic verification code and the fingerprint as authentication factors, the authentication time of the transaction authentication increases to a certain extent. In the transaction authentication, the comparison inside the mobile terminal would take some time, such as the fingerprint comparison and attestation in the trusted execution environment. And at the same time the increase of the interaction times between the web server and the authentication server may cause the growth of the authentication time. But it is because of these verification and interactive process, the FREDP possesses high the high security that other schemes do not have. Our proposed FREDP avoids the fingerprint information disclosure, realizes fingerprint's local storage and local verification. But luckily, we can see in Table 6, the transaction authentication with multifactor only increases 1 second. It is absolutely acceptable for the user. So FREDP makes a good compromise between the convenience and the security. Therefore, it is feasible and necessary to adopt multifactor authentication in the key system.

7.2.2. Experiment Target: Comparing the Total Authentication Time in the Single Server/Two-Server Environment. In this section, we test the basic authentication and the transaction authentication, respectively. In the single server framework, the mobile terminal directly interacts with the authentication server, and all the forwarding and verification work are deployed in the authentication server side. There are the web server and the authentication server in the two-server framework. The web is responsible for communicating with the mobile terminal and forwarding the information. The authentication server is responsible for the verification and processing of the authentication process data. The test result of the total time T_{bsum} in the basic authentication under the single server/two-server authentication is, respectively, shown in Table 7. The test result of the total time T_{tsum} in the transaction authentication under the single server/two-server authentication is, respectively, shown in Table 8. Similarly, the timing interval of the authentication time in Tables 7 and 8 is, from the time that the mobile terminal collects all the authentication factors to the time that the mobile terminal receives the authentication result.

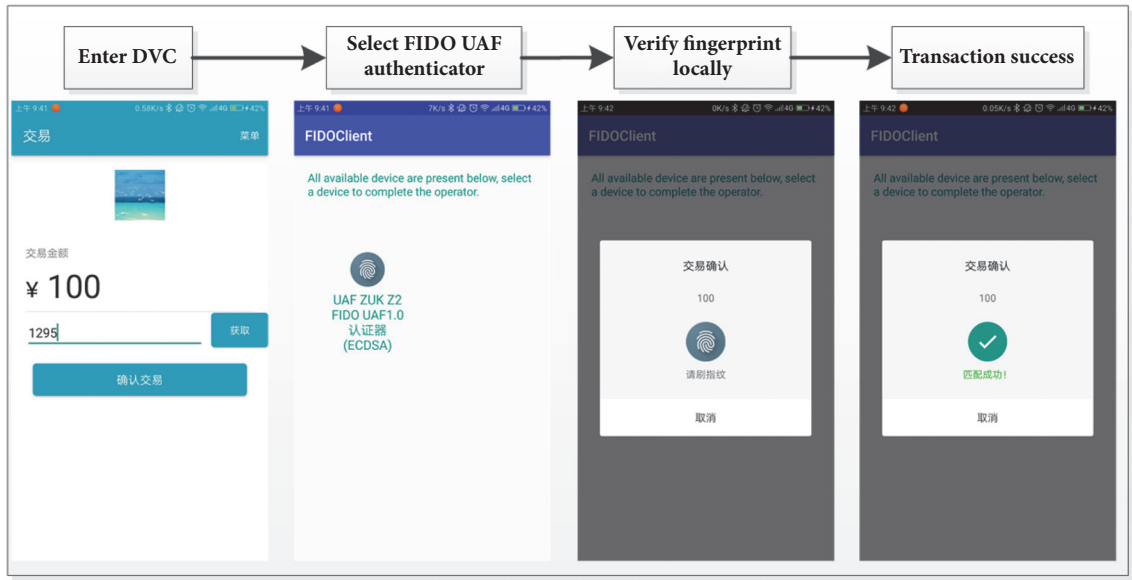


FIGURE 5: Transaction authentication schematic.

登录
2018-04-10 09:39:32,883 INFO [server.modules.user.web.TestLoginController] - username:test03,password:1234
2018-04-10 09:39:38,578 INFO [server.modules.user.web.TestLoginController] - Login OK!
注册请求
2018-04-10 09:40:12,797 INFO [server.modules.user.web.TestLoginController] - Get from client:{"op":"Reg","previousRequest":"test03"}
2018-04-10 09:40:21,784 INFO [server.modules.user.web.TestLoginController] - Get from authentication server:{"statusCode":1200,"uafRe
2018-04-10 09:40:22,003 INFO [server.modules.user.web.TestLoginController] - Send to client:ReturnUAFRequest [statusCode=1200, uafReq
AgU21nbmFOdXJlIEFsZ29yaXRobTogUOhBMjU2d2l0aFJTQSwgTO1EIDOgMS4yLjgOMC4xMTM1NDkuMS4xLjExCgogIEtleTogIFN1biBSUOEgcHVibGljIGtleSwgMjAOOCBi
k6IFtGcm9tOiBUaHUgTWF5IDA0IDEz0jQ5OjU4IENTVCAyMDE3LAogICAgICAgICAgICAgICBUbzogTW9uIE1heSAwNCAxMzo0OTo1OCBDU1QgMjAyMFOKICBJc3N1ZXI6IENO1
AwMDAwMDEwMC5jcmxdCiAgIENSTElzc3VlcjpbQO49Y2EtMTQ1LCBDPUNOXQpdXQoKWzRdOiBPYmplY3RJZDogMi41LjI5LjE1IENyaXRpY2FsaXR5PWZhbHNlCktleVVzYW1l
AwNDA6IEI3IDA4IEI2IDg4IEZDIDcxIEQwIDkxICAgNjcgQkMgNDkgMkY gOTY gQTkgOTcgOTkgIC4uLi4uc54uZy5JLy4uLi4NCjAwNTA6IEJDIDI4IENCIDIyIE45IEJEIEE4:
ZBIEIyIDE2IDQ3ICAgOEUgOEQgQjgg0TUgMjcgQTAgNjggMjAgIEwuLi5qLi5HLi4uLicuaCANCjAwRjA6IEVEIERGIEVFIDAzIDc3IDAzIEQyIDNBICAgNUEgQTAgMEEgRUYg(
注册响应
2018-04-10 09:40:53,837 INFO [server.modules.user.web.TestLoginController] - Get from client:{"uafResponse":"{\"uafProtocolMessage\":'
3YW5nQG5va25va2xhYnMuY24wWTATBgcqhkjOPQIBBggqhkjOPQMBBwNCAAQsCSTWZx3DdOKol3w2HRM6ww_SqWqiaK8XFKGRBDQeW7bmHbFbaLzMdRwuicqTrpG1Q94pmo-hD_
UTRNekl5TVRjNE16WTRPRF16TURjeU9ESTNPVFl5TmpJeE9UY3hNekOOT1RVeU1UVTBORGswTORNME5ETTJNakE1T1RnNU9EQXhPRFkxTmprNU1qVTNOVGN6TURrMk56ZzBNak
WTFPRFkOTWpZNU1qUXdNall5TURRME5ERUtJQOJ3ZFdKc2FXTWdaWGh3YjI1bGJuUTZJRFkxT1RNMONpQWdWbUZzYVdScGRIazZJRnRHY205dE9pQ1VhSFVnVFdGNU1EQTBJRE
NDFMakk1TGpNeElFTnlhWFJwWTJGc2FYUjVQV1poYkhObENrT1NURVJwYzNSeWFXSjFkR2x2YmxCdmFXNTBjeUJiQ21BZ1cwUnBjM1J5YVdKMWRHbHZibEJ2YVc1ME9nb2dJQOI
QTFRUOF4TONBZO1TNHVMaTVoSVMOdUxpNTBMQzVhTGcwSO1EQXhNRG9nTkRNZO5EWWdOVFVnTORZZO1DQWdJQOFnSUNBZO1DQWdJQOFnSUNBZO1DQWdJQOFnSUNBZO1DQWdJQOI
NHVMaTR1TGkOdVNDNE5DakF3TmpBNkIEVTVJRFJDSURkQklFRTRJRGszSURBNElEbEVJRGd4SUNBZ1JVUWdOVVFnUVRZZO5rTWdRamNnTXprZ1JUZ2dOa1FnSUZsTGVpNHVMaTI
VWdOekFnTVRVZO1FWWdRaO1nUTBNZO1DNWJMaTR1VmkOckxtRjFjQzR1TGkOTkNqQXdSVEE2SURSRE1FTTFJRGxCSURFMO1EWkJJRU15SURFMklEUTNJQOFnTOVVZO9FUWdRam
2018-04-10 09:41:03,966 INFO [server.modules.user.web.TestLoginController] - Send to client:{"statusCode":1200,"description":"OPERATI(
认证请求
2018-04-10 09:41:50,826 INFO [server.modules.user.web.AuthenticationController] - Get from client:GetUAFRequest [op=Auth, previousReq
2018-04-10 09:41:51,197 INFO [server.modules.user.web.AuthenticationController] - Get from authentication server:ReturnUAFRequest [st
n\":0,\"attachmentHint\":0,\"tcDisplay\":0,\"authenticatorVersion\":0}]}]}]","additionalData":""}, op=Auth, lifetimeMillis=300000]
2018-04-10 09:41:51,197 INFO [server.modules.user.web.AuthenticationController] - Send to client:{"statusCode":1200,"uafRequest":"{\"
"001A#2121\\\"],\\\"keyIDs\\\":[\\\"A4BR1VZvYmeUW3EfHwKM6uo_1wzQUP_ZA_IC5QHeg9U\\\"],\\\"userVerification\\\":0,\\\"keyProtection\\\":(
4NDMxMjA5NTk3NzYyNzg5MzA5MjYwMTg5ODQwNjA3NzM2MDc2OTY2OTEOMzE5NDY5OTU2NDMxMDk5NDkxMDA1MDc5NzIyNDk4NTMOMjU4NzA3NTQxMzcOMjY2NjE1ODg4ODAyO(
gei4uLgOKXQpdCgpbMlO6IE9iamVjdElkOiAyLjUuMjkuMTkgQ3JpdGljYWxpdHk9ZmFsc2UKQmFraWNDb25zdHJhaW5OczpbCiAgQOE6ZmFsc2UKICBQYXRoTGVuOjIxNDcOO(
gICAgICAgICAgICAgICAgICAgQOZVLgOKXQpdCgpdCiAgQWxnb3JpdGhtOiBbUOhBMjU2d2l0aFJTQVOKICBTaWduYXR1cmU6CjAwMDA6IDkxIEY5IEJEIEZEIEY1IDY1IEVFIDQyI(
5IDE2IEIwICAgMUYgNjggQjEgMTEgREUgNTAgMOIgOUIgIC5KLi4rLi4uLmguLi5QOy4NCjAwQTA6IEIyIEQOIEU1IDc2IEMyIEIwIEFCIDU1ICAgMDYgODUgNDYgRkIgODkgRI
认证响应
2018-04-10 09:42:26,401 INFO [server.modules.user.web.AuthenticationController] - Get from client:{"uafResponse":"{\"uafProtocolMessa
CVGRXSnFaVO4wT21CRFRqMTNkM2N1Wm1sa2IzUmxjM1F1WTI5dExDQkRQVU5FQ21BZ1UybG5ibUYwZFhKbE1FRnNaMjl5YVhSb2JUb2dVMGhCTWpVMmQybDBhRkpUUVN3Z1Qwbl
Jd01ERTNOREUyTURVMU9EWTFNemN3T1RneE1qUTJNakUyT1RjMk1qZ3pNam85IXpJME5qWXhPVGcyTnpFek9UVTROVEF3TkRrMO1UWTBNVEkOT1Rnek9UZzVNRGN3T1RJMU56W2
FZO1EVWdPVGNnUVVNZO1DQWdJQOFnSUNBZO1DQWdJQOFnSUNBZO1DQWdJQOFnSUNBZO1DQWdlaTR1TGcwS1hRcGRDZ3BiTlYwNkdFOW1hbVZqZEVsaO9pQ1
WdWRBb2dJRVJoZEdGZ1JXNWphWEJvWlhKdFpXNTBDaUFnUzJWNVgwRm5jbVZsYldWRBcGRDZ3BiTlYwNkdFOW1hbVZqZEVsaO9pQX1MalV1TWprdU1UUWdRMOpwZEdsallXe
ZZ1FqRWdNRU1nTjBJZO56SWdOak1nSUVBdUxpNHVMbGN1VWk1VOxpNHVjbU1OQ2pBd05EQTZJRUkzSURBNElFSTJJRGc0SUVaRE1EY3hJRVF3SURreE1DQWdOamNuUWtNZO5Ea
KQklEazBJREOOSURWRO1EbEdJQOFnT0RJZO5FWWdSRUVnUmtJZ1EwUWdSalFnTlRNZO5FRWdJQzR1TGkOdU9GOHVMazh1TGkOdVUwbO5DakF3UXpBNklFUXpJRUO:SURCQO1FT
IcE9TR2h4VG1rMA\\\",\\\"upv\\\":{\\\"major\\\":1,\\\"minor\\\":0}}]}\"}
2018-04-10 09:42:26,590 INFO [server.modules.user.web.AuthenticationController] - Send to client:{"statusCode":1200,"description":"OPI

FIGURE 6: Log of the web server.

As shown in Tables 7 and 8, whether in the basic authentication or the transaction authentication, the total authentication time in the two servers takes 1 seconds more than in the single server. It is acceptable for the user. And the mode of the two-server authentication shares the authentication risk, which avoids all verification information leakage because the server is attacked in the single server. We adds a layer of "shelter" to improve the security of the verification information and the reliability of the authentication process. Therefore, the design of the two servers is necessary and reasonable.

8. Conclusion

In allusion to the security problem of the fingerprint leakage in the mobile payment environment, we propose a novel multifactor two-server authentication scheme under mobile

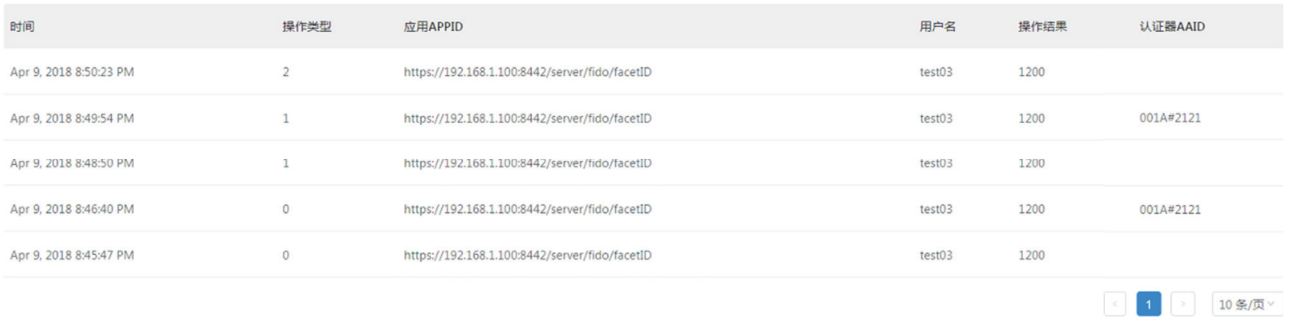


Figure 7: Console log of the web server.

Table 7: Comparison of single factor/multifactor authentication in the basic authentication.

	Single Server (ms)	Two Servers (ms)
T_{bsum1}	191	252
T_{bsum2}	193	277
T_{bsum3}	196	283
T_{bsum}	193.3	270.7

Table 8: Comparison of single factor/multifactor authentication in the basic authentication.

	Single Server (ms)	Two Servers (ms)
T_{tsum1}	288	364
T_{tsum2}	228	381
T_{tsum3}	295	306
T_{tsum}	270.3	383.7

computing MTSAS. In the MTSAS, the user's biometric characteristics cannot leave the user device. And the server side never stores the user's fingerprint information. Particularly, the authentication server is applied by the private cloud. The use of the two servers can obviously lower the security risk of server attack. Moreover, we stand on the user's shoes. MTSAS provides the different authentication factors depending on the privacy level of the authentication. Thus our work has a certain degree of contribution to the mobile payment security.

Unfortunately, the problem of the loss of the user device still troubled us. In this conditions, apparently, it may take the problem of data redundancy in the cloud when the user registration again. We will pay more attention to the authentication problem in this part.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported by National Natural Science Foundation of China (61671360 and 61672415), the Key Program of NSFC-Tongyong Union Foundation under Grant U1636209, the National Key Basic Research Program (2017YFB0801805), the Key Program of NSFC Grant U1405255, the Natural Science Basic Research Plan in Shaanxi Province of China (2017JM6082), and the Opening Project of Science and Technology on Communication Networks Laboratory (KX172600024)

References

[1] H. Suo, Z. Liu, J. Wan, and K. Zhou, "Security and privacy in mobile cloud computing," in *Proceedings of the 2013 9th International Wireless Communications and Mobile Computing Conference, IWCMC 2013*, pp. 655–659, Sardinia, Italy, July 2013.

[2] W. Song and X. Su, "Review of Mobile cloud computing," in *Proceedings of the IEEE 3rd International Conference on Communication Software and Networks (ICCSN '11)*, pp. 1–4, Xi'an, China, May 2011.

[3] V. Chang, Y.-H. Kuo, and M. Ramachandran, "Cloud computing adoption framework: A security framework for business clouds," *Future Generation Computer Systems*, vol. 57, pp. 24–41, 2016.

[4] G. Mathew and S. Thomas, "A novel multifactor authentication system ensuring usability and security," *International Journal of Security, Privacy and Trust Management*, vol. 2, no. 5, pp. 21–30, 2013.

[5] D.-J. He, M.-D. Ma, Y. Zhang, C. Chen, and J.-J. Bu, "A strong user authentication scheme with smart cards for wireless communications," *Computer Communications*, vol. 34, no. 3, pp. 367–374, 2011.

[6] G. Wang, Q. Liu, J. Zhou, and J. Z. Chen, "A multi-factors identity authentication scheme in classified environment," *Advanced Materials Research*, vol. 765-767, pp. 1734–1738, 2013.

[7] A. Bhargav-Spantzel, A. C. Squicciarini, S. Modi, M. Young, E. Bertino, and S. J. Elliott, "Privacy preserving multi-factor authentication with biometrics," *Journal of Computer Security*, vol. 15, no. 5, pp. 529–560, 2007.

[8] A. Tiwari, S. Sanyal, A. Abraham et al., "A multi-factor security protocol for wireless payment - secure web authentication using mobile devices," *Computer Science*, 2011.

[9] X. Yi, S. Ling, and H. Wang, "Efficient two-server password-only authenticated key exchange," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 9, pp. 1773–1782, 2013.

[10] Y. Yang, R. H. Deng, and F. Bao, "A practical password-based two-server authentication and key exchange system," *IEEE Transactions on Dependable and Secure Computing*, vol. 3, no. 2, pp. 105–114, 2006.

[11] C. Lee, T. Lin, and R. Chang, "A secure dynamic ID based remote user authentication scheme for multi-server environment using smart cards," *Expert Systems with Applications*, vol. 38, no. 11, pp. 13863–13870, 2011.

[12] D. F. L. Souza, A. M. F. Burlamaqui, and G. L. S. Filho, "A multi factor authentication approach based on biometrics, optical interference and chaotic maps," *IEEE Latin America Transactions*, vol. 15, no. 9, pp. 1700–1708, 2017.

[13] H. Saevanee, N. L. Clarke, and S. M. Furnell, "Multi-modal behavioural biometric authentication for mobile devices," *IFIP Advances in Information and Communication Technology*, vol. 376, pp. 465–474, 2012.

[14] P. Baraki and V. Ramaswamy, "Biometric authentication of a user using online dynamic signature," in *Proceedings of the 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016*, pp. 576–581, Bangalore, India, July 2016.

[15] H. Shen, C. Gao, D. He, and L. Wu, "New biometrics-based authentication scheme for multi-server environment in critical systems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 6, no. 6, pp. 825–834, 2015.

[16] X. Li, K. Wang, J. Shen, S. Kumari, F. Wu, and Y. Hu, "An enhanced biometrics-based user authentication scheme for multi-server environments in critical systems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 7, no. 3, pp. 427–443, 2016.

[17] Z. Han, L. Yang, and Q. Liu, "A Novel Multifactor Two-Server Authentication Scheme under the Mobile Cloud Computing," in *Proceedings of the 2017 International Conference on Networking and Network Applications (NaNA)*, pp. 341–346, Kathmandu, Nepal, October 2017.

[18] A. Roy, N. Memon, and A. Ross, "MasterPrint: exploring the vulnerability of partial fingerprint-based authentication systems," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 9, pp. 2013–2025, 2017.

[19] Y. Xue, Y.-a. Tan, C. Liang, Y. Li, J. Zheng, and Q. Zhang, "RootAgency: A Digital signature-based root privilege management agency for cloud terminal devices," *Information Sciences*, vol. 444, pp. 36–50, 2018.

[20] D. Pointcheval and S. Zimmer, "Multi-factor authenticated key exchange," in *International Conference on Applied Cryptography and Network Security*, vol. 5037 of *Lecture Notes in Computer Science*, pp. 277–295, 2008.

[21] S. Layeghian Javan and A. Ghaemi Bafghi, "An anonymous mobile payment protocol based on SWPP," *Electronic Commerce Research*, vol. 14, no. 4, pp. 635–660, 2014.

[22] S. H. Khan and M. A. Akbar, "Multi-factor authentication on cloud," in *Proceedings of the 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–7, Adelaide, Australia, November 2015.

[23] C.-C. Chang and T.-F. Cheng, "A robust and efficient smart card based remote login mechanism for multi-server architecture," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 8, pp. 4589–4602, 2011.

[24] C. T. Li, C. Y. Weng, and C. I. Fan, "Two-factor user authentication in multi-server networks," *International Journal of Security & Its Applications*, vol. 6, no. 2, 2012.

[25] M. Burrows, M. Abad, and M. Needham, "A logic of authentication," *Proceedings of the Royal Society A Mathematical, Physical and Engineering Sciences*, vol. 426, no. 1871, pp. 233–271, 1989.

[26] L. Xu and F. Wu, "Cryptanalysis and improvement of a user authentication scheme preserving uniqueness and anonymity for connected health care," *Journal of Medical Systems*, vol. 39, no. 2, article 10, 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 8131367, 12 pages
https://doi.org/10.1155/2018/8131367

Research Article

An Anonymous Authentication Protocol Based on Cloud for Telemedical Systems

Wenmin Li,[1] Shuo Zhang ,[1] Qi Su,[2] Qiaoyan Wen,[1] and Yang Chen[1]

[1]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*State Key Laboratory of Cryptology, Beijing 100878, China*

Correspondence should be addressed to Shuo Zhang; zs1990@vip.qq.com

Received 10 April 2018; Accepted 2 August 2018; Published 2 September 2018

Academic Editor: Joseph Liu

Telecare medical information systems (TMIS) enable patients to access healthcare delivery services conveniently. With the explosive development occurring in cloud computing and services, storage of personal medical and health information outsourcing to cloud infrastructure has been a potential alternative. However, this has entailed many considerable security and privacy issues. In order to address the security loopholes, we propose a promising solution satisfying the requirements of cloud computing scenarios for telemedical systems. The proposed scheme could provide both data confidentiality and message authenticity while preserving anonymity. Furthermore, the formal security proof demonstrates that the proposed scheme is resistant to various attacks. The performance comparisons show the proposal's workability and it is well suited to adoption in telemedical services.

1. Introduction

With the explosion of cloud computing and services, there has been a growing trend to use the cloud for large-scale data storage and management. It is a new style of computing that offers dynamically scalable network services to external customers delivered over the Internet. Cloud computing provides a powerful underlying architecture for telemedicine, which is an emerging treatment mode for delivering appropriate healthcare services remotely. It facilitates medical practitioners and patients to establish communication over public networks and patients can acquire the medical services via electronic networks conveniently. This will significantly lower the social and economic expenses, while enhancing the medical quality and efficiency.

Cloud computing introduces a new way for medical systems to store and manage medical data, which is complex task. As wearable devices are becoming more and more powerful, patients can obtain their health information timely. They also could upload and access their medical records to the cloud through mobile devices. This can help medical institutions to quickly obtain patients' physical condition in urgent cases for proper medical diagnosis and treatment process. Any delay in the access to medical record at the time of emergency would cause severe errors, which profoundly affects patient's therapeutic process. In the cloud based telecare medical information systems, the cloud database is responsible for storing patients' critical medical data and updates it as the medical treatment availed by the patient. However, the storage of patients' electronic medical records such as personal information, medical records, and physiological parameters in the medical server may result in the exposure of patients' privacy. Cloud computing offers expansively developing prospects of new and better models of healthcare; it also raises some security issues due to new potential ways for data theft. And hence, safeguarding security and patients' privacy in cloud based telecare medical information systems are very significant. Authentication mechanism is a prerequisite to verify the legality of all participants and tackle the illegal access in distributed systems, such as wireless interface systems [1, 2], multiple server architecture based systems [3], smart card based system [4], and mobile radio systems [5, 6]. Furthermore, the anonymous authentication could protect users anonymity and prevent the disclosure of private information [7, 8]. Therefore, a secure authentication protocol is a proper

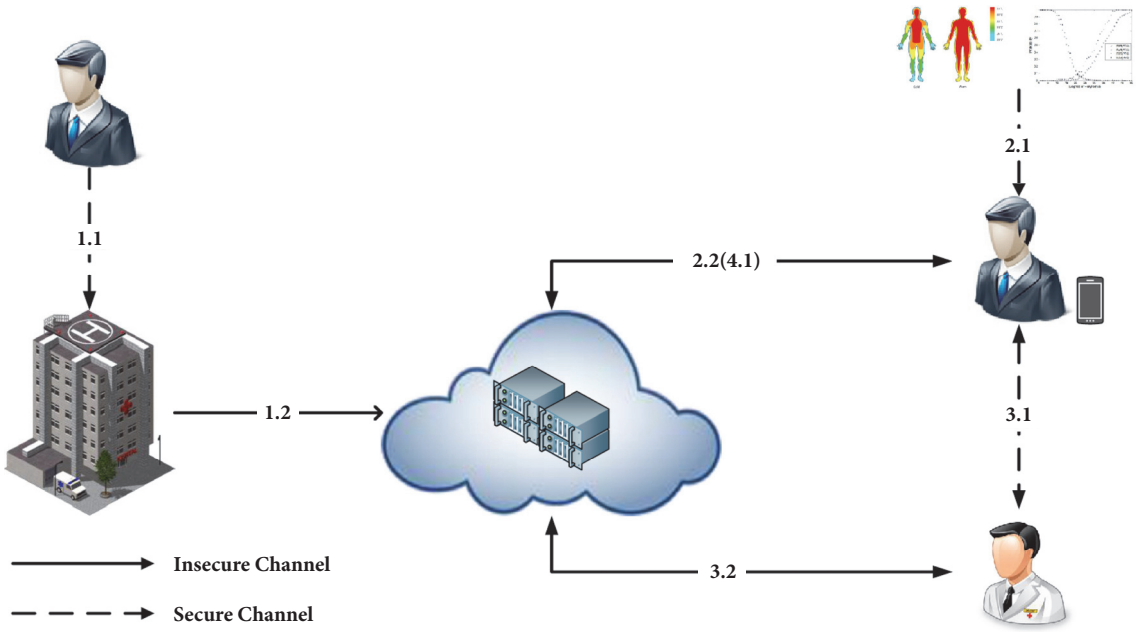


Figure 1: Structure of cloud based TMIS.

solution to provide security and privacy for TMIS [9–11]. Hitherto, authentication protocol for integration telemedical systems in cloud computing environment recently has drawn significant attention from academia [12–20].

In 2012, Padhy et al. [12] introduced a cloud based model for rural healthcare systems. In 2013, Banerjee et al. [13] presented a new architecture for cloud based healthcare application to serve patients in emergency. Nevertheless, their scheme is unable to offer confidentiality of transmitted data. One year later, Chen et al. [14] proposed a medical data exchange protocol in cloud computing environment. In their scheme, patients and doctors could be convenient to access medical resources outsourced in the cloud. Unfortunately, their scheme could not resist impersonation attacks or provide patient anonymity. To fix the defects, a modification was developed in the same year [15]. In 2016, Chiou et al. [16] showed that their scheme still lacks privacy protection and message authentication. Then, the authors proposed a new privacy authentication scheme based on cloud for TMIS which provided a "real" and complete telemedicine system. However, in 2007, Mohit et al. [17], Cheng et al. [20], and Li et al. [18] identified Chiou et al.'s protocol that failed to preserve patients' privacy and forward security and suffers from mobile device stolen attack, respectively. Meanwhile, Mohit et al. [17] and Cheng et al. [20] both presented an improved mechanism for cloud-assisted medical care systems. Recently, Li et al. [19] pointed out that Mohit et al.'s proposal also was susceptible to health report revelation and inspection report forgery attacks. In Cheng et al.'s scheme [20], the inputs of bilinear maps are generators in the corresponding cyclic groups, rather than random numbers of integer field Z_q. This will bring about errors in the authentication process.

In this paper, we design a telemedical information model based on cloud authentication which allows patients to remotely access medical services with privacy. Further, we discuss its security and prove that it can withstand various attacks. Compared with the state of the art, our scheme provides formal security proofs and achieves better efficiency in terms of computation cost. Performance and functionality analysis shows that it is more secure and practical for cloud based telemedicine system.

The remaining of this paper is organized as follows. Section 2 describes our robust cloud based authentication scheme for TMIS, together with formally proving its security in Section 3. Subsequently, we compare the performance with the previous schemes in Section 4. Finally, we draw the conclusions in Section 5.

2. The Proposed Scheme

In this section, we present an anonymous authentication scheme on the basis of cloud for medical environment. There are five participants in our scheme: including patients P, healthcare center H, doctors D, cloud C, and sensors S. Healthcare center is trusted medical center. The cloud servers possess the jurisdiction to store patients' medical data which can be accessed by patients and doctors remotely. Sensors can collect and measure the patient's health information timely. In Figure 1, we depict the structure of the cloud based authentication system for TMIS simply.

Our scheme consists of four phases which are described as follows. In order to initialize this protocol, the key generation center (KGC) chooses a multiplication cyclic group G and a generator $g \in G$ with order p, where p is a large prime number. Then KGC selects random numbers $k_x \in Z_p$ ($x = \{H, P, D, C\}$) and computes $PK_x = g^{k_x} \bmod p$. Finally, KGC issues the public key and secret key pairs (PK_x, k_x) to the participants.

We list the used notations of the proposal as follows.

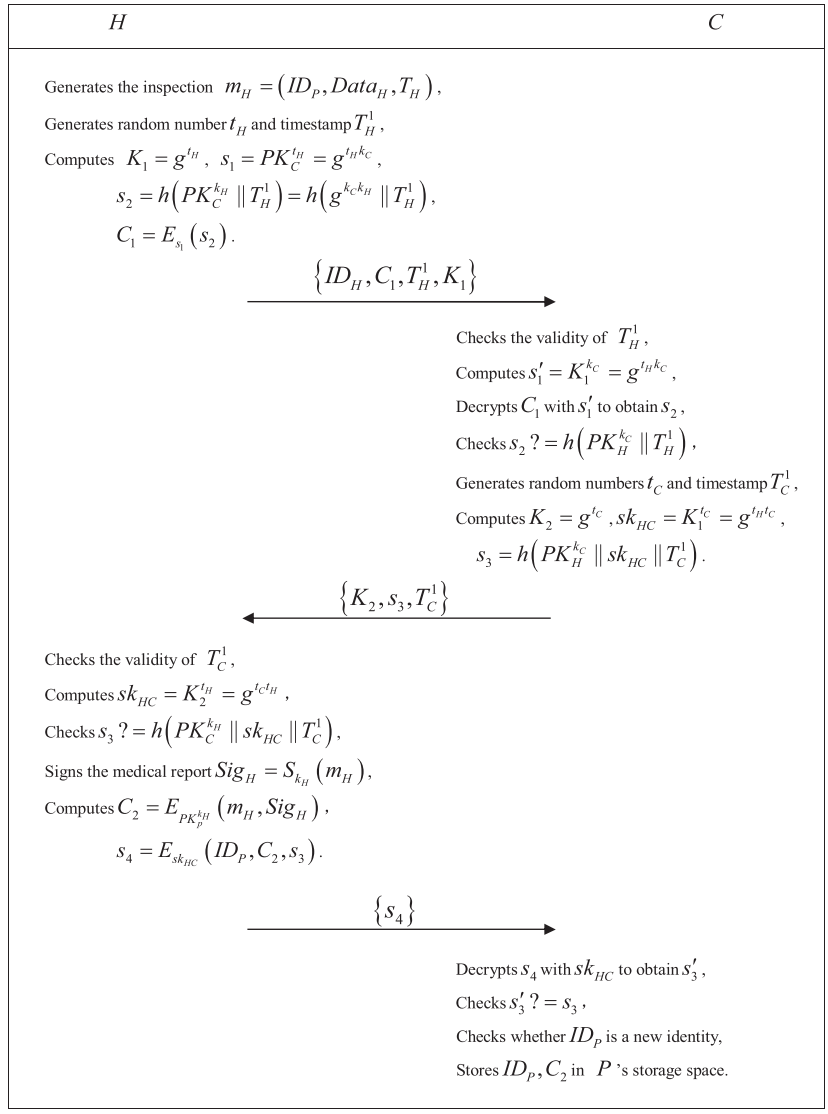


Figure 2: Healthcare center uploading phase.

(i) ID_x: the identity of x

(ii) m_x: the health report of x

(iii) $Data_x$: the health data of x

(iv) sk_{xy}: the session key shared between x and y

(v) PK_x: the public key of x

(vi) k_x: the secret key of x

(vii) Sig_x: signature signed by x

(viii) T_x^i: the ith timestamp generated by x

(ix) G: a multiplicative cyclic group of prime order q

(x) g: the generator of G

(xi) $h(\cdot)$: a one-way hash function

(xii) $\oplus$: exclusive-OR operation

(xiii) $\|$: string concatenation operation

(xiv) $E_{key}(M)$: encrypting M using symmetric encryption with key

(xv) $D_{key}(C)$: decrypting the ciphertext C with key

(xvi) $S_{key}(M)$: signing M with secret key key

(xvii) $V_{key}(Sig_x)$: verifying the signature Sig_x

2.1. Healthcare Center Uploading Phase. As shown in steps 1.1 and 1.2 in Figure 1, patient P makes a health inspection in the healthcare center H and H uploads the generated inspection record to the cloud server C. In Figure 2, we will further describe the authentication process of the phase.

Step 1. After generating the inspection report, H selects a random number t_H and computes $K_1 = g^{t_H}$, $s_1 = PK_C^{t_H} = g^{t_H k_C}$, $s_2 = h(PK_C^{k_H} \| T_H^1) = h(g^{k_C k_H} \| T_H^1)$, $C_1 =$

$E_{s_1}(s_2)$, where T_H^1 is the current timestamp. After that, it sends $\{ID_H, C_1, T_H^1, K_1\}$ to the cloud C.

Step 2. On receiving $\{ID_H, C_1, T_H^1, K_1\}$ from H, C checks the validity of T_H^1. If it is invalid, C aborts the received messages; otherwise, it computes $s_1' = K_1^{k_C} = g^{t_H k_C}$ and decrypts C_1 to recover the values s_2. Subsequently, C calculates $s_2' = h(PK_H^{k_C} \parallel T_H^1)$ and compares it with the decrypted s_2. If the equivalence holds, the legitimacy of H is assured. Then C generates a random number t_C and acquires the timestamp T_C^1 to compute $K_2 = g^{t_C}$, $sk_{HC} = K_1^{t_C} = g^{t_C t_H}$, and $s_3 = h(PK_H^{k_C} \parallel sk_{HC} \parallel T_C^1) = h(g^{k_H k_C} \parallel sk_{HC} \parallel T_C^1)$. Finally, C transmits the authentication message $\{K_2, s_3, T_C^1\}$ to H.

Step 3. Upon receiving the reply message, H checks the validity of T_C^1. If $T_H^2 - T_C^1 \geq \triangle T$, H terminates this phase; otherwise, it computes $sk_{HC} = K_2^{t_H} = g^{t_C t_H}$, $s_3' = h(PK_C^{k_H} \parallel sk_{HC} \parallel T_C^1) = h(g^{k_H k_C} \parallel sk_{HC} \parallel T_C^1)$ and compares s_3' with the received s_3. If they are not equal, the uploading phase is given up by H; else, H signs P's medical report $m_H = (ID_P, Data_H, T_H)$ with its secret key: $Sig_H = S_{k_H}(m_H)$. Note that T_H is the current timestamp when P makes health inspection. After that, H calculates $C_2 = E_{PK_P^{k_H}}(m_H, Sig_H) = E_{g^{k_H k_P}}(m_H, Sig_H)$, $s_4 = E_{sk_{HC}}(ID_P, C_2, s_3)$ and uploads s_4 to the cloud.

Step 4. On receiving s_4, C decrypts it with sk_{HC} to recover ID_P, C_2, s_3' and verifies whether s_3' is equal to s_3 or not. If it is true, the healthcare center is authentic. After that, it verifies whether P is a new user or not. If P is a new user, C stores (ID_P, C_2) in a new storage space; else, it stores C_2 in P's database.

2.2. Patient Uploading Phase. As shown in steps 2.1 and 2.2 in Figure 1, patient P collects health information m_S measured by body sensors S and he could upload the health data to the cloud. In Figure 3, we will depict the detailed process.

Step 1. When P's mobile phone collects the measured information, then it generates the timestamp T_P^1 and a random number a_P to compute $K_3 = g^{a_P}$, $s_5 = PK_C^{a_P} = g^{a_P k_C}$, $s_6 = h(PK_C^{k_P} \parallel T_P^1) = h(g^{k_C k_P} \parallel T_P^1)$, $C_3 = E_{s_5}(ID_P, s_6)$. Subsequently, P transmits $\{C_3, T_P^1, K_3\}$ to C.

Step 2. After receiving the messages, C verifies the freshness of T_P^1 by checking whether $T_C^2 - T_P^1 \leq \triangle T$ or not. If it is valid, C decrypts C_3 to obtain the values ID_P and s_6 with the computed $s_5' = K_3^{k_C} = g^{a_P k_C}$. After that, C calculates $s_6' = h(PK_P^{k_C} \parallel T_P^1)$ and verifies whether the equation $s_6' = s_6$ holds. If it does, P is legitimate user. Then C selects random numbers a_C and computes $K_4 = g^{a_C}$, $sk_{PC} = K_3^{a_C} = g^{a_C a_P}$ and $s_7 = h(PK_P^{k_C} \parallel sk_{PC} \parallel T_C^3) = h(g^{k_P k_C} \parallel sk_{PC} \parallel T_C^3)$, where T_C^3 is the acquired timestamp. Finally, C transmits the $\{K_4, s_7, C_2, T_C^3\}$ to patient P.

Step 3. On receiving response, P checks the validity of T_C^3. If T_C^3 is invalid, P terminates the procedure. If T_C^3 is fresh, P computes $sk_{PC} = K_4^{a_P} = g^{a_C a_P}$, $s_7' = h(PK_C^{k_P} \parallel sk_{PC} \parallel T_C^3) = h(g^{k_P k_C} \parallel sk_{HC} \parallel T_C^3)$ and verifies s_7 is valid by checking whether $s_7' = s_7$ holds. If so, P decrypts C_2 with the computed $PK_H^{k_P}$ to recover $m_H = (ID_P, Data_H, T_H)$ and Sig_H. Subsequently, he/she verifies the validity of H's signature Sig_H. If Sig_H is valid, P chooses a random number sn and computes $C_4 = E_{PK_D^{snk_P}}(m_H, m_S, Sig_H) = E_{g^{snk_D k_P}}(m_H, m_S, Sig_H)$, where $m_S = (ID_P, Data_S, T_S)$ is the collected measured data. Note that T_S is the current timestamp when the body sensors S monitor P's physical condition. Then P calculates $s_8 = E_{sk_{PC}}(ID_P, C_4, s_7)$ and uploads s_8 to the C.

Step 4. On receiving the reply message, C decrypts s_8 with sk_{PC} and obtains ID_P, C_4, s_7'. After that, the cloud server verifies P's validity by checking whether s_7' equals to s_7 or not. If so, C stores C_4 in P's storage space to replace C_2; otherwise, it resumes the procedure.

2.3. Treatment Phase. As shown in steps 3.1 and 3.2 in Figure 1, D is appointed by P and obtains P's identity ID_P and appointment sequence value sn. Subsequently, D can download P's inspection report and measured health information from C, and he/she also can upload the diagnosing records with his/her signature to C. The details of the execution steps are further illustrated in Figure 4.

Step 1. D selects a random number b_D and computes $K_5 = g^{b_D}$, $s_9 = PK_C^{b_D} = g^{b_D k_C}$, $s_{10} = h(PK_C^{k_D} \parallel T_D^1) = h(g^{k_C k_D} \parallel T_D^1)$, $C_5 = E_{s_9}(s_{10})$, where T_D^1 is the acquired current timestamp. Then D transmits $\{ID_D, C_5, T_D^1, K_5\}$ to the cloud C.

Step 2. After receiving messages from D, C checks $T_C^4 - T_D^1 ? \leq \triangle T$. If it is invalid, C terminates the phase; otherwise, it computes $s_9' = K_5^{k_C} = g^{b_D k_C}$ and decrypts C_5 to obtain the values s_{10}. Later, C verifies $s_{10}' = h(PK_D^{k_C} \parallel T_D^1)$ with the decrypted s_{10} to confirm the legitimation of D. If they are equal, C generates random numbers b_C and the timestamp T_C^5 and computes $K_6 = g^{b_C}$, $sk_{CD} = K_5^{b_C} = g^{b_C b_D}$, $s_{11} = h(PK_D^{k_C} \parallel sk_{CD} \parallel T_C^5) = h(g^{k_C k_D} \parallel sk_{CD} \parallel T_C^5)$. Finally, C sends $\{K_6, s_{11}, C_4, T_C^5\}$ to D.

Step 3. On receiving $\{K_6, s_{11}, C_4, T_C^5\}$, D checks the freshness of T_C^5. If so, D computes $sk_{CD} = K_6^{b_D} = g^{b_C b_D}$, $s_{11}' = h(PK_C^{k_D} \parallel sk_{CD} \parallel T_C^5) = h(g^{k_C k_D} \parallel sk_{CD} \parallel T_C^5)$ and compares s_{11}' with the received s_{11} to assure C's authenticity. If they are not equal, the phase is terminated by himself/herself; otherwise, D uses the appointment sequence number sn to compute $PK_P^{snk_D}$ and decrypts C_4 with it to recover m_H, m_S, Sig_H. Subsequently, D verifies whether the signature Sig_H is valid or not by checking $m_H ? = V_{PK_H}(Sig_H)$. If it is valid, D diagnoses P's symptom on the basis of (m_H, m_S) and generates the diagnostic records $m_D = (ID_P, Data_D, T_D)$, where T_D is the timestamp when the doctor generates m_D. After that, D uses his/her private

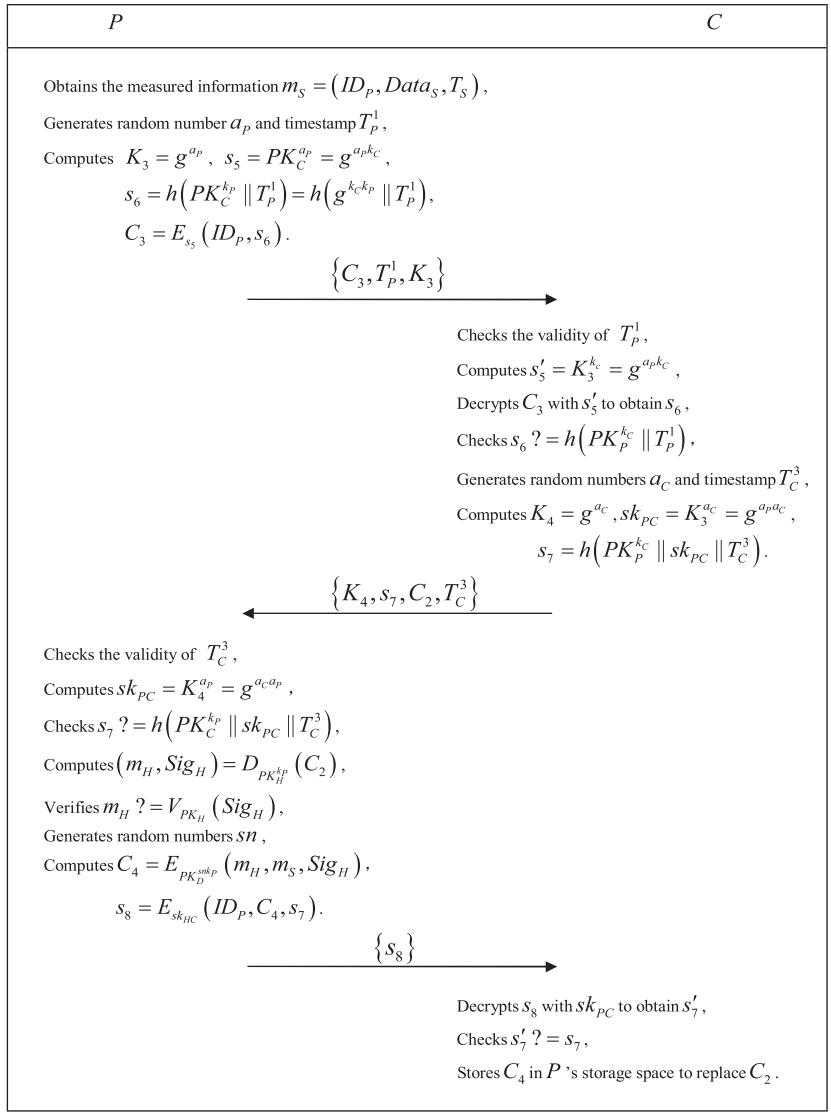


Figure 3: Patient uploading phase.

key k_D to sign m_D: $Sig_D = S_{k_D}(m_D)$. Then, D calculates $C_6 = E_{PK_P^{snk_D}}(m_H, m_S, m_D, Sig_D)$, $s_{12} = E_{sk_{CD}}(ID_P, C_6, s_{11})$ and sends s_{12} to the cloud.

Step 4. Upon receiving s_{12}, C decrypts it with sk_{CD} and obtains ID_P, C_6, s_{11}'. Later, it checks $s_{11}' ? = s_{11}$. If the equation holds, the validity of D is confirmed; otherwise, this phase fails. After that, C replaces C_4 with C_6 and stores it in P's storage space.

2.4. Checking Report Phase. As shown in step 4.1 in Figure 1, patient P can access the cloud to obtain the medical record via the mobile phone. In Figure 5, we depict the detailed process of the phase.

Step 1. P generates the timestamp T_P^2 and a random number v_P to compute $K_7 = g^{v_P}$, $s_{13} = PK_C^{v_P} = g^{v_P k_C}$, $s_{14} = h(PK_C^{k_P} \| T_P^2) = h(g^{k_C k_P} \| T_P^2)$, $C_7 = E_{s_{13}}(ID_P, s_{14})$. Then, P transmits the request $\{C_7, T_P^2, K_7\}$ to the cloud C.

Step 2. C verifies the freshness of T_P^2 after receiving the request from P. If so, C decrypts C_7 with computed $s_{13}' = K_7^{k_C} = g^{v_P k_C}$ and obtains the ID_P, s_{14}'. Subsequently, C calculates $s_{14}' = h(PK_P^{k_C} \| T_P^2)$ and verifies $s_{14}' ? = s_{14}$. If they are equal, C computes $K_8 = g^{v_C}$, $sk_{PC'} = K_7^{v_C} = g^{v_C v_P}$ and $s_{15} = h(PK_P^{k_C} \| sk_{PC'} \| T_C^6) = h(g^{k_P k_C} \| sk_{PC'} \| T_C^6)$, where v_C and T_C^6 are generated random value and the acquired timestamp, respectively. After that, C sends the $\{K_8, s_{15}, C_6, T_C^6\}$ to P.

Step 3. On receiving response $\{K_8, s_{15}, C_6, T_C^6\}$, P checks the validity of T_C^6. If $T_P^3 - T_C^6 \geq \triangle T$, P aborts the session; otherwise, he/she calculates $sk_{PC'} = K_8^{v_P} = g^{v_C v_P}$, $s_{15}' = h(PK_C^{k_P} \|$

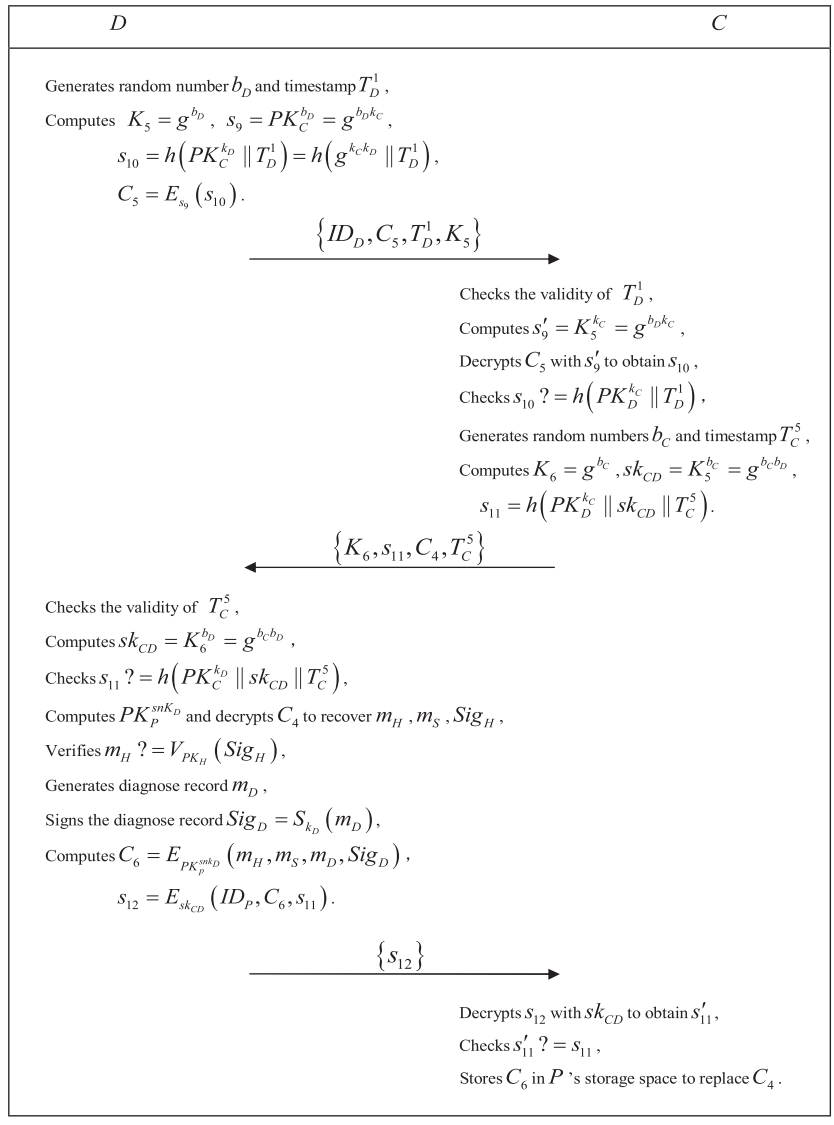


Figure 4: Treatment phase.

$sk_{PC'} \| T_C^6) = h(g^{k_P k_C} \| sk_{PC'} \| T_C^6)$ and verifies whether $s_{15}' = s_{15}$ holds. If yes, P computes $(m_H, m_S, m_D, Sig_D) = D_{PK_D^{snkp}}(C_6)$ and he/she verifies whether $m_D ? = V_{PK_D}(Sig_D)$. If so, P takes the diagnosis report of doctor D for the health information m_H and m_S. Finally, P calculates $C_8 = E_{PK_P^{cn}}(m_H, m_S, m_D)$ and $s_{16} = E_{sk_{PC'}}(ID_P, C_8, s_{15})$ and uploads s_{16} to the C, where cn is a pre-generated key to encrypt (m_H, m_S, m_D). Note that cn is the pre-generated key for all the medical staff. If the patient P has an accident, they can decrypt C_8 to obtain the health information and make treatment decisions.

Step 4. Upon receiving the reply s_{16}, C computes $(ID_P, C_8, s_{15}') = D_{sk_{PC'}}(s_{16})$ and verifies whether s_{15}' equals s_{15} or not. If $s_{15}' = s_{15}$, C stores C_8 in P's storage space to replace C_6; otherwise, the phase is failed.

3. Security Proof

In this section, we will prove our scheme to be secure in standard model. We reduce the security of our authentication scheme to cryptography basic elements [21, 22]. At first, in order to achieve this goal, we will introduce the definitions of security, a structured security model, and the basic assumptions. Then we use all of them to prove the result.

Definition 1 (semantic security). For arbitrary security parameter N, if and only if any polynomial time adversary has a negligible advantage against the scheme, we say the scheme has semantic security.

The definitions are inherited and modified from the methodology of Bellare, Pointcheval, and Rogaway [23] and the game-based structure [24] is used to prove this scheme achieving semantic security.

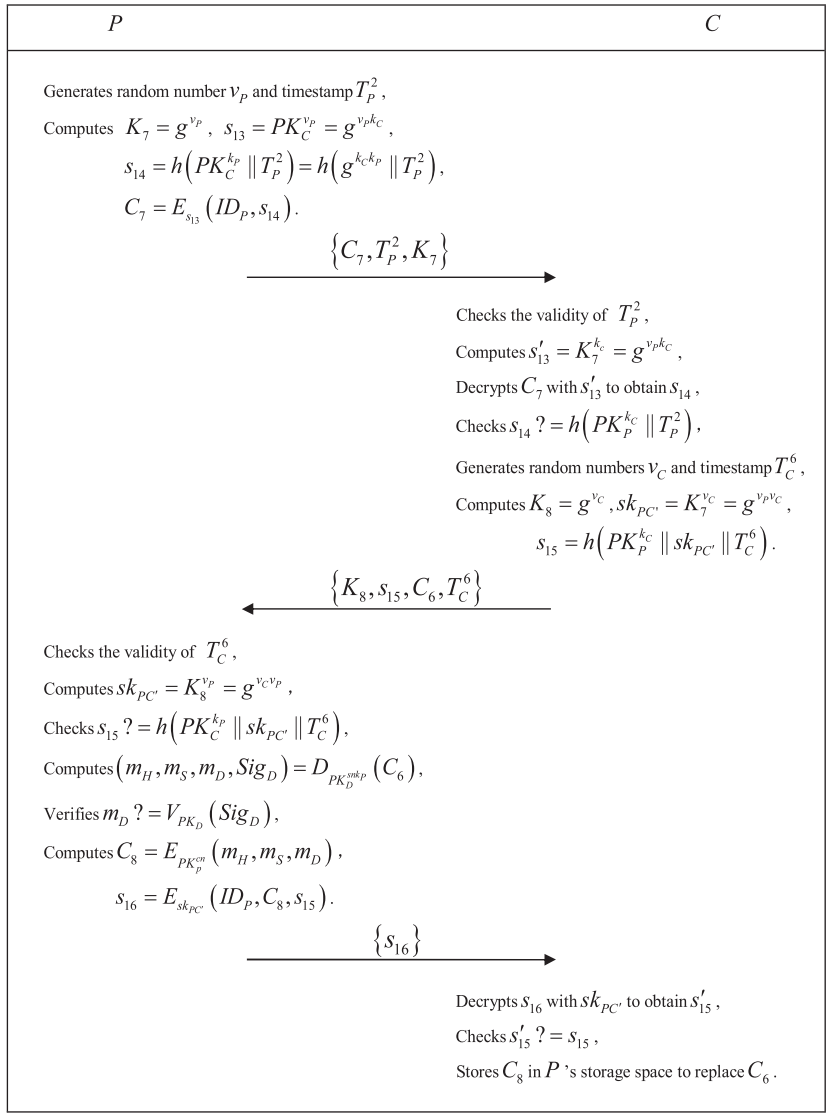


Figure 5: Checking report phase.

Security Model. In the security model, the adversary plays a game with an oracle. The oracle runs the real protocol and answers the queries of adversary to simulate the real interaction of participants. After a range of queries, the adversary gets different capabilities. When the adversary finishes the training and obtains enough messages, oracle should answer the test query once. Finally, we judge if the adversary wins or loses by what the adversary gets. The adversary and the oracle are denoted by $\mathcal{A}$ and $\mathcal{B}$, respectively.

Init: before replying to queries of $\mathcal{A}$, $\mathcal{B}$ generates the system parameters including security parameter N, a multiplication cyclic group G, and a generator $g \in G$ with order p, where p is a large prime number related N. Then $\mathcal{B}$ selects random numbers $k_x \in Z_p$ and computes $PK_x = g^{k_x} \bmod p$ for $x \in \{H, P, D, C\}$. We notice that in a complete system P and D are not unique. Then $\mathcal{B}$ prepares public key and secret key pairs denoted by (PK_{P_i}, k_{P_i}) and (PK_{D_j}, k_{D_j}) for P_i and D_j, where $i \in [1, n]$ and $j \in [1, m]$. $\mathcal{B}$ marks up all P and D with void state. Then it maintains a list of SID recording simulated conversations. $SID^u_{P_i,D_j}$ represents the uth conversation involved P_i and D_j. Noticeably, any SID has a void state before being invoked.

After the init phase, $\mathcal{A}$ is allowed to make queries for simulating the real protocol.

Corrupt(P_i): $\mathcal{B}$ gives k_{P_i} back to $\mathcal{A}$ and marks up the state of P_i with corrupted.

Corrupt(D_j): $\mathcal{B}$ gives k_{D_j} back to $\mathcal{A}$ and marks up the state of D_j with corrupted.

Reveal($SID^u_{P_i,D_j}$): this query simulates abuse of session keys sk.

(1) If the state of $SID^u_{P_i,D_j}$ is void, $\mathscr{B}$ executes a real conversation and gives sk (being used to encrypt m) to $\mathscr{A}$. Then the result of $SID^u_{P_i,D_j}$ is changed into revealed.
(2) Else, if the state of $SID^u_{P_i,D_j}$ is not void, $\mathscr{B}$ answers $\mathscr{A}$ according to the conversation in the list and adds revealed to its state.

Execute($SID^u_{P_i,D_j}$): this query simulates the passive adversary.

(1) When one or both of P_i and D_i are corrupted, $\mathscr{B}$ checks the state of $SID^u_{P_i,D_j}$.
   (a) If void, $\mathscr{B}$ executes a real conversation, gives to $\mathscr{A}$ data transferred over the network, and changes the state of $SID^u_{P_i,D_j}$ into executed.
   (b) Else, if not void, $\mathscr{B}$ answers $\mathscr{A}$ using the conversation in the list and adds executed into its state.
(2) When neither of P_i and D_i is corrupted, $\mathscr{B}$ checks the state of $SID^u_{P_i,D_j}$.
   (a) If void, $\mathscr{B}$ executes a real conversation, gives to $\mathscr{A}$ data transferred over the network, and changes the state of $SID^u_{P_i,D_j}$ into executed.
   (b) Else, if not void, $\mathscr{B}$ answers $\mathscr{A}$ using the conversation in the list and adds executed into its state.

Send(P_i, D_i, Message): this query simulates active attacks.

(1) When one or both of P_i and D_i are corrupted, $\mathscr{B}$ executes a real conversation to answer $\mathscr{A}$.
   (a) If message is verified successfully, $\mathscr{B}$ answers $\mathscr{A}$ as the real protocol and adds this conversation into list with sent state.
   (b) Else, $\mathscr{B}$ outputs that $\mathscr{A}$ is refused.
(2) When neither of P_i and D_i is corrupted, $\mathscr{B}$ answers $\mathscr{A}$ as follows.
   (a) If message is verified successfully, $\mathscr{B}$ answers $\mathscr{A}$ as the real protocol and adds this conversation into list with sent state.
   (b) Else, $\mathscr{B}$ outputs that $\mathscr{A}$ is refused.

Test(m_0, m_1): $\mathscr{B}$ chooses P_i and D_j whose both states are void. Then it randomly chooses a coin b and simulates a real conversation with inputting m_b. $\mathscr{A}$ guesses the result of b.

The complexity assumptions needed for proving security of our scheme are reviewed as follows:

Assumption 2 (DDH). Let G be a cyclic group with order p. $g \in G$ is a generator and $a, b \in Z_p^*$. It is difficult to judge if an element in G equals g^{ab} by just given g^a and g^b.

We denote upper bound of adversaries' advantage against DDH as ϵ_{DDH}. So ϵ_{DDH} should be negligible if the assumption is right.

Assumption 3 (hash). There exists a secure irreversible hash function which achieves strong collision resistance.

We denote the advantage of adversaries against the hash function as ϵ_{hash}. So ϵ_{hash} should be negligible if the assumption is right.

Assumption 4 (signature). There exists a secure digital signature scheme.

We denote advantage of adversaries against this signature scheme as ϵ_{sign}. So ϵ_{sign} should be negligible if the assumption is right.

Assumption 5 (encryption). There is a symmetry encryption that achieves CPA security.

We denote advantage of adversaries against this encryption scheme as ϵ_{enc}. So ϵ_{enc} should be negligible if the assumption is right.

Proof. A PPT adversary $\mathscr{A}$ is attacking the protocol. We use a series of games to bound the advantage of $\mathscr{A}$. The advantage of $\mathscr{A}$ in Game i is defined as

$$Adv_i \overset{def}{=\!=} |2 \cdot Pr[\mathscr{A} \text{ succeeds in Game } i] - 1|. \quad (1)$$

The games used to bound the advantages of $\mathscr{A}$ are listed in the following. We analyse the advantage difference in nearly games and bound them. In Game 0, it would be the real protocol.

Game 0. $\mathscr{A}$ interacts with the initial security model.

Game 1. In this game, we modify Execute queries. When the states of P_i, D_j, and SID are all void, $\mathscr{B}$ simulates a real protocol but replaces $s_1, sk_{HC}, s_5, sk_{PC}, s_9, sk_{DC}, s_{13}$ and $sk_{PC'}$ with random numbers in G.

Lemma 1. $|Adv_0 - Adv_1| \leq negl$.

Proof. We just replace the sks of traditional DH protocol with random numbers. The advantage difference between two games is caused by DDH problem. And hence, Lemma 1 is proved by DDH assumption right.

Game 2. This game is based on Game 1 and we also modify Execute queries. When the states of P_i, D_j, and SID are all void, $\mathscr{B}$ simulates a real protocol but replaces $s_2, s_3, s_6, s_7, s_{10}, s_{11}, s_{14}, s_{15}$ with uniform random numbers in the range of hash function.

Lemma 2. $|Adv_1 - Adv_2| \leq negl$.

Proof. We just replace the real hash results with random numbers. Without the knowledge of inputs, the probability that $\mathscr{A}$ can distinguish the real hash results and random numbers is less than the advantage of $\mathscr{A}$ that captures the hash. And hence, if the hash function is secure, the probability is negligible.

Table 1: Comparisons of properties.

	Chen et al.'s [15]	Chiou et al.'s [16]	Cheng et al.'s [20]	Ours
Resistance of impersonation attack	Yes	Yes	Yes	Yes
Prevention of replay attack	Yes	Yes	Yes	Yes
Preserving patient privacy	No	Yes	No	Yes
Complete mutual authentication	No	No	No	Yes
Perfect forward secrecy	Yes	Yes	Yes	Yes
Confidentiality	Yes	Yes	No	Yes
Perfect backward secrecy	Yes	Yes	Yes	Yes

Game 3. This game is based on Game 2 and we modify Execute queries. When the states of P_i, D_j, and *SID* are all void, $\mathscr{B}$ simulates a real protocol but replaces $C_1, C_2, s_4, C_3, C_4, s_8, C_5, C_6, s_{12}, C_7, C_8, s_{16}$ with uniform random numbers in the range of encryption.

Lemma 3. $|Adv_2 - Adv_3| \leq negl$.

Proof. We just replace the symmetric encryption results with random numbers. Without knowing inputs, the probability that $\mathscr{A}$ can distinguish the real encryption results and random numbers is less than the advantage of $\mathscr{A}$ that captures the CPA security symmetric encryption. Thereby, if the symmetric encryption is secure, the probability is negligible.

Game 4. This game is based on Game 3 and we modify Send queries. When the states of P_i and D_j are both void, $\mathscr{B}$ simulates a real protocol but replaces $s_1, sk_{HC}, s_5, sk_{PC}, s_9, sk_{DC}, s_{13}$ and $sk_{PC'}$ with random numbers in G.

Lemma 4. $|Adv_3 - Adv_4| \leq negl$.

Proof. We just replace the *sk*s of traditional DH protocol with random numbers. The advantage difference between two games is caused by DDH problem. Therefore, Lemma 4 is proved by DDH assumption right.

Game 5. This game is based on Game 4 and we modify Send queries. When the states of P_i and D_j are both void, $\mathscr{B}$ simulates a real protocol but replaces $s_2, s_3, s_6, s_7, s_{10}, s_{11}, s_{14}, s_{15}$ with uniform random numbers in the range of hash function.

Lemma 5. $|Adv_4 - Adv_5| \leq negl$.

Proof. We just replace the real hash results with random numbers. Without knowing inputs, the probability that $\mathscr{A}$ can distinguish the real hash results and random numbers is less than the advantage of $\mathscr{A}$ that captures the hash. If the hash function is secure, the probability is negligible.

Game 6. This game is based on Game 5 and we also modify Send queries. When the states of P_i and D_j are both void, $\mathscr{B}$ simulates a real protocol but replaces $C_1, C_2, s_4, C_3C_4, s_8, C_5, C_6, s_{12}, C_7, C_8, s_{16}$ with uniform random numbers in the range of encryption.

Lemma 6. $|Adv_5 - Adv_6| \leq negl$.

Proof. We just replace the encryption results by random numbers. Without the knowledge of inputs, the probability that $\mathscr{A}$ can distinguish the real encryption results and random numbers is less than the advantage of $\mathscr{A}$ that captures the CPA security symmetric encryption. If the symmetric encryption is secure, the probability is negligible.

In Game 6, we notice that P_i and D_j of test query neither are corrupted. For any *SID* without being corrupted, the Execute and Send queries are all randomized. So in Game 6, the advantage of $\mathscr{A}$ is zero. So, we can compute the Adv_0 as follows:

$$Adv_0 = |Adv_0 - Adv_6| \leq \epsilon_{DDH} + \epsilon_{hash} + \epsilon_{enc}, \quad (2)$$

which is a negligible value. □

4. Performance and Functionality Analysis

Herein, we evaluate the performance and functionality of the proposed scheme and compare it with three related schemes for cloud based telemedicine systems, including Chen et al.'s scheme [15], Chiou et al.'s scheme [16], and Cheng et al.'s scheme [20].

The comparisons on the key security properties among these systems are given in Table 1. It is visible that our scheme could achieve all security properties and it is superior to the rest three related schemes. Chen et al.'s scheme [15] fails to provide anonymity and complete mutual authentication, while Chiou et al.'s scheme [16] could not achieve the complete mutual authentication. Furthermore, Cheng et al.'s scheme [20] could not preserve users' privacy, complete mutual authentication, and confidentiality. Note that the proposed scheme offers important security features and it is better suitable for cloud based telemedicine environment.

Meanwhile, we present the comparisons of efficiency in terms of computation loading among these schemes in Table 2. Compared with the other three related schemes, the proposed scheme needs not perform the bilinear pairing and could provide more additional security features. Furthermore, our scheme achieves the provably security in the standard model.

More detailed efficiency comparisons are shown in Figures 6 and 7. We implement the cloud of authentication schemes for cloud based telemedicine systems in Python 3.5.2 using an Intel(R) Core(TM) i5-4590 CPU @ 3.30GHZ with 3300MB RAM and Ubuntu 16.04 system. The simulations of platform for healthcare center, patients, and doctors are

Table 2: Comparisons of computation loading.

		Chen et al.'s [15]	Chiou et al.'s [16]	Cheng et al.'s [20]	Ours
HUP	H	$T_s+T_a+2T_m+2T_p+T_E+3T_h$	$T_s+2T_p+2T_E+4T_h$	$2T_p+2T_E+4T_h+T_s+T_m$	$5T_e + 2T_h + T_s + 3T_E$
	C	$2T_m + 2T_p + T_E + 3T_h$	$T_p + T_E + 3T_h$	$3T_h + T_p + T_E$	$4T_e + 2T_h + 2T_E$
PUP	P	$T_s+T_a+2T_m+2T_p+2T_E+3T_h$	$T_s+3T_p+2T_E+7T_h$	$3T_p+2T_E+6T_h+T_s+T_m$	$T_s + 6T_e + 4T_E + 2T_h$
	C	$2T_m + 2T_p + T_E + 3T_h$	$T_p + 4T_h$	$T_p + 5T_h$	$4T_e + 2T_E + 2T_h$
TP	D	$2T_s+2T_m+2T_p+3T_E+3T_h$	$2T_s+2T_p+2T_E+4T_h$	$2T_s + 2T_E + 2T_p + 5T_h$	$5T_e+4T_E+2T_h+2T_s$
	C	$2T_m+2T_p+T_E+3T_h$	$T_p + 4T_h$	$T_p + 4T_h$	$4T_e+2T_E+2T_h$
CP	P	N/A	$T_s+T_p+2T_E+4T_h$	$T_s+T_p+2T_E+4T_h$	$6T_e+4T_E+2T_h+T_s$
	C	N/A	T_p+4T_h	T_p+4T_h	$4T_e+2T_E+2T_h$

HUP: healthcare center uploading phase.
PUP: patient uploading phase.
TP: treatment phase.
CP: checking report phase.
T_h: time consumption for executing a hash function.
T_E: time consumption for executing the symmetric encryption/decryption operation.
T_e: time consumption for executing a modular exponent operation.
T_p: time consumption for executing a bilinear pairing operation.
T_s: time consumption for executing/verifying a signature.
T_a: time consumption for executing an asymmetric encryption/decryption operation.
T_m: time consumption for executing a multiplication operation.

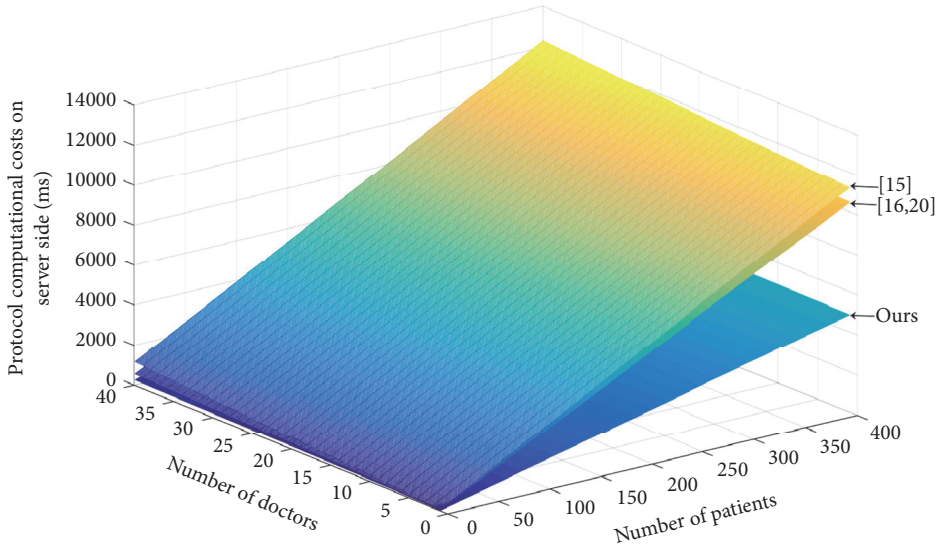


Figure 6: Cost time on cloud computing.

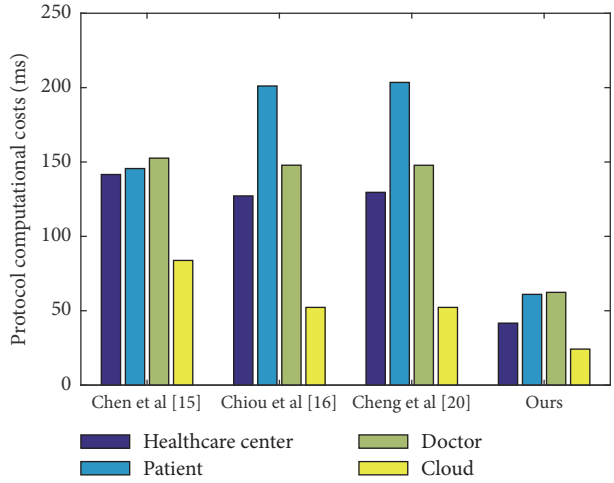


Figure 7: Cost time on healthcare center, patients, doctors, and cloud.

implemented in Python 3.5.2 using an Intel(R) Core(TM) i5-4590 CPU at 1.65GHZ with 1540MB RAM and Ubuntu 16.04 system. The one-way hash function used is SHA-256, and the symmetric encryption/decryption algorithm is advanced encryption standard. We use the ElGamal signature scheme and ElGamal encryption scheme with 1024-bit security parameter for digital signature algorithm and the asymmetric encryption/decryption algorithm, respectively. Moreover, the bilinear paring is simulated in two MNT asymmetric groups, "MNT224".

Figure 6 shows the main cost on the cloud computing of interacting with multiple patients and doctors for authentication simultaneously. It demonstrates that our proposal costs less time for the cloud to authenticate doctors and patients. Figure 7 illustrates the main cost on healthcare center, patients, doctors, and cloud for one round authentication in healthcare center uploading phase, patient uploading phase, treatment phase, and checking report phase, respectively. From Figure 7 we can conclude that our scheme is the most efficient to finish one round mutual authentication.

5. Conclusion

In this article, we proposed an anonymous authentication scheme based on cloud for medical environment, which provided both data confidentiality and message authenticity. Subsequently, we stated that the proposed scheme was provably secure in the standard model. The comparisons with existing competitive protocols also observe that our scheme is suitable for the cloud based telecare medical information systems.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by NSFC (Grant Nos. 61602045, 61502044).

References

[1] W. Shi and P. Gong, "A new user authentication protocol for wireless sensor networks using elliptic curves cryptography," *International Journal of Distributed Sensor Networks*, vol. 9, no. 4, Article ID 730831, pp. 51–59, 2013.

[2] Q. Jiang, J. Ma, G. Li, and L. Yang, "An efficient ticket based authentication protocol with unlinkability for wireless access networks," *Wireless Personal Communications*, vol. 77, no. 2, pp. 1489–1506, 2014.

[3] D. Guo and F. Wen, "Analysis and improvement of a robust smart card based-authentication scheme for multi-server architecture," *Wireless Personal Communications*, vol. 78, no. 1, pp. 475–490, 2014.

[4] W. Shi and D. He, "A security enhanced mutual authentication scheme based on nonce and smart cards," *Journal of the Chinese Institute of Engineers*, vol. 37, no. 8, pp. 1090–1095, 2014.

[5] D. He, N. Kumar, M. K. Khan, and J.-H. Lee, "Anonymous two-factor authentication for consumer roaming service in global mobility networks," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 4, pp. 811–817, 2013.

[6] D. Wang, H. Cheng, D. He, and P. Wang, "On the challenges in designing identity-based privacy-preserving authentication schemes for mobile devices," *IEEE Systems Journal*, vol. 12, no. 1, pp. 916–925, 2018.

[7] D. He, N. Kumar, and N. Chilamkurti, "A secure temporal-credential-based mutual authentication and key agreement scheme with pseudo identity for wireless sensor networks," *Information Sciences*, vol. 321, Article ID 11403, pp. 263–277, 2015.

[8] Q. Jiang, F. Wei, S. Fu, J. Ma, G. Li, and A. Alelaiwi, "Robust extended chaotic maps-based three-factor authentication scheme preserving biometric template privacy," *Nonlinear Dynamics*, vol. 83, no. 4, pp. 2085–2101, 2016.

[9] D. He, S. Zeadally, and L. Wu, "Certificateless public auditing scheme for cloud-assisted wireless body area networks," *IEEE Systems Journal*, vol. 12, no. 1, pp. 64–73, 2018.

[10] Q. Jiang, M. K. Khan, X. Lu, J. Ma, and D. He, "A privacy preserving three-factor authentication protocol for e-Health clouds," *The Journal of Supercomputing*, vol. 72, no. 10, pp. 3826–3849, 2016.

[11] D. He, N. Kumar, H. Wang, L. Wang, R. K. Choo, and A. Vinel, "A provably-secure cross-domain handshake scheme with symptoms-matching for mobile healthcare social network," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 633–645, 2016.

[12] R. P. Padhy, M. R. Patra, and S. C. Satapathy, "Design and implementation of a cloud based rural healthcare information system model," *American Journal of Computer Science and Technology*, vol. 2, no. 1, pp. 149–157, 2012.

[13] A. Banerjee, P. Agrawal, and R. Rajkumar, "Design of a cloud based emergency healthcare service model," *International Journal of Applied Engineering Research*, vol. 8, no. 19, pp. 2261–2264, 2013.

[14] C.-L. Chen, T.-T. Yang, and T.-F. Shih, "A secure medical data exchange protocol based on cloud environment," *Journal of Medical Systems*, vol. 38, no. 9, article 112, 2014.

[15] C.-L. Chen, T.-T. Yang, M.-L. Chiang, and T.-F. Shih, "A privacy authentication scheme based on cloud for medical environment," *Journal of Medical Systems*, vol. 38, article 143, 2014.

[16] S. Y. Chiou, Z. Ying, and J. Liu, "Improvement of a privacy authentication scheme based on cloud for medical environment," *Journal of Medical Systems*, vol. 40, no. 4, pp. 1–15, 2016.

[17] P. Mohit, R. Amin, A. Karati, G. P. Biswas, and M. K. Khan, "A stand mutual authentication protocol for cloud computing based health care system," *Journal of Medical Systems*, vol. 41, no. 4, pp. 1–13, 2017.

[18] C. T. Li, D. H. Shih, and C. C. Wang, "On the security of a privacy authentication scheme based on cloud for medical environment," in *International Conference on Information Science and Applications*, vol. 424 of *Lecture Notes in Electrical Engineering*, pp. 241–248, Springer, Singapore, Singapore, 2017.

[19] C.-T. Li, D.-H. Shih, and C.-C. Wang, "Cloud-assisted mutual authentication and privacy preservation protocol for telecare medical information systems," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 191–203, 2018.

[20] X. F. Cheng, X. L. Zhang, and J. F. Ma, "ICASME: an improved cloud-based authentication scheme for medical environment," *Journal of Medical Systems*, vol. 41, no. 3, pp. 1–14, 2017.

[21] 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, 2016.

[22] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, "Zipfs law in passwords," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

[23] M. Bellare, D. Pointcheval, and P. Rogaway, "Authenticated key exchange secure against dictionary attacks," in *Advances in Cryptology—Eurocrypt 2000*, vol. 1807 of *Lecture Notes in Computer Science*, pp. 139–155, Springer, Berlin, Germany, 2000.

[24] A. Groce and J. Katz, "A new framework for efficient password-based authenticated key exchange," in *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS '10)*, pp. 516–525, ACM, Chicago, Ill, USA, October 2010.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 5973285, 11 pages
https://doi.org/10.1155/2018/5973285

Research Article

Oblivious Transfer via Lossy Encryption from Lattice-Based Cryptography

Zengpeng Li,[1] **Can Xiang**,[2] **and Chengyu Wang**[3]

[1]*College of Computer Science and Technology, Qingdao University, Qingdao 266071, China*
[2]*College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China*
[3]*School of CyberSpace Security, Beijing University of Posts and Telecommunications, Beijing 100871, China*

Correspondence should be addressed to Can Xiang; xiangcan1987@sina.com

Received 9 April 2018; Accepted 10 July 2018; Published 2 September 2018

Academic Editor: Jian Shen

Authentication is the first defence line to prevent malicious entities to access *smart mobile devices* (or SMD). Essentially, there exist many available cryptographic primitives to design authentication protocols. *Oblivious transfer* (OT) protocol is one of the important cryptographic primitives to design authentication protocols. The first lattice-based OT framework under universal composability (UC) model was designed by dual mode encryption and promoted us to find an alternative efficient scheme. We note that "lossy encryption" scheme is an extension of the dual mode encryption and can be used to design UC-secure OT protocol, but the investigations of OT via lossy encryption over the lattice are absent. Hence, in order to obtain an efficient authentication protocol by improving the performance of the UC-secure OT protocol, in this paper, we first design a multibit lossy encryption under the decisional learning with errors (LWE) assumption and then design a new variant of UC-secure OT protocol for authenticated protocol via lossy encryption scheme. Additionally, our OT protocol is secure against semihonest (static) adversaries in the common reference string (CRS) model and within the UC framework.

1. Introduction

Oblivious transfer (OT) is an important cryptographic primitive which can be used for designing secure multiparty computing and privacy-preserving schemes, such as authenticated key exchange and password-based authentication key exchange (PAKE) [1]. Apparently, authentication is the first line of defence to prevent unauthorized access from illegitimate entities (including both devices and users). Very recently, the issues of privacy-preserving for *smart mobile devices* (SMD), *Internet of things* (IoTs), *wireless sensor networks* (WSNs), and *cloud storage auditing* are arousing the public attention [2–4]. In this case, it is becoming more important to protect the private information for mobile computing environments by utilizing the technique of authentication [5, 6]. However, the literature which focuses on the problem of how to use the basic cryptographic primitive to design authentication protocols is relatively few and the research of improving the performance of the primitives is fewer. Thus, to design the efficient authentication protocols in the future, this paper focuses on how to improve the performance of universal composable- (UC-) secure OT protocol via lattice-based cryptography.

To our knowledge, OT protocol was originally proposed by Rabin [7]. Since then various cryptography schemes and protocols are designed by using OT, e.g., [1, 8]. Informally speaking, there exist two players (the sender **S** and the receiver **R**) in OT protocol. On the one hand, the sender can send two (or more) values to the receiver. However, the sender does not know which value will be received by the receiver, and the receiver only knows the received value and remains oblivious to the other values. In a word, they are oblivious to other's true behaviour.

Importantly, Gertner et al. [9] pointed out the relationship between the public key encryption (PKE) scheme and the OT protocol at FOCS'00. In some indistinguishability against chosen plaintext attack (IND-CPA) secure PKE schemes, if the public key generated by the key generation algorithm is

indistinguishable from the public key sampled from a uniform distribution, then we can use the PKE scheme to design an OT protocol [9]. In this setting, Peikert-Vaikuntanathan-Waters (PVW) [10] has constructed an *efficient, universally composability and generally realizable* OT via "dual mode encryption" under *worst-case* lattice assumption (LWE) at CRYPTO'08. Loosely speaking, there are two types of public key in "dual mode encryption". One type is injective keys; it is real public key and behaves normally. The other one is "lossy" or "messy" key, it is lossy public key, and it loses some information of the plaintext. Moreover, there exist two important properties for the dual mode encryption. The first one is statistically close; namely, the distributions of ciphertext for any two plaintexts under a lossy key are statistically close. The second one computationally indistinguishable; namely, the injective key is computationally indistinguishable from the lossy key. (Importantly, no efficient adversary can tell the difference between normal keys and lossy keys.)

Along with this line, the notation of "*lossy encryption*" was proposed by Bellare, Hofheinz, and Yilek (BHY) [11] on EUROCRYPT'09. Actually, the lossy encryption is an extension of the meaningful/meaningless encryption [12] and dual mode encryption [10]. In a nutshell, a "lossy" (or "messy" in [10]) cryptosystem is one which also has two modes according to two types of public keys. Concretely, **(1)** In the normal mode, the ciphertext is generated by encrypting the plaintext under an injective key. **(2)** In the lossy (or "messy") mode, the ciphertext is independent of the plaintext. Actually, the operability property was proposed by [11]; they basically can open a ciphertext generated under a lossy key for any plaintext by adopting a possibly inefficient algorithm. Meanwhile, the injective key is computationally indistinguishable from the lossy key. Actually, our work is along this line and we embark on this question:

> *How to design a (string)* OT *protocol via multibit lattice-based "lossy encryption" rather than "dual mode encryption"?*

To solve this issue, we note that, after the polynomial time solvers in the nonclassical quantum computation model was pointed out by Shor [13] for discrete logarithm and integer factorization, most researchers seek to find the various alternative computational assumption; thus lattice-based (e.g., learning with errors, LWE) cryptography draws attention. Over the last decade, lattice has emerged as a very attractive foundation separately for cryptography. Specially, Regev scheme [14] and Gentry-Peikert-Vaikutanathan (GPV, a.k.a., dual Regev) scheme [15] are important lattice-based schemes to remain secure even against quantum computer attacks.

From the above observations and inspired by the work of Peikert et al. [10], we still work along this line and construct a multibit LWE-based lossy encryption scheme which has two types of public keys.

1.1. Our Contributions and Techniques Overview. Although many would consider OT protocol a breakthrough for multiparty computation, nowadays, OT protocols are plagued by several well-known pain-points among which performance (string OT) and security (postquantum attacks) are perhaps the most visible and most often debated points. However, most existing OT protocols adopt a variety of standard number-theoretic assumptions; only a few works focus on designing the protocol under *worst-case* lattice assumption such as [10, 16, 17]. Here we fill in some of the missing details in the high-level description.

(1) We use the lossy encryption scheme to replace the dual mode encryption, then we design the OT protocol via lossy encryption over the lattice. More concretely, we note that, Peikert et al. [10] proposed the framework of OT protocol by using dual mode cryptography. Actually, the lossy encryption is an extension of the meaningful/meaningless encryption [12] and dual mode encryption [10] and has the obvious property of two types of public keys. The crux of this issue is how to obtain multibit lossy encryption scheme. In this paper, we construct *lossy encryption* via multibit Gentry-Peikert-Vaikuntanathan (GPV, a.k.a., dual Regev scheme), i.e., MGPV scheme. In particular, the public key in MGPV with many LWE instances rather than a simple matrix of LWE instance.

(2) Moreover, we design a multibit public key encryption scheme (i.e., MGPV scheme) by following the methodology of Li et al. [18]. Actually, the semantically secure multibit public key encryption scheme via subset sum problem (SSP) proposed by Lyubashevsky et al. [16] and multiple secrets Gentry-Sahai-Waters scheme via LWE assumption proposed by Li et al. [19] promoted us to explore the functions of the public key with a sequence of LWE instances. Importantly, the public matrix $\mathbf{A}$ contains many LWE instances, each one is used to protect the secret key. In this setting, the decrypter can decrypt the plaintext in a bit-by-bit manner.

(3) Lastly, we attempt to explore the potential application of our UC-secure OT protocol for PAKE in SMDs. To our knowledge, PAKE is an important tool to design authentication protocol, which can help SMDs to enable adequate user authentication and prevent unauthorized use of an unattended and lost, etc. Inspired by the OT-based PAKE [1], extending our multibit OT protocol via lossy encryption to multibit PAKE is a natural result. This solution is aimed at helping us apply SMDs for authentication and other security services.

1.2. Paper Organization. In Section 2 we formally define and present some related notations. In Section 3 we describe our multibit encryption scheme (hereafter MGPV) via LWE assumption. In Section 4 we describe our lossy encryption scheme (hereafter LE) via MGPV scheme. In Section 5 we describe a oblivious transfer protocol via the constructed LE scheme. In Section 6, we explore the potential application PAKE in SMDs. Finally, we give a conclusion in Section 7.

2. Preliminaries

Below, we introduce some necessary notations.

2.1. Notation. Throughout our paper, vectors is denoted by bold lower-case letters, e.g., $\mathbf{a}$, and matrices were denoted by upper-case letters, e.g., $\mathbf{A}$. The matrix $\mathbf{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\} \in \mathbb{R}^n$ contains n linearly independent vectors. The basis $\mathbf{B}$ can be used to generate the n-dimensional lattice Λ as follows:

$$\Lambda = L(\mathbf{B}) = \left\{ \mathbf{Bc} = \sum_{i \in [n]} c_i \cdot \mathbf{b}_i : \mathbf{c} \in \mathbb{Z}_q^n \right\}. \quad (1)$$

Below we give a variant of leftover hash lemma.

Lemma 1 (see [20] Lemma 2.1). *We first denote the statistical distance between the distribution $\mathbf{A}$ and $\mathbf{B}$ by $\Delta(\mathbf{A}, \mathbf{B})$. If the parameters satisfy the following conditions, i.e., $\lambda \in \mathbb{Z}$, $n \in \mathbb{N}$, $q \in \mathbb{N}$, $m \geq n \log q + 2\lambda$, $\mathbf{A}$ is a uniform random matrix over $\mathbb{Z}_q^{m \times n}$, $\mathbf{r} \xleftarrow{R} \{0,1\}^m$ and $\mathbf{y} \xleftarrow{R} \mathbb{Z}_q^n$. Then we have that*

$$\Delta\left(\left(\mathbf{A}, \mathbf{A}^T \cdot \mathbf{r}\right), (\mathbf{A}, \mathbf{y})\right) \leq 2^{-\lambda} \quad (2)$$

2.2. Gaussian Distribution. We denote the truncated discrete Gaussian distribution over $\mathbb{Z}^m$ with parameter σ by $D_{\mathbb{Z}_q^m, \sigma}$ and let $D_{\mathbb{Z}^m, \sigma}$ be $\sqrt{m} \cdot \sigma$-bounded.

Remark 2. If define $D_\sigma \longleftarrow \mathbb{Z} \Longrightarrow e \longleftarrow D_\sigma, |e| \leq \sigma$, then $D_\sigma^m \longleftarrow \mathbb{Z}^m \Longrightarrow \mathbf{e} \longleftarrow D_\sigma^m, |e| \leq \sqrt{m} \cdot \sigma$. Throughout the paper, we suppose $\sigma \geq 2\sqrt{m}$. Therefore, if $\mathbf{e} \longleftarrow D_\sigma^m$ then we have, on average, that $\|\mathbf{e}\| \approx \sqrt{m} \cdot \sigma$.

2.3. Learning with Errors

Definition 3 (LWE distribution). The LWE distribution $\mathcal{A}_{\mathbf{s},\chi}$ over $\mathbb{Z}_q^n \times \mathbb{Z}_q$ is sampled by sampling $\mathbf{a} \in \mathbb{Z}_q^n$ uniformly, choosing $e \longleftarrow \chi$, and outputting $(\mathbf{a}, b = \langle \mathbf{s}, \mathbf{a} \rangle + e \pmod q)$ for a secret vector $\mathbf{s} \in \mathbb{Z}_q^n$.

Below we describe the decision version.

Definition 4 (decision-$\mathsf{LWE}_{n,q,\chi,m}$). Sampled m samples $(\mathbf{a}_i, b_i)$ independently over $\mathbb{Z}_q^n \times \mathbb{Z}_q$. For every sample $(\mathbf{a}_i, b_i)$ the following distributions are indistinguishable (with nonnegligible advantage). (1) $(\mathbf{a}_i, b_i) \longleftarrow \mathcal{A}_{\mathbf{s},\chi}$, (2) $(\mathbf{a}_i, b_i)$ sampled from the uniform distribution.

2.4. Inhomogeneous Short Integer Solution. In this subsection, we review the Inhomogeneous Short Integer Solution (ISIS) problem as follows.

Definition 5 (ISIS). Given an integer q, a public matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, a vector $\mathbf{b} \in \mathbb{Z}_q^n$, and a real β, then find an integer vector $\mathbf{e} \in \mathbb{Z}^m$ such that $\mathbf{Ae} = \mathbf{b} \pmod q$ and $\|\mathbf{e}\|_2 \leq \beta$.

2.5. Lossy Encryption

Definition 6 ((perfectly) lossy encryption [21]). An encryption scheme $(\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$ is called "lossy" if there exists a probabilistic polynomial time (PPT) algorithm $\mathsf{KeyGen}_{\mathsf{lossy}}$ that takes 1^λ as input and outputs pk_{lossy} such that

(i) the distribution pk_{lossy} is computationally indistinguishable from a public key pk generated by $\mathsf{Gen}(1^\lambda)$;

(ii) for every two equal-length messages m_0 and m_1, the distributions $\mathsf{Enc}(pk_{\mathsf{lossy}}, m_0)$ and $\mathsf{Enc}(pk_{\mathsf{lossy}}, m_1)$ are identically distributed for every $pk_{\mathsf{lossy}} \longleftarrow \mathsf{KeyGen}_{\mathsf{lossy}}(1^\lambda)$.

Remark 7. If an encryption scheme is lossy then it is semantically secure.

It is given by a tuple of PPT algorithms

$$\left\{\mathsf{KeyGen}_{\mathsf{real}}, \mathsf{KeyGen}_{\mathsf{lossy}}, \mathsf{Enc}, \mathsf{Dec}\right\}. \quad (3)$$

The details are as follows:

(i) $\mathsf{KeyGen}_{\mathsf{real}}(1^\lambda, \mathsf{inj})$ takes as input a security parameter λ and outputs either the real public key along with the secret key (pk_{real}, sk) or the injective key.

(ii) $\mathsf{KeyGen}_{\mathsf{lossy}}(1^\lambda, \mathsf{lossy})$ takes as input λ and outputs a lossy public key and $\perp$ instead of sk, i.e., $(pk_{\mathsf{lossy}}, \perp)$.

(iii) $\mathsf{Enc}(pk, m)$ takes as input either pk_{real} or pk_{lossy} and message m and outputs a ciphertext C.

(iv) $\mathsf{Dec}(sk, C)$ takes as input a secret key sk and a ciphertext C and outputs either a message m or $\perp$.

Lemma 8 (see [15]). *If consider all but a $2q^{-n}$ fraction of all matrix $\mathbf{A}$ over $\mathbb{Z}_q^{m \times n}$ along with any $s \geq \omega(\sqrt{\log m})$ and $m \geq 2n \log q$, then the distribution of $\mathbf{u}^T = \mathbf{e}^T \cdot \mathbf{A} \pmod q$ is statistically close to uniform over $\mathbf{Z}_q^n$, where $\mathbf{e} \sim \mathcal{D}_{\mathbb{Z},s}^m$.*

Lemma 9 (see [22]). *We first denote the distribution $\mathcal{D}$ over $\mathbb{Z}_q^n$ with min-entropy k. If $\mathbf{C} \longleftarrow \mathbb{Z}_q^{l \times n}$ is a uniform matrix and $\mathbf{s} \in \mathbb{Z}_q^{n \times n}$ is sampled from the distribution $\mathcal{D}$ for any $\varepsilon > 0$ and $l \leq (k - 2\log(1/\varepsilon) - O(1))/\log q$, then the joint distribution of $(\mathbf{C}, \mathbf{C} \cdot \mathbf{s})$ is ε-close to the uniform distribution over $\mathbb{Z}_q^{l \times n} \times \mathbb{Z}_q^l$.*

Lemma 10 (see [22]). *Consider a distribution "Lossy" for $\overline{\mathbf{A}} \longleftarrow Lossy \approx_c \mathbf{U} \longleftarrow \mathbb{Z}_q^{m \times n}$. If given $\mathbf{s} \longleftarrow \mathbb{Z}_q^n$ and $\mathbf{e} \longleftarrow \mathcal{A}_{\mathbb{Z},\beta \cdot q}^{m \times n}$, there exists $\widetilde{\mathbf{H}}_\infty^\varepsilon(\mathbf{s} \mid \overline{\mathbf{A}}, \overline{\mathbf{A}} \cdot \mathbf{s} + \mathbf{x}) \geq n$ for $\varepsilon = negl(\lambda)$. Then the Lossy distribution is as follows:*

(i) *Choose $\mathbf{D} \longleftarrow \mathbb{Z}_q^{m \times k}$, $\mathbf{C} \longleftarrow \mathbb{Z}_q^{k \times n}$, and $\mathbf{Z} \longleftarrow \mathcal{D}_{\alpha q}^{m \times n}$, where $\alpha/\beta = negl(\lambda)$ and $k \cdot \log q \leq n - 2\lambda + 2$.*

(ii) *Let $\overline{\mathbf{A}} = \mathbf{DC} + \mathbf{Z}$.*

(iii) *Output $\overline{\mathbf{A}}$.*

2.6. The Universal Composability (UC) Framework. The UC framework first defines a PPT environment machine $\mathcal{Z}$ and then uses the machine to oversee the execution of a protocol in one of two worlds. The detailed description of the executions was presented by Canetti [23], and there exist two world ensembles $\mathsf{EXEC}_{\pi,\mathcal{A},\mathcal{Z}}$ and $\mathsf{IDEAL}_{F,\mathcal{S},\mathcal{Z}}$ for real world and ideal world, respectively.

Definition 11 (see [10] Def.2.1). If there exists a simulator $\mathcal{S}$ for any adversary $\mathcal{A}$ such that for all environments $\mathcal{Z}$

$$\mathsf{IDEAL}_{F,\mathcal{S},\mathcal{Z}} \overset{c}{\approx} \mathsf{EXEC}_{\pi,\mathcal{A},\mathcal{Z}}, \tag{4}$$

then we can say that a protocol π is UC-realize a functionality F.

3. Multibit GPV Scheme

This paper aims to obtain an efficient multibit OT protocol via the multibit lossy encryption. But before designing the multibit lossy encryption scheme, we first present how to obtain the building block multibit GPV scheme of the lossy encryption scheme. Below, we follow the multibit FHE framework of Li et al. [22], and we first develop the multibit GPV (MGPV) scheme. Notably, most of existing LWE-based encryption schemes focus on how to enrich the functions of the single-bit encryption that was originally proposed by Regev [14], such as chosen-plaintext-attacker (CPA) secure schemes [10, 24], chosen-ciphertext-attacker (CCA) secure schemes [25, 26], fully homomorphic encryption (FHE) schemes [20], and oblivious transfer [10, 17].

Notably several recent works (such as [10, 27]) have formally shown the properties of multibit encryption. Fortunately, a multibit Regev [14] scheme was provided by Peikert et al. [10], which is called "pack ciphertext method" and was used as a crux tool to construct multibit FHE schemes [27]. Similarly, many works extended the work of Gentry-Peikert-Vaikutanathan (GPV) to the multibit scheme in the same way [18]. Along with this line, Lindner and Peikert [24] considered a new scheme under the multibit setting, where it is possible to encrypt multiple bits at one-time and makes PKE even more efficient. However, all of the mentioned schemes are constructed by a straightforward concatenation method with the inefficient performance. An important question is raised naturally.

> *Is it possible to explore a new method to design the multiple bits* GPV *encryption under the* LWE *assumption instead of the method of straightforward concatenation?*

We formally explore this important question in this section and we believe that the multibit GPV based on the public key with a sequence of LWE instances might offer many advantages over other approaches. The main ideas behind our method to design the MGPV scheme is described in following sections.

3.1. MGPV *Scheme.* Below we describe the MGPV scheme and its properties.

(i) $\mathsf{params} \longleftarrow \mathsf{MGPV.Setup}(1^\lambda)$:

(1) Take λ as input and output the common parameter $\mathsf{params} = (n, q, \chi, m, t)$, and let $l = \lfloor \log q \rfloor + 1$. We remark that this Setup algorithm is identical to the GPV [15] scheme except that we let a parameter t be the number of secret keys.

(ii) $(pk, sk) \longleftarrow \mathsf{MGPV.KeyGen}(\mathsf{params})$:

(1) Sample $\mathbf{e}_i \longleftarrow \chi^{n\times 1}$, $i \in [t]$ and output $sk_i := \mathbf{e}_i' \longleftarrow (\mathbf{I}_i, -\mathbf{e}_i^T)^T = (0, \cdots, 1, \cdots, 0 \mid -e_{i,1}, \cdots, -e_{i,n}) \in \chi^{n+t}$, and the i-th position is 1.

(2) Choose a matrix $\mathbf{B} \longleftarrow \mathbb{Z}_q^{m\times n}$ and compute $\mathbf{u}_i = \mathbf{B}\mathbf{e}_i \in \mathbb{Z}_q^{m\times 1}$, then we set $\mathbf{A} = [\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_t \mid \mathbf{B}] \in \mathbb{Z}_q^{m\times(n+t)}$.

(3) Output $pk \longleftarrow \mathbf{A}$ and $sk \longleftarrow \{\mathbf{e}_1', \cdots, \mathbf{e}_t'\}$.

(iii) $\mathbf{c} \longleftarrow \mathsf{MGPV.Enc}(\mathsf{params}, pk, \mathbf{m})$:

(1) Set $\mathbf{m} := (m_1, \cdots, m_t)$, $m_i \in \{0, 1\}$ and define $\mathbf{m}' := (\mathbf{m} \mid \mathbf{0}) \in \mathbb{Z}_q^{1\times(n+t)}$.

(2) Sample $\mathbf{x}^T = (\mathbf{x}_1^T \mid \mathbf{x}_2^T) = (x_{1,1}, \cdots, x_{1,t} \mid x_{2,1}, \cdots, x_{2,n}) = (\mathbf{x}_1^T \longleftarrow \chi^{1\times t}, \mathbf{x}_2^T \longleftarrow \chi^{1\times n}) \in D_{\mathbb{Z}^{1\times(n+t)}}$, then choose $\mathbf{r} \longleftarrow \mathbb{Z}_q^{m\times 1}$.

(3) Compute $\mathbf{c} = \mathbf{A}^T \cdot \mathbf{r} + \lfloor q/2 \rfloor \cdot \mathbf{m}^T + \mathbf{x} \in \mathbb{Z}_q^{(n+t)\times 1}$, where the size of ciphertext is $O((n+t)\log^2 q)$.

(iv) $\mathbf{m}' \longleftarrow \mathsf{MGPV.Dec}(\mathsf{params}, sk, \mathbf{c})$:

(1) In order to make the reader understand the structure of the secret ky matrix $\mathbf{S} = (sk_1, \cdots, sk_t) \in \{0, 1\}^{(n+t)\times t}$, the detailed form of the matrix is as follows:

$$\mathbf{S} = \left(\mathbf{e}_1', \cdots, \mathbf{e}_t'\right) = \left(\begin{array}{c|c|c} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ \hline -e_{1,1} & \cdots & -e_{t,1} \\ \vdots & \ddots & \vdots \\ -e_{1,n} & \cdots & -e_{t,n} \end{array}\right). \tag{5}$$

(2) Then compute and output

$$\begin{aligned} \langle \mathbf{c}, \mathbf{S} \rangle &= \left\langle \mathbf{A}^T \cdot \mathbf{r} + \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m}^T + \mathbf{x}^T, \mathbf{S} \right\rangle \\ &= \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m}^T \cdot \mathbf{S} + \mathbf{x}^T \cdot \mathbf{S} \\ &= \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m}' + \left(x_{1,1} - \mathbf{x}_2\mathbf{e}_1, \cdots, x_{1,t} - \mathbf{x}_2\mathbf{e}_t\right) \\ &\quad \left(\mathrm{mod}\, q\right). \end{aligned} \tag{6}$$

We note that the magnitude of the vector $(x_{1,1} - \mathbf{x}_2\mathbf{e}_1, \cdots, x_{1,t} - \mathbf{x}_2\mathbf{e}_t)$ can be regarded as the form of $t \cdot |(x_{1,i} + \mathbf{x}_2^T \cdot \mathbf{e}_i)|$ for $i \in [t]$. If $\|\langle \mathbf{c}, \mathbf{S}\rangle\| \leq t(n+1)B < q/4$, then set $m_i = 1$ and otherwise set $m_i = 0$. Output $\mathbf{m}' = (m_1, \cdots, m_t)$.

Remark 12. We stress that the ciphertext can be decrypted in a bit-by-bit manner. Once we have the secret key matrix $\mathbf{S}$, we can choose the i-th column of $\mathbf{S}$ to recover the i-th bit of the plaintext. In more detail,

(1) we use i-th column vector $\mathbf{s}_i$ from $\mathbf{S}$ to get the i-th position bit of message;

(2) compute and output $\langle \mathbf{c}, \mathbf{s}\rangle = \langle \mathbf{A}^T\mathbf{r} + \lfloor q/2 \rfloor \mathbf{m} + \mathbf{x}^T, \mathbf{s}\rangle = \lfloor q/2 \rfloor m_i + (x_{1,i} + \mathbf{x}_2^T \cdot \mathbf{e}_i) \pmod q$.

If $\|\langle \mathbf{c}, \mathbf{S}\rangle\| \leq (n+1)B < q/4$, then set $m_i = 1$ and otherwise set $m_i = 0$. Output m'.

3.2. Correctness. In this subsection, we analyze the magnitude of the noise.

Lemma 13 (correctness). *Consider the decryption algorithm decrypts in a bit-by-bit manner. If the ciphertext is* $\mathbf{c} = \mathbf{A}^T \cdot \mathbf{r} + \lfloor q/2 \rfloor \cdot \mathbf{m}^T + \mathbf{x}^T \pmod q \in \mathbb{Z}_q^{(m+t)\times 1}$ *under the i-th column secret key* $sk_i = \mathbf{e}_i' \in \mathbb{Z}_q^{(m+t)\times 1}$*, then we have that*

$$\begin{aligned} \langle \mathbf{c}, \mathbf{e}_i \rangle &= \left\lfloor \frac{q}{2} \right\rfloor \cdot m_i + \left(x_{1,i} + \mathbf{x}_2^T \cdot \mathbf{e}_i\right) \\ &= \left\lfloor \frac{q}{2} \right\rfloor \cdot m_i + error \pmod q, \end{aligned} \tag{7}$$

with $|error| < E \leq \lfloor q/2 \rfloor/2$*. Hence, for the secret key matrix* $sk = \mathbf{S} \in \mathbb{Z}_q^{(m+t)\times t}$*, we get the following result:*

$$\langle \mathbf{c}, \mathbf{S}\rangle = \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m}' + \mathbf{x}^T \cdot \mathbf{S} = \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m}' + t \cdot error \pmod q, \tag{8}$$

with $|t \cdot error| < t \cdot E \leq \lfloor q/2 \rfloor/2$*. Hence, there exists* $\mathbf{m}' \longleftarrow \mathsf{Dec}(sk, \mathbf{m})$.

Proof. Consider the following parameters $\mathbf{x}^T = (\mathbf{x}_1^T \longleftarrow \chi^{1\times t}, \mathbf{x}_2^T \longleftarrow \chi^{1\times m})$ and $\forall x_i \longleftarrow \chi$, $|x_i| \leq B$ (where $B \ll q$). Thus, we can get

$$\begin{aligned} \left\langle \mathbf{c}, \mathbf{e}_i' \right\rangle &= \mathbf{r}^T \cdot \mathbf{A}^T \cdot \mathbf{e}_i' + \left\lfloor \frac{q}{2} \right\rfloor \mathbf{m}^T \cdot \mathbf{e}_i' + \mathbf{x}^T \cdot \mathbf{e}_i' \\ &= 0 + \left\lfloor \frac{q}{2} \right\rfloor \cdot m_i + \left(x_{1,i} + \mathbf{x}_2^T \cdot \mathbf{e}_i\right) \\ &= \left\lfloor \frac{q}{2} \right\rfloor \cdot m_i + error_D \pmod q, \end{aligned} \tag{9}$$

with $\|error_D\| \leq \|x_{1,i}\| + \|\mathbf{x}_2^T \cdot \mathbf{e}_i\| \leq B + mB \leq E_D$; the norm of $x_{1,i} + \mathbf{x}_2^T \cdot \mathbf{e}_i$ is bounded by $(m+1)B$, where E_D is denoted as the norm of error elements.

Hence, we can easily obtain the result $\|t \cdot error_D\| \leq t \cdot E_D$ for $\langle \mathbf{c}, \mathbf{S}\rangle = \lfloor q/2 \rfloor \cdot \mathbf{m}' + t \cdot error_D \pmod q$. □

3.3. Security

Theorem 14. *Regarding the following two distributions* $\mathcal{X}$ *and* $\mathcal{Y}$*,*

(i) *the distribution* $\mathcal{X}$ *is denoted as matrices* $[\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_t \mid \mathbf{B}]$ *on* $m \times (t+n)$*, where* $\mathbf{B} \in \mathbb{Z}_q^{m\times n}$ *is a uniform matrix for all* $1 \leq i \leq t$*,* $\mathbf{u}_i = \mathbf{B}\mathbf{e}_i \pmod q$*, and* $\mathbf{e}_i$ *is sampled from* χ^n*.*

(ii) *the distribution* $\mathcal{Y}$ *is denoted as the uniform on* $\mathbb{Z}_q^{m\times(t+n)}$*.*

If the (n, q, χ, m)-ISIS *assumption is hard for the parameters* $m > n \in \mathbb{N}$*,* $q \in \mathbb{N}$*,* $\chi \longleftarrow \mathbb{Z}$*, and* $t = O(\log(n))$ *being an integer, then the distribution* $\mathcal{X}$ *is computationally indistinguishable from* $\mathcal{Y}$*.*

The following theorem formalizes the key result used to show the security of MGPV scheme. We show the scheme is *IND-CPA* secure by using Theorem 14.

Theorem 15. *If the* ISIS *assumption and* LWE *assumption hold for the parameters* params $= (n, q, \chi, m, t)$*, then the* MGPV *scheme is IND-CPA-secure.*

Proof. The high-level proof is as follows:

(i) Firstly, armed with the ISIS assumption, the matrix $\mathbf{A} = [\mathbf{u}_1, \cdots, \mathbf{u}_t \mid \mathbf{B}] \in \mathbb{Z}_q^{n\times(m+t)}$ is computationally indistinguishable from a uniform random matrix by applying the Theorem 14.

(ii) Secondly, the matrix $\mathbf{Ar} + \mathbf{e}$ is indistinguishable from uniform under the LWE assumption and the leftover hash lemma.

This concludes the proof of the theorem. □

3.4. Oblivious Transfer via MGPV. In this subsection, we instantiate the OT protocol under the LWE assumption that provides security for the sender against an honest-but-curious receiver and security for the receiver against a cheating sender.

OT protocol contains two phases, as shown in Figure 1, the initialization phase (i.e., Setup) and the transfer phase (i.e., Transfer) [10].

(i) The Setup phase: the *sender* $\mathbf{S}$ owns 2 elements $\mathbf{m}_0$ and $\mathbf{m}_1$. The *receiver* samples a choice bit $b \in \{0, 1\}$.

(ii) The Transfer phase:

(1) At the beginning of each transfer, the *receiver* $\mathbf{R}$ has an input choice bit b, and he invokes the KeyGen$(\cdot)$ algorithm and outputs a pair (pk_b, sk_b), then he draws a vector as pk_{1-b} from the distribution $\mathbb{Z}_q^{n\times(n+t)}$, then $\mathbf{R}$ sends the pair (pk_0, pk_1) to the *sender* $\mathbf{S}$.

(2) Upon receiving the pair (pk_0, pk_1), the *sender* $\mathbf{S}$ inputs 2 elements $\mathbf{m}_0$ and $\mathbf{m}_1$ and invokes the Enc$(\cdot)$ algorithm to encrypt them under the

Receiver $\mathbf{R}(1^\lambda, b)$		*Sender* $\mathbf{S}$
$(pk_b, sk_b) \leftarrow \mathsf{KeyGen}(1^\lambda, b)$;		
Set $pk_{1-b} \leftarrow \mathbb{Z}_q^{n\times(n+t)}$	$\xrightarrow{(pk_0, pk_1)}$	*Draw* $\mathbf{m}_0, \mathbf{m}_1 \in \{0,1\}^t$
		$\mathbf{c}_0 \leftarrow \mathsf{Enc}(pk_0, \mathbf{m}_0)$
	$\xleftarrow{(\mathbf{c}_0, \mathbf{c}_1)}$	$\mathbf{c}_1 \leftarrow \mathsf{Enc}(pk_1, \mathbf{m}_1)$
$\mathbf{m}_b \leftarrow \mathsf{Dec}(sk_b, \mathbf{c}_b)$		

* $\mathscr{F}_{CRS}^{\mathscr{D}}$ runs with parties $P_1, \cdots, P_n$. The common reference string functionality $\mathscr{F}_{CRS}^{\mathscr{D}}$ [10]

Figure 1: Function $\mathscr{F}_{CRS}^{\mathscr{D}}$, * [10].

pk_0, pk_1, respectively, then outputs the ciphertext $\mathbf{c}_0$ and $\mathbf{c}_1$ and sends back to the receiver $\mathbf{R}$.

(3) Upon receiving the pair $\mathbf{c}_0, \mathbf{c}_1$, the receiver invokes the $\mathsf{Dec}(\cdot)$ algorithm and outputs $\mathbf{m}_b$.

3.5. Security. The OT protocol is a simple application of MGPV scheme.

Theorem 16 (see [16] Theorem 5.2). *We say that the OT protocol is secure for the receiver if the* $\mathsf{LWE}_{n,q,m,\chi}$ *problem is hard.*

Proof. The above security proof is simple, so we omit further details and recommend the reader to find further details from the proof of Theorem 5.2 in Lyubashevsky et al. at TCC'10 [16]. □

Theorem 17 (see [16] Theorem 5.3). *We say that the above OT protocol is secure for the sender against an honest-but-curious receiver, if the* $\mathsf{LWE}_{n,q,m,\chi}$ *assumption is hard for the input message length t of the sender.*

Proof. The detailed proof can be found from the Theorem 5.3 in Lyubashevsky et al. at TCC'10 [16]. □

4. Lossy Encryption (LE Scheme)

The notation of "lossy encryption" was proposed by Bellare-Hofheinz-Yilek (BHY) [11]. Actually, the lossy encryption is an extension of the meaningful/meaningless encryption [12] and dual mode encryption [10]. At a high level, a "lossy" (or "messy" in [10]) cryptosystem is one which also has two modes according to two types of public keys. Concretely, **(1)** in the normal mode, the ciphertext is generated by encrypting the plaintext under an injective key. **(2)** In the lossy (or "messy") mode, the ciphertext is independent of the plaintext. Actually, the operability property was proposed by [11]; they basically allow a possibly inefficient algorithm to open a ciphertext generated under a lossy key to any plaintext. Meanwhile, the injective key is computationally indistinguishable from the lossy key.

4.1. Multibit Lossy Encryption Scheme. Gentry et al. [15] proposed the dual Regev scheme to design the identity-based encryption (IBE) with the random oracle. Then, Agrawal et al. [28] used it to design the IBE scheme in the standard model. In this paper, we construct the LWE-based lossy encryption from multiple bits GPV. However, the process of encryption is different from GPV. In our construction, we only sample the noise vector $\mathbf{e}$ one-time rather than twice $\mathbf{AS}$ in the GPV scheme. The concrete construction is as follows:

(i) $(\mathsf{crs}, \mathsf{params}) \longleftarrow \mathsf{Setup}(1^\lambda)$:

(1) Set $m \geq 2n \cdot \log q$ and secure parameter λ. Since Lemma 10, we set $k \cdot \log q \leq n - 2 \cdot \lambda + 2$, $l \leq (k - 2\log(1/\varepsilon) - O(1))/\log q$, $q \geq 5rm$, $r \geq \omega(\sqrt{\log m})$, $\beta \leq 1/(r\sqrt{m} \cdot \omega(\sqrt{\log m}))$, $\beta \cdot q > O(2 \cdot \sqrt{n})$, and $\alpha/\beta = negl(\lambda)$. To satisfy these requirements, q should be superpolynomial of the secure parameter λ, moreover, t as described in MGPV scheme.

(2) Output $\mathsf{params} := (m, n, q, \chi, k, l, t)$ and $\mathsf{crs} := \mathbf{B}$, where $\mathbf{B} \longleftarrow \mathbb{Z}_q^{m\times n}$.

(ii) $(pk_{\mathsf{real}}, sk) \longleftarrow \mathsf{KeyGen}_{\mathsf{real}}(\mathsf{params})$:

(1) For $i \in [t]$, $\mathbf{e}_i \longleftarrow \chi^{n\times 1}$, then we have that $\mathbf{u}_i := \mathbf{A}\mathbf{e}_i \in \mathbb{Z}_q^{m\times 1}$. Compose all $\mathbf{e}_i$ together, then we have that $\mathbf{E} = [\mathbf{e}_1, \cdots, \mathbf{e}_t] \in \mathscr{D}_{\mathbb{Z},r}^{n\times t}$.

(2) Hence, $pk_{\mathsf{real}} := \mathbf{A} = [\mathbf{u}_1, \cdots, \mathbf{u}_t \mid \mathbf{B}] \in \mathbb{Z}_q^{m\times(n+t)}$.

(3) Output $sk := \mathbf{S} = [\mathbf{I} \mid \mathbf{E}]$ as described above.

(iii) $(pk_{\mathsf{lossy}}, \perp) \longleftarrow \mathsf{KeyGen}_{\mathsf{lossy}}(\mathsf{params})$:

(1) Choose $\mathbf{D} \longleftarrow \mathbb{Z}_q^{m\times k}$, $\mathbf{C} \longleftarrow \mathbb{Z}_q^{k\times n}$, $\mathbf{Z} \longleftarrow \mathscr{D}_{\mathbb{Z},\alpha\cdot q}^{m\times n}$, and $\mathbf{U} \in \mathbb{Z}_q^{m\times t} := [\mathbf{u}_i', \ldots, \mathbf{u}_t']$, where $\mathbf{u}_i' \longleftarrow \mathbb{Z}_q^{m\times 1}$.

(2) Output $sk := \perp$, $pk_{\mathsf{lossy}} := (\mathbf{u}_1', \cdots, \mathbf{u}_t' \mid \mathbf{DC} + \mathbf{Z})$.

(iv) $\mathbf{c} \longleftarrow \mathsf{Enc}(pk, \mathbf{m})$:

(1) Denote $\mathbf{m}' = [\mathbf{m} \mid \mathbf{0}] \in \{0,1\}^{1\times(n+t)}$.

(2) Choose random vectors $\mathbf{r} \longleftarrow \mathbf{Z}_q^{m\times 1}$, $\mathbf{x} \longleftarrow \mathscr{D}_{\mathbb{Z},\beta\cdot q}^{(n+t)\times l}$.

(3) Compute and output ciphertexts: $\mathbf{c} := \mathbf{A}^T \cdot \mathbf{r} + \lfloor q/2 \rfloor \mathbf{m}'^T + \mathbf{x} \in \mathbb{Z}_q^{(n+t)\times 1}$.

(v) $\mathbf{m}' \longleftarrow \mathsf{Dec}(sk, \mathbf{c})$: Compute and output: $\langle \mathbf{c}, \mathbf{S} \rangle = \lfloor q/2 \rfloor \mathbf{m} + \mathbf{x}^T \cdot \mathbf{S}$.

In order to construct the oblivious transfer, we need to design a verification algorithm $\mathsf{Verify}(\cdot)$ for the sender $\mathbf{S}$, who is similar to the $\mathsf{FindMessy}(\cdot)$ in [10], and will use the $\mathsf{Verify}(\cdot)$ to verify that the public key pk_σ from the receiver $\mathbf{R}$ is pk_{real} or pk_{lossy}, in more detail:

(i) $b \longleftarrow \mathsf{Verify}(pk_\sigma, \sigma)$: for $\sigma \in \{0,1\} := \{\mathsf{real}, \mathsf{lossy}\}$, the key generation takes a chosen decryptable branch $\sigma \in \{0,1\}$ as a parameter, and the resulting secret key sk_σ corresponds to branch σ of public key pk_σ. Then, we use $b \in \{0,1\}$ to distinguish the two messages. Actually, messages encrypted on branch $b = \sigma$ can be decrypted using sk_σ, while those on the other branch cannot.

Below, we show that this scheme fulfills the properties of lossy encryption.

Proposition 18. *Correctness on Real Keys. For all* (pk_{real}, sk) *generated by* $\mathsf{KeyGen}_{\mathsf{real}}(1^\lambda)$ *and all message* $\mathbf{m}$,

$$\begin{aligned}\mathsf{Dec}\left(sk, \mathsf{Enc}\left(pk_{\mathrm{real}}, \mathbf{m}\right)\right) &= \mathsf{Dec}\left(\mathbf{S}, \mathsf{Enc}\left(\mathbf{A}, \mathbf{m}'\right)\right)\\ &= \mathsf{Dec}\left(\mathbf{S}, \left(\mathbf{A}^T\mathbf{r} + \left\lfloor \frac{q}{2} \right\rfloor \mathbf{m}'^T + \mathbf{x}\right)\right)\\ &= \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m} + error\end{aligned} \tag{10}$$

the algorithm $\mathsf{Dec}(\cdot)$ *will get the correct message with overwhelming probability.*

We need to remark that, considering the parameters $q \geq 5rm$ and $\beta \leq 1/(r\sqrt{m} \cdot \omega(\sqrt{\log q}))$ which were denoted in [15]. Then $\mathsf{Dec}(sk, \mathbf{c})$ decrypts correctly with overwhelming probability (over the random choices of $\mathsf{KeyGen}_{\mathsf{real}}(1^\lambda)$ and $\mathsf{Enc}(pk, \mathbf{m})$).

Proposition 19. *Lossiness of Encryption with Lossy Keys. In more detail*

$$\begin{aligned}\mathsf{Enc}\left(pk_{\mathsf{lossy}}, \mathbf{m}\right) &= \mathsf{Enc}\left(\underline{(\mathbf{u}_1, \ldots, \mathbf{u}_t \mid \mathbf{DC} + \mathbf{Z})}, \mathbf{m}'\right)\\ &= \underline{(\mathbf{u}_1, \ldots, \mathbf{u}_t \mid \mathbf{DC} + \mathbf{Z})}^T \cdot \mathbf{r} + \left\lfloor \frac{q}{2} \right\rfloor\\ &\quad \cdot \mathbf{m}' + \mathbf{x} \pmod q\end{aligned} \tag{11}$$

Parse public key $\mathbf{A}$ *as* $pk1 := (\mathbf{u}_1, \ldots, \mathbf{u}_t)$ *and* $pk2 := \mathbf{B}$, *by Lemma 10,* $\widetilde{\mathbf{H}}_\infty^\varepsilon(\mathbf{r} \mid \overline{\mathbf{B}}, \overline{\mathbf{B}} \cdot \mathbf{r} + \mathbf{x}) \geq m$ *since* $m \leq (\lambda - 2\log(1/\varepsilon) - O(1))/\log q$, *and by Lemma 9, given* $(\mathbf{DC}+\mathbf{Z})\cdot\mathbf{r}+\mathbf{x}$, $\underline{(\mathbf{u}_1, \ldots, \mathbf{u}_t)}^T \cdot \mathbf{r}$ *is* ε*-close to* $\mathcal{U}(\mathbb{Z}_q^{t\times 1})$. *When* $\varepsilon = negl(\lambda)$, $\underline{(\mathbf{u}_1, \ldots, \mathbf{u}_t)}^T \cdot \mathbf{r} \approx_s \mathcal{U}(\mathbb{Z}_q^{t\times 1})$ *given* $(\mathbf{DC} + \mathbf{Z}) \cdot \mathbf{r} + \mathbf{x}$. *Therefore,* $\forall \mathbf{m} \in \mathcal{M}$, *given* $(\mathbf{DC} + \mathbf{Z})\mathbf{r} + \mathbf{x}$, *i.e.,*

$$\underline{(\mathbf{u}_1, \ldots, \mathbf{u}_t)}^T \cdot \mathbf{r} + \left\lfloor \frac{q}{2} \right\rfloor \cdot \mathbf{m} \approx_s \mathcal{U}\left(\mathbb{Z}_q^{t\times 1}\right) \tag{12}$$

for any lossy keys pk_{lossy} *generated by* $\mathsf{KeyGen}_{\mathsf{lossy}}(1^\lambda)$ *and any two messages* $\mathbf{m}_0 \neq \mathbf{m}_1$, *holds*

$$\mathsf{Enc}\left(pk_{\mathsf{lossy}}, \mathbf{m}_0\right) \approx_s \mathsf{Enc}\left(pk_{\mathsf{lossy}}, \mathbf{m}_1\right) \tag{13}$$

Proposition 20. *Indistinguishability between Real Public Key and Lossy Public Key.* pk_{real} *is* $(\underline{\mathbf{u}_1, \ldots, \mathbf{u}_t} \mid \mathbf{B})$, *and* pk_{lossy} *is* $(\underline{\mathbf{u}'_1, \ldots, \mathbf{u}'_t} \mid (\mathbf{DC} + \mathbf{Z}))$. *Since* $m \geq 2n\log q$, *by Lemma 8, sample* $\mathbf{U}_1 \longleftarrow \mathbb{Z}_q^{m\times t}$ *and* $\mathbf{U}_2 \longleftarrow \mathbb{Z}_q^{m\times n}$.

$$\left(\underline{\mathbf{B} \cdot \mathbf{e}_1, \ldots, \mathbf{B} \cdot \mathbf{e}_t} \mid \mathbf{B}\right) \approx_s (\mathbf{U}_1, \mathbf{U}_2) \tag{14}$$

Under the hardness of LWE, $(\mathbf{U}_1, \mathbf{U}_2) \approx_c (\underline{\mathbf{u}_1, \ldots, \mathbf{u}_t} \mid \mathbf{DC} + \mathbf{Z})$,

$$\therefore \left(\underline{\mathbf{u}_1, \ldots, \mathbf{u}_t} \mid \mathbf{DC} + \mathbf{Z}\right) \approx_c \left(\underline{\mathbf{B} \cdot \mathbf{e}_1, \ldots, \mathbf{B} \cdot \mathbf{e}_t} \mid \mathbf{B}\right). \tag{15}$$

i.e., pk_{real} *and* pk_{lossy} *are computationally indistinguishable.*

5. OT via Lossy Encryption

In this section, inspired by David et al. [21] UC-secure OT protocol via lossy encryption using the McEliece assumption over code-based cryptography, we present an UC-secure OT protocol via lattice-based lossy encryption using LWE and ISIS assumption.

5.1. Our Construction: UC-Secure OT *for Ideal Functionalities.* Before describing our construction, we first denote the ideal functionalities $\mathcal{F}_{CRS}^{\mathcal{D}}$ and $\mathcal{F}_{\mathsf{OT}}$. In more detail, the CRS functionality $\mathcal{F}_{CRS}^{\mathcal{D}}$ outputs a string with a fixed distribution as depicted in Algorithm 1.

As shown in Algorithm 2, the two-party functionality OT contains a sender $\mathbf{S}$ with input x_0, x_1 and receiver $\mathbf{R}$ with an input $\sigma \in \{0,1\}$. Importantly, the OT functionality $\mathcal{F}_{\mathsf{OT}}$ captures requirement of OT specification.

5.2. Our Construction: OT *from Lossy Encryption.* Below we describe our main contribution, a various OT protocol from lossy encryption in Figure 2.

Simulating the communication with $\mathcal{Z}$: the simulator $\mathcal{S}$ writes every input value from $\mathcal{Z}$ into the input tape of the adversary $\mathcal{A}$. $\mathcal{S}$ copies every output value written by $\mathcal{A}$ to his output tape. The environment $\mathcal{Z}$ can read the output tap.

Simulating R is corrupted: $\mathcal{S}$ simulates the view of the receiver without considering which mode of the protocol and does the following: running the Setup algorithm in messy mode and letting $(\mathsf{crs}, t) \longleftarrow \mathsf{Setup}^{\mathsf{Lossy}}(1^\lambda)$. If the parties query the $\mathcal{F}_{CRS}^{\mathsf{mode}}$, then it obtains the feedback $(\mathsf{sid}, \mathsf{crs})$.

(i) Once the adversary $\mathcal{A}$ generates a message $(\mathsf{sid}, \mathsf{ssid}, pk_{\mathsf{real}}, pk_{\mathsf{lossy}})$, $\mathcal{S}$ extracts the choice bit b of the corrupted receiver and lets $b \longleftarrow \{0,1\}$, then $\mathcal{S}$ sends the command $(\mathsf{sid}, \mathsf{ssid}, \mathsf{receiver}, 1-b)$ to the $\mathcal{F}_{\mathsf{OT}}$, then $\mathcal{F}_{\mathsf{OT}}$ returns the output $(\mathsf{sid}, \mathsf{ssid}, \mathbf{m}_{1-b})$ to $\mathcal{S}$, and $\mathcal{S}$ then stores it along with b.

(ii) Once the dummy simulator $\mathbf{S}$ is activated by the command $(\mathsf{sid}, \mathsf{ssid})$, $\mathcal{S}$ then simulates the $\mathbf{S}$'s behaviour, and looks up the corresponding bit b for $\mathbf{m}_b$ and $\mathbf{m}_{1-b}$ and then computes $\mathbf{c}_b \longleftarrow \mathsf{Enc}(pk, \mathbf{0}^t)$ and $\mathbf{c}_{1-b} \longleftarrow \mathsf{Enc}(pk, \mathbf{m}_{1-b})$ and sends $(\mathsf{sid}, \mathsf{ssid}, \mathbf{c}_0, \mathbf{c}_1)$ to $\mathcal{A}$.

$\mathcal{F}_{CRS}^{\mathcal{D}}$ is parameterized by an algorithm $\mathcal{D}$, and $\mathcal{F}_{CRS}^{\mathcal{D}}$ can interact with parities $P_1, \cdots, P_n$.
(i) Upon receiving a command (sid, P_i, P_j) from the party P_i, first let $\mathsf{crs} \longleftarrow \mathcal{D}(1^\lambda)$, then send the message $(\mathsf{sid}, \mathsf{crs})$ to P_i and send the message (crs, P_i, P_j) to the adversary;
(ii) Upon receiving a command (sid, P_i, P_j) from the party P_j (and only P_j), then send P_j and the adversary the message $(\mathsf{sid}, \mathsf{crs})$, and halt.

Algorithm 1: The CRS functionality $\mathcal{F}_{CRS}^{\mathcal{D}}$ from [10].

$\mathcal{F}_{\mathsf{OT}}$ interacts with a receiver **R** and a sender **S**.
(i) Upon receiving a command $(\mathsf{sid}, \mathsf{sender}, x_0, x_1)$ from **S**, store the pair (x_0, x_1) for $x_i \in \{0,1\}^l$. (Notably, the length of the string l is fixed and all parties know);
(ii) Upon receiving a command $(\mathsf{sid}, \mathsf{receiver}, \sigma)$ from **R**, then check if $(\mathsf{sid}, \mathsf{sender}, \cdots)$ was previously sent and send the message (sid, x_σ) to **R**, and send the adversary $\mathcal{S}$ the message (sid) and halt. Otherwise, send nothing to **R**.

Algorithm 2: The oblivious transfer functionality $\mathcal{F}_{\mathsf{OT}}$.

Simulating S is corrupted: $\mathcal{S}$ does the following without considering which mode of the protocol: running the real (injective) mode KeyGen algorithm and letting $(\mathsf{crs}) \longleftarrow \mathsf{Setup}(1^\lambda)$. If the parties query the ideal functionality $\mathcal{F}_{CRS}^{mode}$, then $\mathcal{F}_{CRS}^{mode}$ returns $(\mathsf{sid}, \mathsf{crs})$ to them.

(i) Once the dummy **R** is activated on by the command $(\mathsf{sid}, \mathsf{ssid})$, $\mathcal{S}$ then simulates the behaviour of **R** and computes $(pk_{\mathsf{real}}, sk_{\mathsf{real}}) \longleftarrow \mathsf{KeyGen}(1^\lambda, inj)$ and $(pk_{\mathsf{lossy}}, sk_{\mathsf{lossy}}) \longleftarrow \mathsf{KeyGen}(1^\lambda)$, and then $\mathcal{S}$ sends $(\mathsf{sid}, \mathsf{ssid}, pk_{\mathsf{real}}, pk_{\mathsf{lossy}})$ to $\mathcal{A}$ and stores $(\mathsf{sid}, \mathsf{ssid}, pk_{\mathsf{real}}, sk_{\mathsf{real}}, sk_{\mathsf{lossy}})$.
(ii) When $\mathcal{A}$ replies with a message $(\mathsf{sid}, \mathsf{ssid}, \mathbf{c}_0, \mathbf{c}_1)$, the $\mathcal{S}$ looks up the corresponding $(pk_{\mathsf{real}}, sk_{\mathsf{real}})$ and $(pk_{\mathsf{lossy}}, sk_{\mathsf{lossy}})$, computes $\mathbf{m}_b \longleftarrow \mathsf{Dec}(sk_b, \mathbf{c}_b)$ for each $b \in \{0,1\}$, and returns $(\mathsf{sid}, \mathsf{ssid}, \mathsf{sender}, \mathbf{m}_0, \mathbf{m}_1)$ to $\widehat{F}_{OT}$.

Simulating the remaining cases: once both parties are corrupted by the adversary, then $\mathcal{S}$ runs $\mathcal{A}$. More concretely, $\mathcal{S}$ internally runs the **S** on input $(\mathsf{sid}, \mathsf{ssid}, \mathbf{m}_0 = \mathbf{0}^t, \mathbf{m}_1 = \mathbf{0}^t)$; meanwhile, it runs the honest **R** on input $(\mathsf{sid}, \mathsf{ssid}, \sigma = 0)$ and honest no matter which party is corrupted. When the corresponding dummy party is activated in the ideal execution, **S** activates the appropriate algorithm and delivers $\mathcal{A}$ all messages between its internal **R** and **S**.

Caim. If $\mathcal{A}$ corrupts **R** in an execution of $\mathsf{LE}^{\mathsf{lossy}}$, i.e., **S** in lossy mode, then we have

$$\mathsf{IDEAL}_{\widehat{\mathcal{F}}_{\mathsf{OT}}, \mathcal{S}, \mathcal{Z}} \approx_s \mathsf{EXEC}_{\mathsf{LE}^{\mathsf{lossy}}, \mathcal{A}, \mathcal{Z}} \tag{16}$$

Proof. Below we give a formal proof, in more detail:

(i) The real world execution can be viewed as the proceed of the following game.
  (a) Firstly, obtain crs by invoking the algorithm $\mathsf{Setup}_1^{\mathsf{lossy}}(1^\lambda)$.
  (b) Secondly, the environment $\mathcal{Z}$ can schedule subsessions arbitrarily. Notably, in each subsession,
    (1) $\mathcal{Z}$ can choose an arbitrary message $(\mathbf{m}_0, \mathbf{m}_1)$ for the honest sender **S**;
    (2) the honest sender **S** sends the ciphertext $\mathbf{c}_b \longleftarrow \mathsf{Enc}(pk, \mathbf{0}^t)$ for each $b \in \{0,1\}$ to $\mathcal{Z}$.

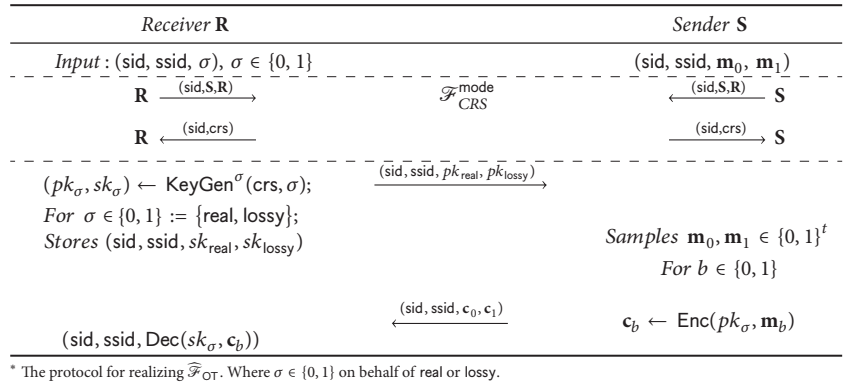
(ii) The ideal world execution can be viewed as the proceed of the following game:
  (a) Firstly, obtain crs by running the algorithm $\mathsf{Setup}^{\mathsf{lossy}}(1^\lambda)$.
  (b) Secondly, the environment $\mathcal{Z}$ schedules subsessions arbitrarily, in each subsession,
    (1) $\mathcal{Z}$ can input an arbitrary pk and the arbitrary input message $(\mathbf{m}_0, \mathbf{m}_1)$ for the dummy sender **S**;
    (2) $\mathcal{S}$ can run the verification algorithm $b \longleftarrow \mathsf{Verify}(\mathsf{crs}, pk_\sigma, \sigma)$ to $\mathcal{Z}$ and learn $\mathbf{m}_{1-b}$ from $\widehat{\mathcal{F}}_{\mathsf{OT}}$. It then sends $\mathcal{Z}$ the $\mathbf{c}_b \longleftarrow \mathsf{Enc}(\mathsf{crs}, pk, 1-b, \mathbf{0}^t)$ and $\mathbf{c}_{1-b} \longleftarrow \mathsf{Enc}(\mathsf{crs}, pk, 1-b, \mathbf{m}_{1-b})$.

We stress that the only difference between the ideal world execution and the real world execution is the generation of $\mathbf{c}_b$ in each subsession. But by lossy key generation in Proposition 19, hence the above two games are statistically indistinguishable. □

Claim. If, in an execution of $\mathsf{LE}^{\mathsf{real}}$ (i.e., **S** in real mode), $\mathcal{A}$ corrupts **S** then we have

$$\mathsf{IDEAL}_{\widehat{\mathcal{F}}_{\mathsf{OT}}, \mathcal{S}, \mathcal{Z}} \approx_s \mathsf{EXEC}_{\mathsf{LE}^{\mathsf{real}}, \mathcal{A}, \mathcal{Z}}. \tag{17}$$

Proof. Below we give a formal proof, in more detail:



* The protocol for realizing $\widetilde{\mathcal{F}}_{\mathrm{OT}}$. Where $\sigma \in \{0,1\}$ on behalf of real or lossy.

FIGURE 2: Protocol $\mathsf{LE}^{\mathsf{mode}}$ for oblivious transfer *.

(i) The real world execution can be viewed as the proceed of the following game:

(a) Firstly, $\mathsf{crs} \longleftarrow \mathsf{Setup}_1^{\mathsf{real}}(1^\lambda)$.

(b) Secondly, the environment arbitrarily schedules some number of subsessions. In each subsession,

(1) $\mathcal{Z}$ chooses an input σ for the honest $\mathbf{R}$, who generates $(pk_\sigma, sk_\sigma) \longleftarrow \mathsf{KeyGen}(\mathsf{crs}, \sigma)$, and sends pk_{real}, pk_{lossy} to $\mathcal{Z}$;

(2) then $\mathcal{D}$ proceeds arbitrarily $(\mathbf{c}_1, \mathbf{c}_2)$ to the honest $\mathbf{R}$ outputs $\mathsf{Dec}(\mathsf{crs}, sk_\sigma, \mathbf{c}_\sigma)$.

(ii) The ideal world execution can be viewed as the proceed of the following game:

(a) Firstly, $(\mathsf{crs}, t) \longleftarrow \mathsf{Setup}^{\mathsf{Dec}}(1^\lambda)$.

(b) Secondly, the environment arbitrarily schedules subsessions, in each subsession,

(1) $\mathcal{Z}$ outputs arbitrary σ which is not known to $\mathbf{S}$;

(2) $\mathcal{S}$ then runs $(pk_{\mathsf{real}}, sk_{\mathsf{real}}) \longleftarrow \mathsf{KeyGen}(\mathsf{crs}, \sigma)$ and $(pk_{\mathsf{lossy}}, sk_{\mathsf{lossy}}) \longleftarrow \mathsf{KeyGen}(\mathsf{crs}, 1-\sigma)$ and sends pk_{real}, pk_{lossy} to $\mathcal{Z}$;

(3) lastly, $\mathcal{S}$ receives the arbitrary ciphertext tuple $(\mathbf{c}_0, \mathbf{c}_1)$ from $\mathcal{Z}$.

□

Remark 21. Actually, the dummy entity $\mathbf{R}$ queries the value of $\mathsf{Dec}(\mathsf{crs}, sk_\sigma, \mathbf{c}_\sigma)$ from the ideal functionality, then the simulator $\mathcal{S}$ provides the messages $\mathbf{m}_b \longleftarrow \mathsf{Dec}(\mathsf{crs}, sk_b, \mathbf{c}_b)$.

The only difference between the two games is method of the public and secret keys. The above two games are statistically indistinguishable by the lossy key generation in Proposition 20.

Claim. There exists the following result:

$$\mathsf{EXEC}_{\pi^{\mathsf{lossy}}, \mathcal{A}, \mathcal{Z}} \approx_c \mathsf{EXEC}_{\pi^{\mathsf{real}}, \mathcal{A}, \mathcal{Z}} \tag{18}$$

for any protocol π^{mode} in the $\mathcal{F}_{CRS}^{\mathsf{mode}}$-hybrid model.

Proof. In the lossy encryption, the output of $\mathcal{F}_{CRS}^{\mathsf{lossy}}$ is computationally indistinguishable from $\mathcal{F}_{CRS}^{\mathsf{real}}$ since the indistinguishability of modes. Moreover, $\mathcal{Z}$ can run the protocol π^{mode} and can receive a polynomial number of samples from either $\mathcal{F}_{CRS}^{\mathsf{lossy}}$ or $\mathcal{F}_{CRS}^{\mathsf{real}}$. Thus, the above two executions are indistinguishable by a standard hybrid argument. □

5.3. Performance. Lattice-based cryptography has been subjecting of intense research appearing recently, bringing groundbreaking advance to the understanding of the subjacent questions. One of the main characteristics of lattice-based cryptography is worst-case to average-case reductions, which provides stronger security against quantum computer attacks. In this paper, we construct a lossy encryption scheme via a variant of multibit GPV scheme, then we construct the universal composable secure OT protocols based on LWE assumption by utilizing the lossy encryption as the building block. Below, a comparison of some related works with our scheme is provided in the Table 1.

As shown in the Table 1, we can easily obtain the following conclusion. We follow the methodology of Li et al. [18] and design a multibit public key encryption scheme, i.e., MGPV scheme. Importantly, the public matrix $\mathbf{A}$ contains many LWE instances; each one is used to protect the secret key. In this setting, the decrypter can decrypt the plaintext either in a bit-by-bit manner or in a one-time manner. Meanwhile, compared the magnitude of ciphertext of PVW scheme [10] with ours, it is easy to see the two schemes with the same magnitude of ciphertext $O(\log_q mn)$. Although the public key size of PVW depends on the parameter n and our scheme's public key size depends on m, the bit decryption of our scheme implies flexible decryption (i.e., multibit decryption), which means that our scheme is more practical in reality.

6. Potential Application: Password-Based Authenticated Key Exchange for Smart Mobile Devices

Nowadays, SMDs, IoTs, and WSNs within the workplace are expanding rapidly. Obliviously, these devices are becoming important tools that offer competitive advantages for the mobile workforce. But they also might be endangered by the

Table 1: The comparison of some related works with our scheme.

Scheme	assumption	message size	bit Dec	one-time Dec	applications
GPV [15]	LWE/SIS	1	✔	✘	PKE&IBE
PVW [10]	LWE	1	✘	✔	PKE&OT
LPS [16]	SSP	t	✔	✔	PKE&OT
Our scheme	LWE	t	✔	✔	PKE&OT

information they can access remotely. In this case, enabling user authentication for SMDs is the first line of defence to prevent the malicious unauthorized user.

Most of related works [5, 29] focus on how to use PAKE as the basic tool to achieve the authentication for SMDs. In particular, Wei et al. [5] proposed a PAKE protocol for wireless body area networks. He et al. [29] proposed an authentication protocol for mobile wireless networks with conditional privacy preservation. However, to our knowledge, related works of lattice-based PAKE for SMDs authentication are limited. Hence, in this section, we explore how to implement PAKE via our OT protocol. Because details of the design and implementation are beyond the scope of the discussion of this paper, thus, we just give a brief of description for the technical line as follows. In more detail, following the technical line of Canetti et al. [1], we first realize OT-based PKAE via LWE assumption instead of computational Diffie-Hellman (CDH) assumption and the hardness of factoring. Next, we can extend the PAKE protocol for privacy-preserving authentication schemes for SMDs.

7. Conclusion

In this paper, we have investigated one of the hot but hard topics in authentication of SMDs, IoTs, and WSNs. As an important building block, OT can be used for designing privacy-preserving authentication protocols. Thus, we focus on an important question how to design on an efficient UC-secure OT protocol for PAKE which can be used to achieve authentication for SMDs. However, an important question that remain is how to implement OT-based PAKE under the LWE assumption following our presented brief technical line. Meanwhile, we believe that this result enriched the postquantum OT protocols. However, it remains open to be secure against adaptive adversaries under the lossy encryption and its variants. We leave these topics for future research.

Data Availability

No data were used to support this study.

Conflicts of Interest

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

Acknowledgments

The research was supported by the National Natural Science Foundation of China (nos. 61802214 and 11701187) and the PhD Start-up Fund of the Natural Science Foundation of Guangdong Province of China (no. 2017A030310522).

References

[1] R. Canetti, D. Dachman-Soled, V. Vaikuntanathan, and H. Wee, "Efficient password authenticated key exchange via oblivious transfer," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 7293, pp. 449–466, 2012.

[2] D. Wang, N. Wang, P. Wang, and S. Qing, "Preserving privacy for free: Efficient and provably secure two-factor authentication scheme with user anonymity," *Information Sciences*, vol. 321, Article ID 11496, pp. 162–178, 2015.

[3] C. Su, B. Santoso, Y. Li, R. H. Deng, and X. Huang, "Universally Composable RFID Mutual Authentication," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 1, pp. 83–94, 2017.

[4] Y. Zhang, J. Yu, R. Hao, C. Wang, and K. Ren, "Enabling Efficient User Revocation in Identity-based Cloud Storage Auditing for Shared Big Data," *IEEE Transactions on Dependable and Secure Computing*, pp. 1-1, 2018.

[5] F. Wei, P. Vijayakumar, J. Shen, R. Zhang, and L. Li, "A provably secure password-based anonymous authentication scheme for wireless body area networks," *Computers and Electrical Engineering*, vol. 65, pp. 322–331, 2018.

[6] L. Zhang, Z. Zhang, and X. Hu, "UC-secure two-server password-based authentication protocol and its applications," in *Proceedings of the 11th ACM Asia Conference on Computer and Communications Security, ASIA CCS 2016*, pp. 153–164, June 2016.

[7] M. O. Rabin, "Probabilistic algorithms in finite fields," *SIAM Journal on Computing*, vol. 9, no. 2, pp. 273–280, 1980.

[8] J. Han, W. Susilo, Y. Mu, M. H. Au, and J. Cao, "AAC-OT: Accountable Oblivious Transfer with Access Control," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2502–2514, 2015.

[9] Y. Gertner, S. Kannan, T. Malkin, O. Reingold, and Viswanathan. M., "The relationship between public key encryption and oblivious transfer," in *Proceedings of the FOCS 2000*, pp. 325–335, IEEE Computer Society Press.

[10] C. Peikert, V. Vaikuntanathan, and B. Waters, "A framework for efficient and composable oblivious transfer," Cryptology ePrint Archive 2007/348, https://eprint.iacr.org/2007/348.

[11] M. Bellare, D. Hofheinz, and S. Yilek, "Possibility and impossibility results for encryption and commitment secure under selective opening," in *Advances in cryptology EUROCRYPT 2009*, vol. 5479, pp. 1–35, Springer, Berlin, Germany, 2009.

[12] G. Kol and M. Naor, "Cryptography and game theory: designing protocols for exchanging information," in *Theory of Cryptography: Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19–21, 2008. Proceedings*, vol. 4948

of *Lecture Notes in Computer Science*, pp. 320–339, Springer, Berlin, Germany, 2008.

[13] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science (SFCS '94)*, pp. 124–134, IEEE Computer Society Press, 1994.

[14] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," in *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC '05)*, pp. 84–93, ACM Press, Baltimore, Md, USA, May 2005.

[15] C. Gentry, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in *Proceedings of the 14th Annual ACM Symposium on Theory of Computing (STOC '08)*, pp. 197–206, Victoria, Canada, May 2008.

[16] V. Lyubashevsky, A. Palacio, and G. Segev, "Public-key cryptographic primitives provably as secure as subset sum," in *Theory of cryptography*, vol. 5978, pp. 382–400, Springer, Berlin, Germany, 2010.

[17] B. David, R. Dowsley, and A. Nascimento, "Universally composable oblivious transfer based on a variant of LPN," in *Proceedings of the CANS 2014*, pp. 143–158, Springer.

[18] Z. Li, C. Ma, and H. Zhou, "Multi-key FHE for multi-bit messages," *Science China Information Sciences*, vol. 61, article 029101, pp. 1–3, 2018.

[19] Z. Li, S. D. Galbraith, and C. Ma, "Preventing Adaptive Key Recovery Attacks on the Gentry-Sahai-Waters Leveled Homomorphic Encryption Scheme," Cryptology ePrint Archive 2016/1146, https://eprint.iacr.org/2016/1146.

[20] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE," in *Proceedings of the IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS '11)*, pp. 97–106, Palm Springs, CA, USA, October 2011.

[21] B. M. David, A. C. A. Nascimento, and J. Müller-Quade, "Universally composable oblivious transfer from lossy encryption and the McEliece assumptions," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 7412, pp. 80–99, 2012.

[22] Z. Li, C. Ma, and D. Wang, "Leakage Resilient Leveled FHE on Multiple Bit Message," *IEEE Transactions on Big Data*, pp. 1-1.

[23] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *Proceedings of the FOCS*, pp. 136–145, IEEE Computer Society Press, 2001.

[24] R. Lindner and C. Peikert, "Better key sizes (and attacks) for LWE-based encryption," Cryptology ePrint Archive 2010/613, https://eprint.iacr.org/2010/613.

[25] C. Peikert, "Public-key cryptosystems from the worst-case shortest vector problem: extended abstract," in *Proceedings of the STOC 2009*, pp. 333–342, ACM Press.

[26] C. Peikert and B. Waters, "Lossy trapdoor functions and their applications," in *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC '08)*, pp. 187–196, ACM Press, 2008.

[27] R. Hiromasa, M. Abe, and T. Okamoto, "Packing messages and optimizing bootstrapping in GSW-FHE," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E99A, no. 1, pp. 73–82, 2016.

[28] S. Agrawal, D. Boneh, and X. Boyen, "Efficient lattice (H)IBE in the standard model," in *Advances in Cryptology—EUROCRYPT 2010*, vol. 6110 of *Lecture Notes in Computer Science*, pp. 553–572, Springer, Berlin, Germany, 2010.

[29] D. He, D. Wang, Q. Xie, and K. Chen, "Anonymous handover authentication protocol for mobile wireless networks with conditional privacy preservation," *Science China Information Sciences*, vol. 60, no. 5, Article ID 052104, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 3403456, 11 pages
https://doi.org/10.1155/2018/3403456

Research Article

Muscle Activity-Driven Green-Oriented Random Number Generation Mechanism to Secure WBSN Wearable Device Communications

Yuanlong Cao,[1] Guanghe Zhang,[1] Fanghua Liu,[1] Ilsun You,[2] Guanglou Zheng,[3] Oluwarotimi Williams Samuel,[4,5] and Shixiong Chen[4,5]

[1]*Jiangxi Normal University, Nanchang, China*
[2]*Department of Information Security Engineering, Soonchunhyang University, Asan, Republic of Korea*
[3]*Security Research Institute, Edith Cowan University, Perth WA 6027, Australia*
[4]*Chinese Academy of Sciences (CAS), Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen, China*
[5]*Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China*

Correspondence should be addressed to Ilsun You; ilsunu@gmail.com

Received 13 April 2018; Accepted 27 June 2018; Published 19 August 2018

Academic Editor: Ding Wang

Wireless body sensor networks (WBSNs) mostly consist of low-cost sensor nodes and implanted devices which generally have extremely limited capability of computations and energy capabilities. Hence, traditional security protocols and privacy enhancing technologies are not applicable to the WBSNs since their computations and cryptographic primitives are normally exceedingly complicated. Nowadays, mobile wearable and wireless muscle-computer interfaces have been integrated with the WBSN sensors for various applications such as rehabilitation, sports, entertainment, and healthcare. In this paper, we propose MGRNG, a novel muscle activity-driven green-oriented random number generation mechanism which uses the human muscle activity as green energy resource to generate random numbers (RNs). The RNs can be used to enhance the privacy of wearable device communications and secure WBSNs for rehabilitation purposes. The method was tested on 10 healthy subjects as well as 5 amputee subjects with 105 segments of simultaneously recorded surface electromyography signals from their forearm muscles. The proposed MGRNG requires only one second to generate a 128-bit RN, which is much more efficient when compared to the electrocardiography-based RN generation algorithms. Experimental results show that the RNs generated from human muscle activity signals can pass the entropy test and the NIST random test and thus can be used to secure the WBSN nodes.

1. Introduction

Over the last few years, the growing interest in the wireless body sensor network (WBSN) has resulted in thousands of peer-reviewed publications. Significant results in this area have enabled many medicine and healthcare applications. A WBSN interconnects tiny and wireless sensor nodes and devices worn on or implanted in the human body that have the capability to acquire physiological signals such as electrocardiography (ECG) [1, 2], electromyography (EMG) [3], and electroencephalography (EEG) [4], as well as data about the physical state of individuals which include walking, running, and seating [5–7]. A wide spectrum of WBSNs applications that include different kinds of wearable devices has been developed and applied for, e.g., physical fitness monitoring, and chronic diseases monitoring. Figure 1 illustrates the most typical usage scenarios for a WBSN-based rehabilitation system with wearable devices. In such a rehabilitation system, the EMG signals can be collected by wearable WBSN devices from an injured patient's upper arm, processed and then used to control an artificial limb in order to facilitate the injured or disabled people's daily works and lives.

Although the WBSN technologies bring obvious and attractive benefits to facilitate people's life activity, there are many concerns and challenges to be addressed. The first important concern of WBSN is related to security and privacy

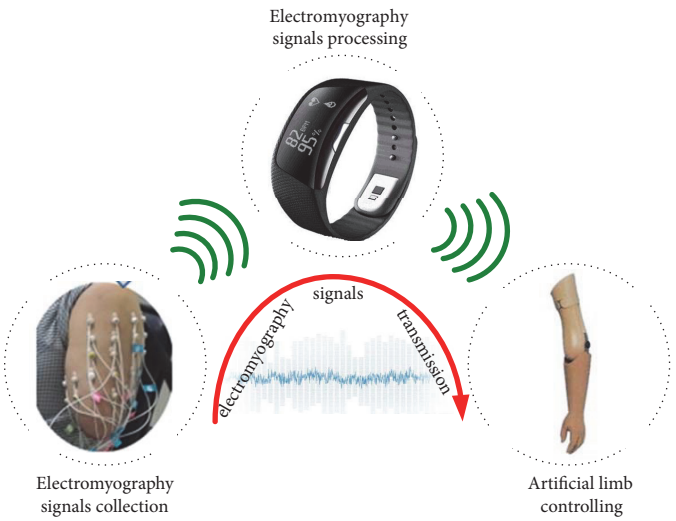


Figure 1: A WBSN-based rehabilitation system with the EMG signals collection, processing, and artificial limb controlling.

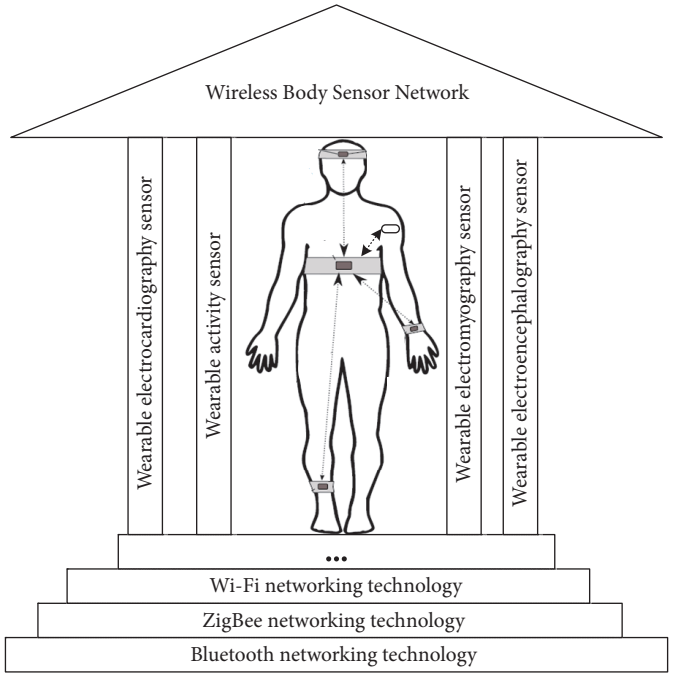


Figure 2: A basic structure of wireless body sensor networks.

challenges. A typical WBSN system consists of a number of wearable biomedical sensors or wireless muscle-computer interfaces worn on or even implanted in a single person; it also involves a base station which can be a wearable computer, a smart watch, or a smart phone. The base station is responsible for collecting physiological signals from a patient through wearable biomedical sensors placed on a human by using of Wi-Fi, ZigBee, and/or Bluetooth networking technologies (as shown in Figure 2). However, due to limited bandwidth resource and computing capabilities, wearable WBSN devices provide less security and authentication system (in comparison to traditional personal computers, laptops, and other computing devices). Consequently, we have reason to believe that attacks targeting wearable WBSN devices with the aim of malicious disruption or worse will rise. As numerous wearable devices make their way into the WBSNs with a much more open structure, they undoubtedly bring a WBSN system of security and privacy challenges.

Aside from the wearable devices, there will also be a rise in the number of targeted attacks focused on interrupting WBSN signal transmission or even stealing personal data. In a WBSN system, wearable devices generally have limited

storage and computation capabilities, they require to pair with other wearable electronic devices to perform most functions. In other word, a WBSN system is more like a tiny switching system [8, 9]. In such a system, a wearable WBSN device needs to communicate with other computing devices (e.g., a mobile phone and a wearable personal computer) via Wi-Fi, Bluetooth, or ZigBee connections for the data to be processed. During the communications between a wearable device and a computing device, these personal data should be properly protected from all forms of security threats posed by potential attackers, since the data normally contains sensitive health information [10–14]. The unauthorized usage of the data will breach the users' privacy and could be even harmful to their life.

The second concern of WBSN is the energy consumption problem. In the WBSN systems, a wearable device generally has extremely limited energy capacity and is frequently constrained by the volume of energy resource available in its battery. Moreover, a wearable device is not a standalone device as it needs to send the collected biomedical data to a computing device using Bluetooth, Wi-Fi, or ZigBee networking technology, which means that apart from the energy consumption of data collection an energy-constrained wearable device needs to spend a considerable amount of energy resource for the biomedical packet transmission. The high power consumption of wearable devices will limit the WBSN lifetime because if one or more wearable devices run out of battery, the whole WBSN system may fail to work. Therefore, a promising WBSN-based design has to take into consideration energy overhead and optimization policy [15].

In this paper, we propose a novel muscle activity-driven green-oriented random number generation mechanism (MGRNG) which uses the human muscle activity as green energy resource to generate random numbers (RNs). The RNs can be used to enhance the privacy of wearable device communications and secure WBSNs for rehabilitation purposes. To this end, we study a different type of physiological signal for the random number (RN) generation purpose, that is, the EMG signals recorded from the muscle(s) via a surface or intramuscular electrode(s) placed on the muscle(s) [3]. The method was tested on 10 healthy subjects as well as 5 amputee subjects with 105 segments of simultaneously recorded surface electromyography signals from their forearm muscles. The proposed MGRNG requires only one second to generate a 128-bit RN, which is much more efficient when compared to the electrocardiography-based RN generation algorithms.

The remainder of this paper is organized as follows. In Section 2, a problem statement with a clear description of the security and privacy issue in a WBSN system is presented in order to explain our motivation. Section 3 presents the EMG signal-based RN generation algorithm, describes the experimental settings and determines performance metrics in order to evaluate the proposed EMG-based RN generation scheme. Section 4 evaluates and analyzes the performance of the proposed solution. Section 5 discusses the limitations of the work and gives some interesting directions for future work. Section 6 concludes the paper.

2. Problem Statement

Nowadays, with the widespread use of WBSNs, the shapes and functionalities of WBSN devices are evolving dramatically. Among them, wearable devices are becoming even more important as sales of wearable devices continue to see year-over-year growth [17]. Typically, wearable WBSN devices facilitate the daily life and works. Unfortunately, wearable WBSN devices become a vulnerable and attractive target of most attacks because of the lack of security mechanisms [18]. Moreover, a wearable WBSN device becomes more and more lucrative target because it collects and shares an amount of sensitive personal data with third-parties [19]. Therefore, wearable WBSN devices bring much more threats to the WBSN systems and make them face more serious security and privacy challenges [20, 21].

In a WBSN system, data collected from the wearable biomedical sensors or other wearable devices is often transmitted to a mobile phone or other portable devices where it is stored and processed in order to provide real-time feedback for users in different areas of applications [22]. However, the personal data collected by wearable WBSN devices may be illegally used by potential attackers [23]. Figure 3 presents a scenario that the health information of individuals maliciously is manipulated by an attacker. As the figure shows, a wearable electromyography sensor transmits the collected personal electromyography data to a wearable watch for data processing and artificial limb controlling. However, the artificial limb can be out of control because of malicious manipulation caused by an attacker.

Apart from malicious manipulation, personal data and biomedical information collected and processed by connected wearable WBSN devices are also increasingly becoming a main target of attacks. A fresh survey has revealed that there is a rise in the number of targeted attacks focused on stealing personal data from the wearable WBSN devices connected to the computing devices [24]. Figure 4 presents a scenario that the health information of individuals is intercepted by potential attackers. Although the amount of personal data and biomedical information collected and processed in the WBSN systems grows daily, these data and information with extremely valuable can be easily captured by hackers and vanished into the black market due to the lack of security and authentication system in the WBSN systems [25, 26].

Considering the fact that health information of individuals is subject to interception and manipulation by potential attackers, the security of data at rest and in transit is a major challenge in the abovementioned WBSN applications. Existing traditional asymmetric security methods such as Rivest Shamir Adleman (RSA), Elliptic Curve Cryptography (ECC), and symmetric methods including Advanced Encryption Standard (AES) and Rivest Cipher 5 (RC5) are of high computational complexity. Hence, these algorithms cannot be directly applied to the WBAN systems since they have limited computing resources. To develop an efficient security mechanism for WBSN systems, previous studies have proposed the use of features extracted from physiological signals

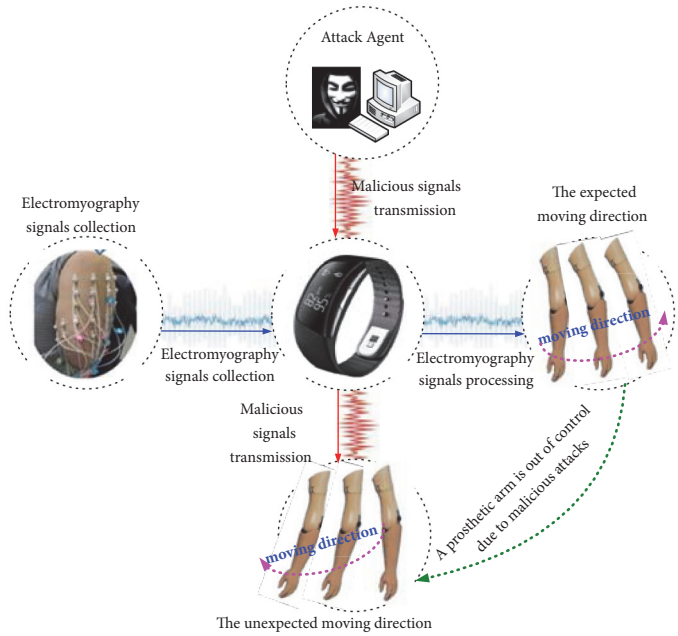


Figure 3: The health information of individuals is subject to manipulation by potential attackers: an example.

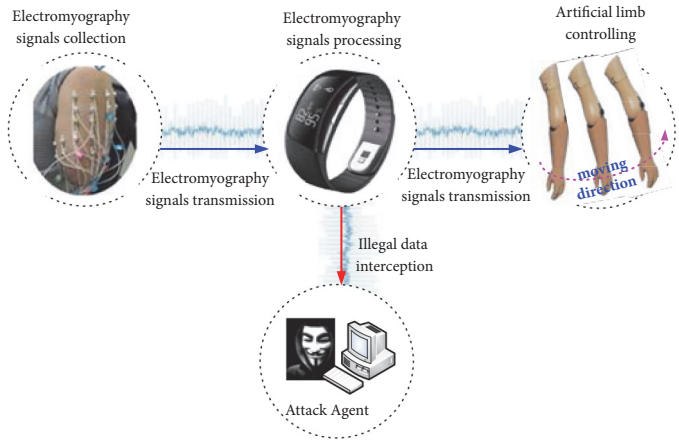


Figure 4: The health information of individuals is subject to interception by potential attackers: an example.

like ECG and EEG collected via wireless sensors and wearable devices for random number generation.

Currently, most of the random numbers generated from ECG features are based on the process of Interpulse Intervals (IPIs) extracted from the signal [16]. In previous studies, a method using IPIs to encode 128-bit random number sequence with reasonable randomness performance shows that if a single IPI is encoded into 4-bits, then at least 32 IPIs have to be extracted from the ECG signal. This implies that the WBSN sensor nodes need to successfully detect at least 33 consecutive heartbeats [16]. Considering that the normal sinus rhythm of an adult lies between 60 and 100 beats per minute (bpm), generating a 128-bit of random number sequence would require about 20 seconds. Hence, the IPI-based methods may not be suitable for real-time applications of WBSN systems since it requires a considerable amount of time.

Furthermore, recent WBSN efforts are devoted to generating random numbers for securing the data associated with WBSN devices, by exploiting the characteristics of EEG signals [4]. Although their research reports reveal that the EEG-based method is comparable to the existing ECG-based methods, it is important to note that the EEG-based random number generation method has a number of issues which may limit its application in real-time systems. These issues are as follows: (a) processing acquired EEG signals requires a relatively high computing resources because the signals have poor signal-to-noise ratio; (b) EEG signal acquisition would generally require precise deployment of dozens of electrodes around the head and the use of various gels, saline solutions,

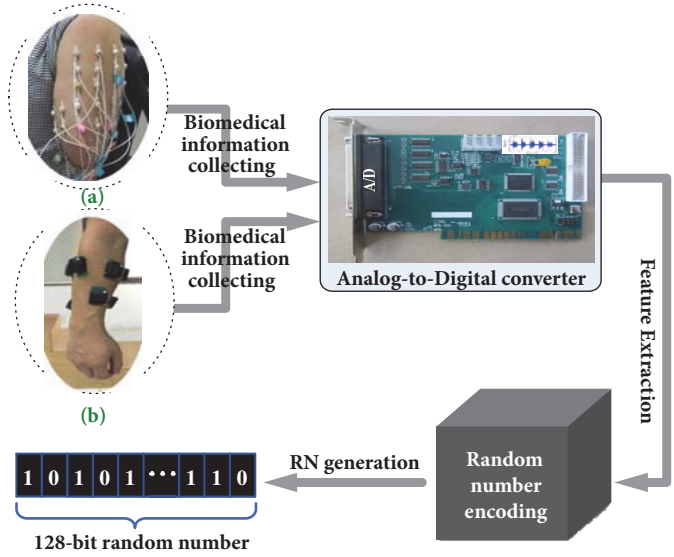


Figure 5: The architecture of MGRNG system.

and/or pastes to keep the electrodes in place, which would severely affect the daily life of the individual and make the user feel uncomfortable.

Compared to the abovementioned IPI-based methods, EMG signals sampled at the rate of 1000Hz can be used to generate a 128-bit random sequence with only one second. More importantly, in comparison to the EEG, EMG signals require less computing resources with a simplified data collection procedure. In this paper, we propose MGRNG, a novel random number generation algorithm which is based on features extracted from EMG recordings. The randomness of the 128-bit RNs was exhaustively examined by using the entropy metric. We believe that these EMG-based RNs can be further used to protect both data at rest and in transit in a WBSN system. More specifically, our solution makes important contributions in the following aspects:

(i) It introduces a green-oriented RN generation method which exploits the human muscle activities as energy resources to generate RNs and improve on wearable device communication by being more energy efficient.

(ii) It designs a fast RN generation scheme which generates a 128-bit random sequence in a timely fashion. The muscle activity-based RNs can be further used to enhance the security and privacy in WBSNs.

3. MGRNG Detail Design

This paper is devoted to exploring the possibility of resolving a critical security issue associated with data storage and transmission among the wearable devices and nodes in WBSN systems. To this end, we propose a novel security method dubbed as MGRNG. MGRN contributes to generating an EMG-based 128-bit RN for secure communication over WBSN systems. Figure 5 presents the architecture of MGRNG, which reveals how the EMG signals are collected (from the upper arm of an amputee subject and the forearm of a healthy subject), then processed, and encoded for RN generation.

3.1. EMG-Based RN Generation. In the MGRNG solution, there are five stages included to generate RN: (i) EMG signal collection, (ii) EMG signal sampling, (iii) EMG signal segmenting, (iv) feature extraction, and (v) RN generation. Figure 6 shows the technological processes of EMG signal-based RNs generation scheme in MGRNG. A detailed description of each of these five stages is presented as follows in order to help the readers understand easily:

(i) *EMG Signal Collection*: In MGRNG, the EMG signals are collected from an upper arm of an amputee subject and forearm of a healthy subject through wearable biomedical sensors placed on these subjects.

(ii) *EMG signal sampling*: The analog EMG signals are converted into digital EMG signals by using an Analog-to-Digital (A/D) converter, and every EMG signal is sampled at the rate of 1000Hz.

(iii) *EMG Signal Segmenting*: These collected EMG signals are divided into 128 segments in order to further used for 128-bit RNs generation.

(iv) *Feature Extraction*: There are many methods can be used for signal denoising and EMG signal extraction. In this study, we choose the wavelet analysis method for signal denoising and EMG signal extraction since as a mature technology, wavelet analysis has proven to be invaluable in signals analysis and processing. More details on the wavelet analysis method can be found in [27].

(v) *RN Generation*: Encoding these extracted EMG features generates a sequence of random numbers (a 128-bit RN) for the purpose of secure WBSN communications.

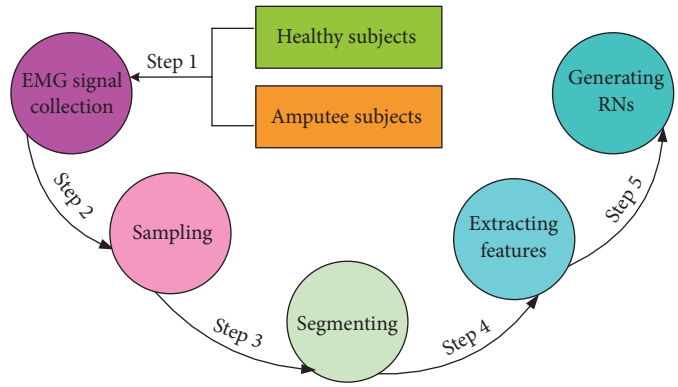


Figure 6: The technological processes of EMG signal-based RN generation.

Definition:
S_{list}: all subjects who are volunteers for EMG collection
WS_{list}: all wearable sensor used to acquire the EMG signals
ES_{list}: all the EMG signals collected from subjects
for $(i = 1, i \leq \text{count}(S_{list}), i++)$ **do**
  for $(j = 1, j \leq \text{count}(WS_{list}), j++)$ **do**
    execute EMG signal acquisition;
  end for
end for
for $(k = 1, k \leq \text{count}(ES_{list}), k++)$ **do**
  take the EMG signal sample;
  divide EMG signals into 128 segments;
  extract EMG features using wavelet analysis method;
  encode the EMG features to generate 128-bit RNs.
end for

Algorithm 1: EMG-based RN generation algorithm.

In this study, we use 256 EMG signals from each subject to generate a number of RNs which are of different lengths. Each EMG signal was encoded into different bits to result in RNs of length varying from 512 bits to 2816 bits. Then the RNs were divided into 128 bit segments. To this end, we developed an algorithm with MATLAB programming tool to detect and exclude the invalid data segments in the EMG recordings. Afterwards, the features were extracted from the preprocessed EMG signals and utilized to generate 128-bit RNs, and a total of 105 (15 subjects including 10 healthy subjects (HS) and 5 disabled subjects (DS), and each subject has 7 sensors) RNs were generated. The experimental protocols were approved by the Shenzhen Institutes of Advanced Technology Institutional Review Board, Chinese Academy of Sciences. All the subjects agreed and gave written informed consent as well as permission for the publication of their photographs and data for scientific and educational purposes. The pseudo code of the EMG-based RN generation algorithm is presented in Algorithm 1.

From Algorithm 1, it can be seen that the EMG-based RN generation method would at least have the following advantages: (i) an EMG is a very low-risk procedure, (ii) the EMG signals can be easily acquired from muscle activities, (iii) compared with most of the traditional pseudorandom number generation solution [7], the EMG-based RN generation method neither requires a random seed nor complex computation operations, and (iv) compared with the classic ECG/EEG-based number generation solutions [4, 16], the EMG-based RN generation method requires less computing resources and shorter time to generate RNs. For these reasons, we believe that the proposed EMG-based RN generation method will become a promising technology used to generate RNs for securing data confidentiality in WBSNs.

3.2. Performance Metric Determination. Considering the fact that the entropy is an important metric to measure and reflect the randomness and uncertainty of matter in a system, in this paper, we select entropy as the performance metric to measure and evaluate the randomness of the EMG-based RNs. We calculate the entropy values for these generated EMG-based RNs by using the following formula:

$$S = -K \times \sum_{i=1}^{N} P_i \ln\left(P_i\right), \tag{1}$$

where K is a constant, P_i is the uncertainty (probability) of the i^{th} RN, and N is the number of RNs.

Moreover, in order to further convince the effectiveness of these generated RNs, we calculate the mean entropy value and standard deviation of RNs generated from each EMG sensor placed on every subject. Assuming the obtained entropy values are $e_1, e_2, e_3, \cdots, e_m$, we calculate the mean value of the entropies by using the following formula:

$$\overline{E_M} = \frac{1}{M} \times \sum_{i=1}^{M} e_i, \tag{2}$$

where E_i is an entropy value, M is the number of entropy values, and $\overline{E_M}$ is the mean entropy of RNs generated from each EMG sensor.

Formula (2) is a general formula for the mean entropy value calculation. In order to reduce the computational complexity, we use an iterative method (see (3)) to calculate the mean entropy value:

$$\overline{E_{M+1}} = \frac{\overline{E_M} \times M + e_{M+1}}{M+1}, \tag{3}$$

We use the previous mean entropy $\overline{E_M}$ and the new entropy value e_{M+1} to calculate the current mean entropy value $\overline{E_{M+1}}$. This means that the current mean entropy is updated according to the newly recorded entropy values.

Similarly, the general formula for calculating the standard deviation (SD) is presented in the following:

$$\sigma_M = \sqrt{\frac{\sum_{i=1}^{M} \left(e_i - \overline{E_M}\right)^2}{M-1}}, \tag{4}$$

Equation (4) is a general formula for calculating the SD value. In order to reduce the computational complexity, we also calculate the SD value utilizing an iterative method presented in the following formula:

$$\sigma_{M+1} = \sqrt{\frac{\sigma_M^2 \times (M-1)}{M} + \frac{\left(e_{M+1} - \overline{E_M}\right)^2}{M+1}}, \tag{5}$$

By using (5), we can calculate the new SD σ_{M+1} using only four variables: the previous SD σ_M, the previous mean entropy value $\overline{E_M}$, the current entropy value e_{M+1}, and the previous records of entropy M.

The mean entropy and the entropy SD values obtained from (3) and (5) can be used as references for evaluating the randomness and effectiveness of these generated EMG-based RNs. Furthermore, we also use 15 tests provided by the US National Institute of Standards and Technology (NIST) to validate the performance of RNs [7]. Specifically, we implement the five commonly used NIST tests (see Table 1) as part of effort to further examine the randomness of our proposed method. In other words, the five NIST tests have been used in this work for evaluating the randomness of the 128 bit RNs and the pass rates have been defined as the ratio of the numbers of p value greater than 0.01 divided by the total generated RNs.

It is worthy to note that the NIST suite consists of 15 different tests [3–5], and some of them require the length of the RNs to be no less than 1,000,000 bits. In this study, we only focus on evaluating an EMG-based RN with a length of only 128 bits; thus we just utilize the five commonly used NIST tests (namely, F Test, B Test, R Test, L Test, and A Test) to verify the randomness of the generated RNs. Moreover, interested readers can refer to our previous work [16] for more detail information on how the NIST tests are conducted for the generated RNs.

Table 1: The five most commonly used NIST tests.

Test cases	Descriptions
F Test	The frequency test
B Test	The frequency test block
R Test	The runs test
L Test	The longest runs ones block test
A Test	The approximate entropy test

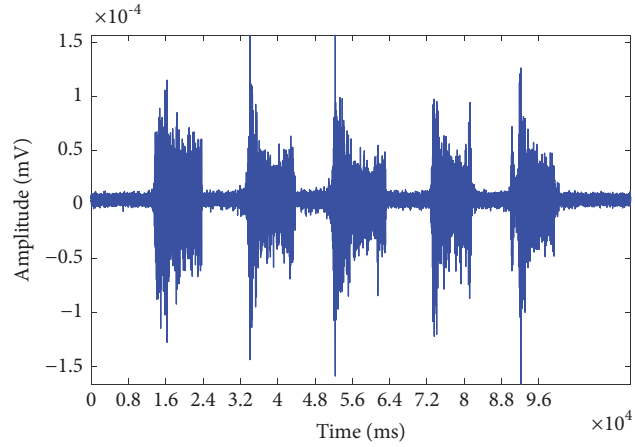


Figure 7: The EMG amplitudes between time $t = 0$ ms and $t = 10$ ms.

4. Results and Analysis

In order to generate 128 bit EMG-based RNs and then evaluate the randomness of them, we pull the EMG samples from total 15 subjects including 10 healthy subjects (HS) and 5 disabled subjects (DS); each subject has 7 wearable biomedical sensors to gather the EMG signals. Figure 7 presents the EMG amplitudes collected from one of these subjects (only EMG amplitudes between time $t = 0$ ms and $t = 10$ ms are illustrated in the figure in order to better show the results).

After EMG feature extraction, feature encoding and RN generation, we calculate the entropy values for these generated RNs by using (1). From the obtained entropy values shown in Table 2, it was observed that the entropy values of the generated RNs varied from 0.9887 to 0.9998 among the 7 EMG sensors for a representative healthy subject. Meanwhile, for an arbitrarily selected amputee, an entropy value that ranged between 0.9914 and 1.0000 was recorded for the generated RNs based on EMG recordings from the seven sensors.

Table 2: The entropy of RNs generated from EMG signals extracted from seven sensors (a healthy subject vs. an amputee subject).

No.	A healthy subject	An amputee subject
1	0.9887	0.9937
2	0.9998	1.0000
3	0.9914	0.9993
4	0.9998	0.9972
5	0.9914	0.9956
6	0.9984	0.9914
7	0.9972	0.9937

Table 3: The entropy of each sensor of health and disabled subjects.

No.	MEHS	VHS	MEAS	VAS
1	0.9967	1.88×10^{-5}	0.9968	6.80×10^{-6}
2	0.9861	3.02×10^{-5}	0.9913	1.70×10^{-5}
3	0.9959	1.51×10^{-4}	0.9964	7.10×10^{-6}
4	0.9948	1.03×10^{-5}	0.9962	5.28×10^{-5}
5	0.9996	2.80×10^{-6}	0.9976	9.13×10^{-5}
6	0.9952	9.40×10^{-6}	0.9967	1.61×10^{-3}
7	0.9973	1.74×10^{-5}	0.9981	1.50×10^{-6}

Table 4: The entropy of each sensor of health and disabled subjects.

Name	Mean of entropy	Variance of entropy
Health subjects (our Best)	0.9996	2.80×10^{-6}
Amputee subjects (our Best)	0.9981	1.50×10^{-6}
Subjects with myocardial infarction [6]	0.9902	2.31×10^{-6}
Subjects with other CVD [6]	0.9899	2.96×10^{-6}
Healthy subjects [6]	0.9893	3.46×10^{-6}

Table 3 presents the average entropy of RNs generated from each EMG sensor (i.e., from sensor 1 to sensor 7) placed on 10 healthy and 5 amputee subjects. It shows that the mean entropy of RNs generated from 10 healthy subjects (MEHS) varies from 0.9861 to 0.9996. Meanwhile, the mean entropy of RNs generated from 5 amputee subjects (MEAS) varies from 0.9913 to 0.9981. The overall average entropy of both categories of subjects per sensor varied from 0.9887 to 0.9986 (i.e., from sensor 1 to sensor 7). The variance of health subjects (VHS) varied from 2.80×10^{-6} to 1.51×10^{-4}, and the variance of amputee subjects (VAS) varied from 1.50×10^{-6} to 1.61×10^{-3}. Moreover, by comparison with the mean entropy values, we found that the mean of entropy of RNs generated from EMG signals in our proposed approach is better than binary previously study [6], as shown in Table 4.

Additionally, we implemented the five most commonly used NIST tests (F Test, B Test, R Test, L Test, and A Test) as part of effort to clarify the randomness of our proposed method. We found out through experiment that, the pass rates of the five NIST tests varied from 0.9857 to 1.0000 for the healthy subjects and from 0.9714 to 1.0000 for the amputee subjects, as shown in Figures 8 and 9, respectively.

Through analyzing the results, we can observe that the entropy values of RNs generated from the representative healthy and amputee subjects shown in Table 2 are observed to be close to 1.00, which reflects a perfect randomness as well as a perfect performance. Also, similar performance can be observed when the entropy of the entire 105 RNs is computed. By comparing the RNs generated from the sensors deployed at different positions of the body, the obtained entropy results show that there is no distinctive difference between RNs from any two sensors. This implies that the EMG sensor position has no effect on the randomness of generated RNs.

Furthermore, we investigate the randomness of the RNs generated based on the features extracted from EMG recordings of the healthy and amputee subjects and then compare the results obtained for both categories of subjects. The rationale behind comparing the RNs of healthy and amputee subjects is because WBSN systems have been widely applied for both the healthcare and the sports training. In healthcare, rehabilitation devices such as the prostheses are now commonly available to help amputees regain their arm functions. And the EMG signal patterns obtained from the amputated arm have been reported to be different from that obtained from the intact arm [28–30]. In this regard, we compare the randomness of the RNs generated from the healthy and amputated subjects to see if the variations in EMG signal patterns between both categories of subjects would be different. From the comparison, it can be observed that there is no distinctive difference between the RNs of the

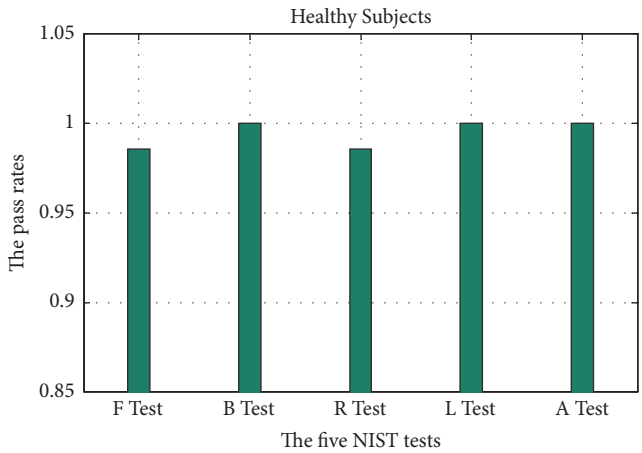


Figure 8: The pass rates of 128-bit RNs using five NIST tests (RNs generated from EMG of healthy subjects).

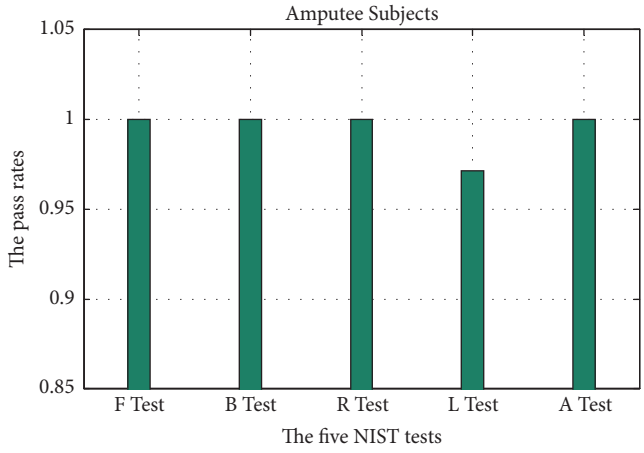


Figure 9: The pass rates of 128-bit RNs using five NIST tests (RNs generated from EMG of disabled subjects).

healthy and amputee subjects. We can see from this result that our proposed method is robust and could be used to secure the data in WBSN systems regardless of the health status of the user (whether the user is healthy or not).

We also compare the pass rates of the RNs generated by the proposed EMG-based RN generation method with those resulted from the Interpulse Intervals- (IPIs-) based RN generation method [16] (the pass rates of the 128-bit RNs generated from IPIs of 20 healthy subjects have been tested). In order to make the comparison more reasonable, the pass rates of RNs from healthy subjects generated by the EMG-based RN generation method and the IPIs-based RN generation method are illustrated only, as shown in Figure 10. As the figure shows, the proposed EMG-based RN generation method outperforms the IPIs-based RN generation method in all cases in terms of RN pass rate. The subfigure in Figure 10 shows the cumulative average pass rates of the two methods. As the subfigure shows, the EMG-based RN generation method achieves a better pass rate in comparison with the IPIs-based RN generation method. More specifically, the average pass rate of the EMG-based RN generation method is about 3.54% higher than that of the IPIs-based RN generation method.

5. Discussion

In Section 4, we demonstrated how the proposed EMG-based RNs generation method is robust, we also investigated the randomness of the EMG-based RNs, and we believe that the EMG-based RNs can be used to enhance the privacy of wearable device communications and secure the data in WBSN systems. In this section, we discuss the limitations of the work and explore some potential directions for future research:

(i) Despite the promising results achieved by the proposed EMG-based RNs generation method, it may not be applicable to persons with neuromuscular disorder, who lack the ability to provide EMG signals from which RNs could be generated.

(ii) It is unknown whether the proposed EMG-based RNs generation method would be robust to variation in muscle contraction level. Therefore, we plan to investigate this issue in our future study.

(iii) It is unknown what the capabilities of the EMG-based RNs are allowed to an attacker. Our future work will study the existing threat models [31] and further define a proper threat model to assess the security capabilities of a WBSN system with the EMG-based RNs.

(iv) It is also unknown whether the proposed EMG-based RNs generation mechanism can be used in the promising anonymous mutual authentication protocols for wearable sensors. This is a very interesting topic and will be carried on in our future research.

(v) As discussed in the previous section, we consider five different NIST tests that could be used to validate the randomness of RNs of 128-bit length. However, as a secure RN generator, it should be better to pass all the 15 NIST tests. We will extend our proposal and evaluate the randomness under those test cases.

(vi) In addition, we believe that a WBSN system with the promising collaboration computing and multipathing technologies [32–36] is an interesting topic worth further investigation. In our future work, we will also apply these promising technologies to optimize the performance of the WBSN systems.

6. Conclusion

In this paper, we present a novel EMG-based RN generation method which uses the human muscle activity as green energy resource to generate 128 bit RNs for securing data communication in a WBSN system. Compared to the previously proposed methods that used ECG features to generate RNs for securing the data in WBSN systems [16], our EMG-based approach could generate random numbers with comparable performance and high speed of generation

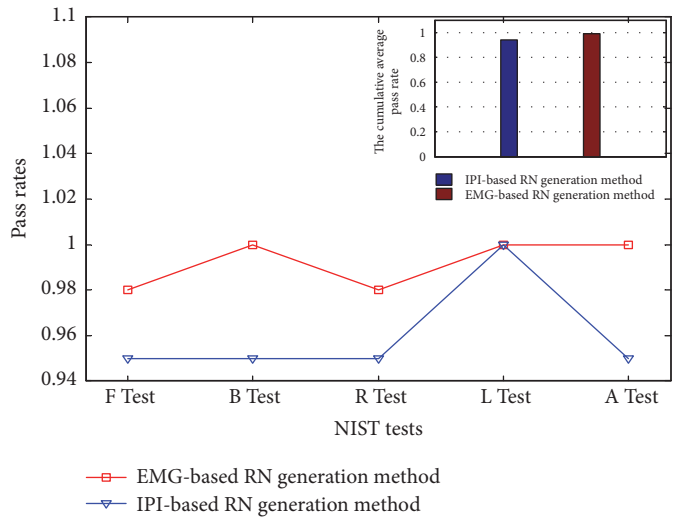


Figure 10: Pass rates of the EMG-based RN generation method as compare to the IPI-based RN generation method [16].

(128 bit per second). In addition, the proposed surface EMG-based RNs generation method can be used for real-time applications for a WBSN system since the signal from a single EMG sensor would be sufficient to provide the required RNs based on our experimental results. In addition, unlike the EEG-based RNs generation method that is sensitive to sensor position as placed on the scalp of a subject, our proposed EMG-based approach is insensitive to sensor location on the muscle, which is an advantage over the recently proposed EEG-based approach [4]. The results from five NIST statistical tests reveal that the RNs generated by our method can potentially be used as authentication identifiers or encryption keys for securing WBSNs.

Data Availability

Some data were omitted due to a confidentiality agreement between the research team and the Shenzhen Institutes of Advanced Technology Institutional Review Board, Chinese Academy of Sciences.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The first two authors contributed equally to this work and share the first authorship.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant no. 61562043 and by the Soonchunhyang University Research Fund. The authors would like to thank the members of the School of Computer and Information Engineering, Jiangxi Normal University, as well as the members of the Research Center for Neural Engineering, Institute of Biomedical and Engineering, Shenzhen Institutes of Advanced Technology, and Chinese Academy of Sciences for their assistance in the experiments.

References

[1] C. C. Y. Poon, Y.-T. Zhang, and S.-D. Bao, "A novel biometrics method to secure wireless body area sensor networks for telemedicine and M-health," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 73–81, 2006.

[2] K. K. Venkatasubramanian, A. Banerjee, and S. K. S. Gupta, "EKG-based key agreement in body sensor networks," in *Proceedings of the 27th IEEE Conference on Computer Communications (IEEE INFOCOM Workshops)*, pp. 1–6, IEEE Xplore, Phoenix, AZ, USA, April 2008.

[3] G. Zhang, O. Samuel, F. Liu et al., "Electromyogram-Based Method to Secure Wireless Body Sensor Networks for Rehabilitation Systems," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC17)*, pp. 1246–1249, JeJu Island, S. Korea, 2017.

[4] J. F. Valenzuela-Valdes, M. A. Lopez, P. Padilla, J. L. Padilla, and J. Minguillon, "Human Neuro-Activity for Securing Body Area Networks: Application of Brain-Computer Interfaces to People-Centric Internet of Things," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 62–67, 2017.

[5] Y.-L. Zheng, X.-R. Ding, C. C. Y. Poon et al., "Unobtrusive sensing and wearable devices for health informatics," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1538–1554, 2014.

[6] O. W. Samuel, X. Li, Y. Geng et al., "Resolving the adverse impact of mobility on myoelectric pattern recognition in upper-limb multifunctional prostheses," *Computers in Biology and Medicine*, vol. 90, pp. 76–87, 2017.

[7] A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST 800-22, 2001.

[8] Y. Zhang, P. Shi, C.-C. Lim, H. Zhu, J. Hu, and Y. Zeng, "Chaotification of a class of linear switching systems based on a Shilnikov criterion," *Journal of The Franklin Institute*, vol. 354, no. 13, pp. 5519–5536, 2017.

[9] Y. Zhang, X. Liu, H. Zhang, and C. Jia, "Constructing chaotic systems from a class of switching systems," *International Journal of Bifurcation and Chaos*, vol. 28, no. 2, 1850032, 9 pages, 2018.

[10] K. K. Venkatasubramanian, A. Banerjee, and S. K. S. Gupta, "PSKA: usable and secure key agreement scheme for body area networks," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 1, pp. 60–68, 2010.

[11] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless Communications Magazine*, vol. 17, no. 1, pp. 80–88, 2010.

[12] M. Li, W. J. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *IEEE Wireless Communications Magazine*, vol. 17, no. 1, pp. 51–58, 2010.

[13] *Summary of Health Insurance Probability and Accountability Act (HIPAA), U.S. Dept. of Health and Human Services, Washington, DC, 2003*.

[14] The European Parliament and the Council of the European Union, "Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector, Official Journal of the European Union, vol.L201, pp.37-47, 2002.

[15] F. Song, D. Huang, H. Zhou, H. Zhang, and I. You, "An Optimization-Based Scheme for Efficient Virtual Machine Placement," *International Journal of Parallel Programming*, vol. 42, no. 5, pp. 853–872, 2014.

[16] G.-H. Zhang, C. C. Y. Poon, and Y.-T. Zhang, "Analysis of using interpulse intervals to generate 128-bit biometric random binary sequences for securing wireless body sensor networks," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 1, pp. 176–182, 2012.

[17] S. Seneviratne, Y. Hu, T. Nguyen et al., "A Survey of Wearable Devices and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2573–2620, 2017.

[18] O. Arias, J. Wurm, K. Hoang, and Y. Jin, "Privacy and security in internet of things and wearable devices," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 1, no. 2, pp. 99–109, 2015.

[19] X. Fafoutis, L. Marchegiani, G. Z. Papadopoulos, R. Piechocki, T. Tryfonas, and G. Oikonomou, "Privacy leakage of physical activity levels in wireless embedded wearable systems," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 136–140, 2017.

[20] Y. Shi, X. Wang, and H. Fan, "Light-weight white-box encryption scheme with random padding for wearable consumer electronic devices," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 1, pp. 44–52, 2017.

[21] S. Wang, R. Bie, F. Zhao, N. Zhang, X. Cheng, and H.-A. Choi, "Security in wearable communications," *IEEE Network*, vol. 30, no. 5, pp. 61–67, 2016.

[22] F. Song, Y. Zhou, Y. Wang, T. Zhao, I. You, and H. Zhang, "Smart collaborative distribution for privacy enhancement in moving target defense," *Information Sciences*, 2018.

[23] J. Liu and W. Sun, "Smart Attacks against Intelligent Wearables in People-Centric Internet of Things," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 44–49, 2016.

[24] Z. Li, Q. Pei, I. Markwood, Y. Liu, and H. Zhu, "Secret key establishment via RSS trajectory matching between wearable devices," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 802–817, 2018.

[25] A. K. Das, M. Wazid, N. Kumar, M. K. Khan, K. R. Choo, and Y. Park, "Design of Secure and Lightweight Authentication Protocol for Wearable Devices Environment," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 4, pp. 1310–1322, 2018.

[26] F. P. Diez, D. S. Touceda, J. M. Sierra Cámara, and S. Zeadally, "Toward self-authenticable wearable devices," *IEEE Wireless Communications Magazine*, vol. 22, no. 1, pp. 36–43, 2015.

[27] MATLAB Wavelet Toolbox User's Guide. https://www.mathworks.com/help/wavelet/examples.html.

[28] G. Zheng, G. Fang, R. Shankaran et al., "Multiple ECG Fiducial Points-Based Random Binary Sequence Generation for Securing Wireless Body Area Networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 655–663, 2017.

[29] M. J. Cler and C. E. Stepp, "Discrete Versus Continuous Mapping of Facial Electromyography for Human-Machine Interface Control: Performance and Training Effects," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 572–580, 2015.

[30] D. Brunelli, E. Farella, D. Giovanelli, B. Milosevic, and I. Minakov, "Design considerations for wireless acquisition of multichannel sEMG signals in prosthetic hand control," *IEEE Sensors Journal*, vol. 16, no. 23, pp. 8338–8347, 2016.

[31] 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.

[32] F. Song, Y.-T. Zhou, K. Kong, Q. Zheng, I. You, and H.-K. Zhang, "Smart collaborative connection management for identifier-based network," *IEEE Access*, vol. 5, pp. 7936–7949, 2017.

[33] Z. Ai, Y. Zhou, and F. Song, "A Smart Collaborative Routing Protocol for Reliable Data Diffusion in IoT Scenarios," *Sensors*, vol. 18, no. 6, p. 1926, 2018.

[34] Y. Cao, F. Song, Q. Liu, M. Huang, H. Wang, and I. You, "A LDDoS-Aware energy-efficient multipathing scheme for mobile cloud computing systems," *IEEE Access*, vol. 5, pp. 21862–21872, 2017.

[35] F. Song, Z. Ai, J. Li et al., "Smart Collaborative Caching for Information-Centric IoT in Fog Computing," *Sensors*, vol. 17, no. 11, p. 2512, 2017.

[36] Z. Ai, Y. Liu, F. Song, and H. Zhang, "A Smart Collaborative Charging Algorithm for Mobile Power Distribution in 5G Networks," *IEEE Access*, vol. 6, pp. 28668–28679, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 4028196, 7 pages
https://doi.org/10.1155/2018/4028196

Research Article

Identity-Based Fast Authentication Scheme for Smart Mobile Devices in Body Area Networks

Chen Wang,[1,2] **Wenying Zheng**,[3] **Sai Ji**,[4] **Qi Liu**,[4] **and Anxi Wang**[4]

[1] *The School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China*
[2] *Guangxi Key Laboratory of Cryptography and Information Security, Guilin, China*
[3] *The School of Applied Meteorology, Nanjing University of Information Science & Technology, Nanjing 210044, China*
[4] *The Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing 210044, China*

Correspondence should be addressed to Wenying Zheng; zhengwy0501@126.com

Received 11 April 2018; Revised 31 May 2018; Accepted 20 June 2018; Published 5 August 2018

Academic Editor: Ding Wang

Smart mobile devices are one of the core components of the wireless body area networks (WBANs). These devices shoulder the important task of collecting, integrating, and transmitting medical data. When a personal computer collects information from these devices, it needs to authenticate the identity of them. Some effective schemes have been put forward to the device authentication in WBANs. However, few researchers have studied the WBANs device authentication in emergency situations. In this paper, we present a novel system named emergency medical system without the assistance of doctors. Based on the system, we propose an identity-based fast authentication scheme for smart mobile devices in WBANs. The scheme can shorten the time of device authentication in an emergency to achieve fast authentication. The analysis of this paper proves the security and efficiency of the proposed scheme.

1. Introduction

Nowadays, people's quality of life is improving with the development of society and technology [1]. People's pursuit of happy life has a broader definition. With the continuous improvement of medical level, the phenomenon of aging has appeared in many developed and developing countries. People are in urgent need of a more complete medical system to ensure health and safety. Real time health monitoring is needed to prevent the possibility of chronic diseases and also for emergency treatment of sudden diseases.

Wireless body area network (WBAN) is a network composed of sensor nodes, personal terminals, and medical cloud platforms [2–6]. WBANs can be used to monitor user signs, feed back on real-time data, provide corresponding treatment plan, and make relevant emergency measures. WBANs can not only monitor medical information and vital signs such as body temperature, pulse, and blood pressure through the attachment of sensor nodes, but also inject drugs with the help of embedded actuators to achieve long-term treatment, remote treatment, and emergency treatment. Some researchers consider that WBANs consist of three parties, including patients, doctors, and cloud servers [7–10]. In fact, in some emergency cases, the participation of doctors may not be able to participate in the treatment timely. These emergency cases include falls, myocardial infarction, and stroke. If all the treatments have to wait for the doctor to confirm, the best treatment period might be missed. In this paper, we think that WBAN is mainly composed of three parts: smart mobile devices (SMDs), personal digital assistants (PDAs), and remote cloud servers (RCSs) [11, 12]. An SMD is a portable sensor or actuator that has certain computing power and can perceive the specific information of the outside world. SMDs are indispensable parts of WBANs. SMDs can be utilized not only as a channel for WBANs to perceive external information, but also as a means for WBANs to intervene in the outside world. PDA can be a kind of mobile computing terminal with personal computers, smart phones, and so on. PDA is responsible for receiving messages collected by SMD or issuing commands to SMD. The RCS is responsible for storing and analyzing medical data and feedback treatment recommendations [13, 14].

Motivation of This Paper. Most of the existing schemes for WBAN do not consider how to implement emergency treatment without doctors' participation, so as to alleviate the sudden exacerbation of the patient's disease. The existing authentication scheme cannot be efficient and fast for such situations.

Our Contributions. The contributions of this paper can be concluded as the following three points. We first discuss a special case of how to perform treatment when a patient meets a sudden illness. The case is taken as the environment of the proposed system and scheme. We then propose an emergency medical system without the assistance of medical staff. Based on the system, we present the novel scheme to achieve fast authentication for smart mobile devices in WBANs. The detailed contributions are listed as follows.

(i) **A special case is discussed.** When the monitored object is suffering from sudden onset of myocardial infarction and stroke, WBANs are needed for emergency treatment. Under such circumstances, the traditional three-party system of patients, doctors, and medical cloud is no longer applicable. This paper discusses the particularity of this case and further studies on it.

(ii) **An emergency medical system has been proposed.** Based on the discussed case, we provide a system design that can be applied to this case. The novel emergency medical system is mainly composed of smart mobile devices, personal digital assistants, and remote cloud servers.

(iii) **An identity-based fast authentication scheme is proposed.** Finally, we propose an identity-based fast authentication scheme for smart mobile devices in WBANs. The scheme can quickly realize the identity authentication of a device and provide a reliable precondition for further encrypted data transmission of the system.

Organizations. The remainder of this paper is organized as follows. Section 2 presents some related works. Section 3 illustrates some preliminaries of this paper, including bilinear pairing, system model, and system components. Section 4 shows the security models of the novel authentication scheme. Section 5 presents the proposed scheme in detail. Section 6 states the security analysis of the proposed scheme. Section 7 presents the performance analysis of the scheme with simulations on PBC. Finally, the conclusions are drawn in Section 8.

2. Related Works

Many researchers have studied the authentication of smart mobile devices in WBANs.

Wang et al. [15] present an overview of attacks, principle, and solutions on the anonymity of two-factor authentication schemes. To improve the current schemes from being stuck with the security-usability tension, the scheme proposed by Wang et al. [16] can resolve the various issues arising from user corruption and server compromise.

Chiou et al. [17] propose a scheme which guarantees anonymity, unlinkability, and message authentication for uses. The proposed scheme also allows patients to directly and remotely consult with doctors in a safe way.

Li et al. [18] present participant authentication in mobile emergency medical care systems for patients supervision. They propose a secure cloud-assisted architecture for accessing and monitoring health in WBANs. Chaotic maps based authentication and key agreement mechanisms are utilized to provide data security and mutual authentication. Based on the proposed scheme, Li et al. [19] design another dynamic identity and chaotic maps based authentication scheme and a secure data protection approach to prevent illegal intrusions for medical systems. They also propose an improved secure authentication and data encryption scheme for the smart devices in medical systems in [20].

Li et al. [21] propose an anonymous mutual authentication for centralized two-hop WBAN. The scheme allows sensor nodes attached to the patient's body to authenticate with the local server/hub node.

Das et al. [22] find that some existing schemes are still vulnerable to privileged-insider attack. So they present a smartcard-based anonymous user authentication scheme for medical systems to be secure against possible known attacks.

To achieve secure and authorized communication, a symmetric key based authentication protocol is designed for medical system by Srinivas et al. [23]. They claimed that the results show that their scheme reaches the level of security requirements and has suitable cost for applications in medical environment.

Some researchers achieve security authentication in WBANs with some novel technologies. For instance, Hayajneh et al. [24] propose a scheme based on the Rabin authentication algorithm. They modify the algorithm to improve its signature signing process for delay-sensitive applications in WBANs. Park et al. [25] propose a selective group authentication scheme using Shamir's threshold technique. They prove that their scheme can achieve efficient user authentication and conditional access authority for devices in medical systems.

Mohit et al. [26] achieve mutual authentication between healthcare center, cloud server, and patients, which can support patient anonymity and resist strong security attacks such as nonrepudiation and confidentiality of data.

Li et al. propose a scheme to resist Denial of Service (DoS) attack [27]. To further solve problems in WBANs, Li et al. [28] provide three protocols for different tiers. The three protocols allow the anonymous authentication among mobile users, controller nodes, and the medical server.

3. Preliminaries

Here are some preliminaries provided for the proposed scheme.

3.1. Bilinear Pairing. Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be two groups of prime order q. Let $\mathbb{G}_1$ be an additive group and $\mathbb{G}_2$ be a multiplicative group. e is set to be as a mapping on $(\mathbb{G}_1, \mathbb{G}_2) : \mathbb{G}_1^2 \longrightarrow \mathbb{G}_2$. The cryptographic bilinear map e satisfies the following properties.

Bilinearity. $e(aP, bQ) = e(P, Q)^{ab}$ for all $P, Q \in \mathbb{G}_1$ and $a, b \in Z_q^*$. This can be expressed in the following manner. For $P, Q, R \in \mathbb{G}_1$, $e(P + Q, R) = e(P, R)e(Q, R)$.

Nondegeneracy. If P is a generator of $\mathbb{G}_1$, then $e(P, P)$ is a generator of $\mathbb{G}_2$. In other words, $e(P, P) \neq 1$.

Computability. e is efficiently computable.

3.2. System Model. In this subsection, we provide the introduction of the novel system in this paper. The system is named as emergency medical system. Figure 1 shows a schematic diagram of the system. The system is composed of sensors, actuators, PDA, and the cloud. Sensor nodes are responsible for collecting medical information in the WBAN. All data from the sensor nodes are compiled by PDA. The cloud will receive a summary of the information that PDA sends to it. According to cloud analysis of the current data and comparison of historical data, a treatment plan is chosen. The treatment plan will be directly sent to various actuators in a concise way, and the whole treatment plan will be sent to PDA.

3.3. System Components. The components of the system include smart mobile devices (SMDs), personal digital assistants (PDAs), and remote cloud servers (RCSs). The three main components are introduced as follows.

Smart Mobile Devices (SMDs). SMDs are sensors or actuators with certain computing ability in WBANs. Sensors are responsible for perceiving vital signs of patients. The important user medical data collected is transmitted by sensors in some specific form. The actuators are responsible for specific treatment operations after receiving instructions, such as injection of adrenaline and electric shock.

Personal Digital Assistants (PDAs). PDAs are personal computers or smart phone taken by the patient. As a link between smart devices and cloud servers, PDA is responsible for transmitting SMD's collected information and RCS's instructions.

Remote Cloud Servers (RCSs). The RCS is often a group of distributed computers with super computing power and large storage space. For ease of interpretation, we usually think that RCS's computing power and storage space are infinite.

4. Security Model

In this section, the security model of this paper is provided. Note that the key generation center (KGC) utilized in this paper is considered as a trusted third party to generate some system parameters [29–32].

4.1. A Forged SMD. We assume that a forged SMD may try to send the wrong message with the legal identity of the original SMD. Once such behavior is successful, it will be very dangerous for patients in the medical system. For example, a patient has no stroke, and the node passes the authentication and sends a stroke message to PDA, which could lead to the final error diagnosis of the patient and the treatment of the patient with the wrong medicine.

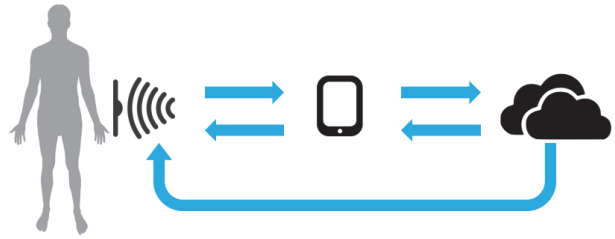


Figure 1: A schematic diagram of the emergency medical system.

4.2. Men-in-the-Middle Attack. Men-in-the-Middle (MITM) attack refers to the attack that the attacker intercepts the message and attempted to tamper with the medical message. This kind of attack will cause the original information to be destroyed, which leads the system to be unable to pass the authentication of SMD. We assume that an MITM attacker have the ability to block the message and implement all necessary calculations.

4.3. Replay Attack. Replay attack means that the attacker collects authentication messages sent before SMD and sends to PDA, trying to pass the authentication by PDA. This attack uses a message that has been authenticated. If the scheme is not well designed, the old authentication message is likely to be used by malicious users to achieve their goals. We assume that the attacker have the ability to obtain the historical authentication message and resubmit it.

5. Our Proposed Authentication Scheme

The proposed scheme is introduced detailedly in this section.

5.1. Overview of the Scheme. The whole scheme consists of three main phases: device fast authentication, secure message transmission, and secure instruction distribution. The whole scheme is shown in Figure 2. The circles in Figure 2 represent SMDs, including sensors and actuators, and the rectangle represents a PDA. A certain amount of SMDs are deployed on the patient. When a sudden illness occurs, one or more SMDs will monitor the change of corresponding parameters and integrate medical data information. Subsequently, SMDs need to prove identities and transmit encrypted information to the PDA. Then, the PDA needs to transmit the message to the RCS. These steps are the device fast authentication and secure message transmission that we mentioned earlier. RCS analyzes the current data with its powerful computing power and compares the data with the stored historical data. After a corresponding treatment decision is formulated, the RCS reaches the treatment instruction by secure instruction distribution phase. In this paper, we focus on the method of device fast authentication. We will provide some feasible solutions of the other phases for reference.

5.2. Device Fast Authentication. This authentication method is the main innovation of our paper. The detailed exposition will be carried out in this subsection. The device fast authentication consists of three algorithms: registration,

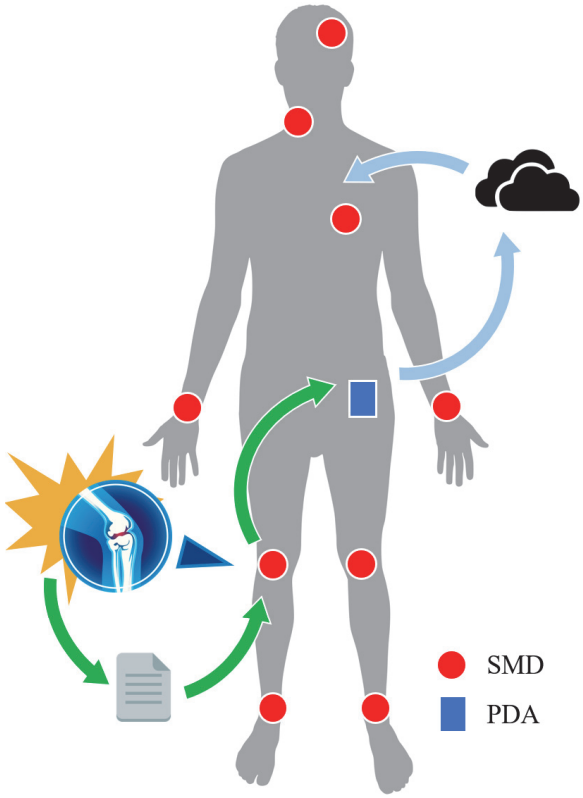


Figure 2: Illustration of the proposed scheme.

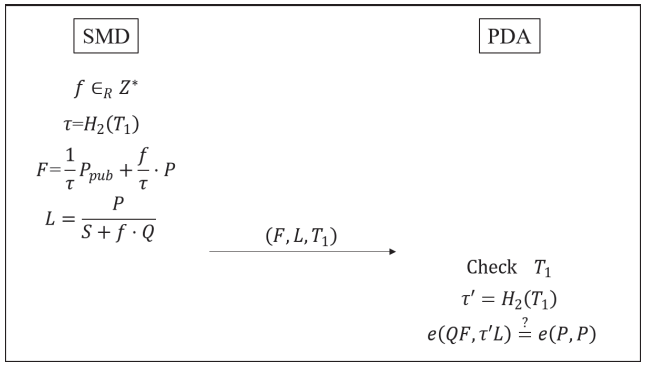


Figure 3: The process of device fast authentication.

authentication information delivery, and identity authentication. We will provide the detailed description of the three algorithms. The authentication information delivery and identity authentication are illustrated in Figure 3.

Registration. When every SMD enters WBAN, it needs to register with KGC.

KGC chooses a random number λ. Let a point P on $\mathbb{G}_1$ be a generator. The system public key is computed as $P_{pub} = \lambda \cdot P$. Choose two hash functions H_1 and H_2 as $H_1 : \{0,1\}^* \longrightarrow \mathbb{G}_1^*$ and $H_1 : \{0,1\}^* \longrightarrow \{0,1\}^l$, where l represents a length of the number. Suppose that the identity of a SMD is ID. The public key of the SMD is $Q = H_1(ID)$ and secret key $S = \lambda \cdot Q$. The system parameter will be written into the memory of the SMD.

Authentication Information Delivery. If the SMD wants to transmit message, its identity needs to be authenticated.

SMD first chooses a random number f. A current timestamp T_1 is recorded and hashed as $\tau = H_2(T_1)$. Authentication information F and L will then be calculated as follows.

$$F = \frac{1}{\tau} P_{pub} + \frac{f}{\tau} \cdot P,$$
$$L = \frac{P}{S + f \cdot Q}, \tag{1}$$

where τ is the hashed timestamp, P_{pub} is the public key of the system, f is the random number chosen by SMD, P represents the generator of the system, Q refers to the hashed ID, and S is the secret key of SMD.

The authentication information file is organized as (F, L, T_1), which is then sent to PDA for authentication.

Identity Authentication. After the authentication information file is received, PDA first checks T_1 to figure out whether the message is delayed when being transmitted. Then, $\tau' = H_2(T_1)$ is performed to calculate the hash value of the timestamp. Finally, PDA determines whether the device is a trusted one by the following formula.

$$e\left(QF, \tau' L\right) \stackrel{?}{=} e\left(P, P\right), \tag{2}$$

where Q is the hashed identity of the SMD, τ' is the hashed timestamp calculated by the PDA itself, and $e(P, P)$ can be computed offline.

5.3. Secure Message Transmission and Secure Instruction Distribution. The two algorithms, which are named as secure message transmission and secure instruction distribution, are both encryption methods. The encryption methods are proven to be safe in WBAN and can be utilized in our system.

We consider two entities in these algorithms: sender and receiver. The registration phase of the sender is the same as what is introduced in the algorithm of device fast authentication. The hashed identities of the sender and receiver are Q_1 and Q_2. The secret keys of the sender and the receiver are S_1 and S_2.

The sender first chooses random number r and computes $R = r \cdot P$. Then, the sender calculates $y = e(1/P, 1/Q_1 P)^r$.

Then the sender computes encrypted message file M with plaintext m:

$$M = m \oplus H_3\left(y\right), \tag{3}$$

where H_3 is a hash function: $H_3 : \{0,1\}^* \longrightarrow \{0,1\}^h$ and h is the length of the message file m.

The final parameter E is calculated as

$$E = \frac{S_1}{Q_2 S_1 R} + \frac{r}{Q_2 S_1 R}. \tag{4}$$

The file (M, E, R) will be sent to the receiver.

The receiver calculates parameter B as

$$B = \frac{S_2}{Q_1 P_{pub} + R}. \tag{5}$$

Let $y' = e(B, E)$ and $m = M \oplus H_3(y')$.

According to the above steps, the collected medical data or the instruction data can be transmitted to the receiver safely and acquired.

6. Security Analysis

In this section, the correctness and security against a forged SMD, MITM attack, and replay attack of our proposal are proved.

6.1. Correctness. The correctness of device fast authentication is proved as follows:

$$\begin{aligned} & e\left(QF, \tau' L\right) \\ &= e\left(Q\left(\frac{1}{\tau} P_{pub} + \frac{f}{\tau} \cdot P\right), \tau'\left(\frac{P}{S + f \cdot Q}\right)\right) \\ &= e\left(\frac{PQ}{\tau}(\lambda + f), \tau' P \frac{1}{\lambda Q + fQ}\right) \\ &= e\left(\frac{Q}{\tau}(\lambda + f) P, \frac{\tau'}{Q} \frac{1}{\lambda + f} P\right) \overset{if \tau = \tau'}{=} e(P, P) \end{aligned} \tag{6}$$

Obviously, if $\tau = \tau'$, $e(QF, \tau' L) = e(P, P)$. In fact, τ and τ' are computed with the same timestamp and hash function, so the authentication process is correct and efficient.

The correctness of the message transmission scheme for secure message transmission and secure instruction distribution is proved as follows:

$$\begin{aligned} y' &= e(B, E) = e\left(\frac{S_2}{Q_1 P_{pub} + R}, \frac{S_1}{Q_2 S_1 R} + \frac{r}{Q_2 S_1 R}\right) \\ &= e\left(\frac{\lambda Q_2}{(Q_1 \lambda + r) P}, \frac{\lambda Q_1 + r}{Q_2 \lambda Q_1 r P}\right) = e\left(\frac{1}{P}, \frac{1}{Q_1 P}\right)^r \\ &= y \end{aligned} \tag{7}$$

It can be seen from the proof that the values of y and y' are equal; obviously the recalculated m by y' is the same as the real one. So the design of the transmission scheme is also correct.

6.2. Security against a Forged SMD. A forged SMD may falsify the authentication information in order to pass the authentication. He can get the current timestamp to calculate τ. He can also select a random number f to calculate $F = (1/\tau) P_{pub} + f/\tau \cdot P$. Actually, the forged SMD has no opportunity to obtain the value of S. So he cannot calculate $L = P/(S + f \cdot Q)$. To sum up, our scheme can resist the attack by a forged SMD.

Table 1: Computational cost comparison.

Phases	SMD	PDA
Authentication information delivery	1H+3M	/
Identity authentication	/	1H+1M+1P

6.3. Security against Men-in-the-Middle Attack. MITM attack may cause medical information to be replaced or tampered in the middle. An adversary may calculate a fake secret key $S' = \lambda' Q$ and choose a new random number f' and replace the value τ with a new hashed timestamp T_2, remarked as τ_2. In fact, he still cannot pass the authentication. The new F' and L' can be calculated as follows:

$$\begin{aligned} F' &= \frac{1}{\tau_2} P_{pub} + \frac{f'}{\tau_2} \cdot P, \\ L' &= \frac{P}{S' + f' \cdot Q}. \end{aligned} \tag{8}$$

The faked file (F', L', T_2) will be sent to the PDA. PDA computes $\tau_2 = H_2(T_2)$. $e(QF, \tau_2 L)$ is computed as follows:

$$\begin{aligned} & e\left(QF, \tau_2 L\right) \\ &= e\left(Q\left(\frac{1}{\tau_2} P_{pub} + \frac{f'}{\tau_2} \cdot P\right), \tau_2\left(\frac{P}{S' + f' \cdot Q}\right)\right) \\ &= e\left(\frac{PQ}{\tau_2}\left(\lambda + f'\right), \tau_2 P \frac{1}{\lambda' Q + f' Q}\right) \\ &= e\left(\frac{Q}{\tau_2}\left(\lambda + f'\right) P, \frac{\tau_2}{Q} \frac{1}{\lambda' + f'} P\right) \\ &= e\left(\left(\lambda + f'\right) P, \frac{1}{\lambda' + f'} P\right) \neq e(P, P) \end{aligned} \tag{9}$$

It is not difficult to see that the new parameters of the MITM can not be certified.

6.4. Security against Replay Attack. An attacker who implements replay attack can try to pass the authentication by collecting files previously sent by SMD and sending an old file to PDA. In fact, there is a timestamp in the file. If the time difference between the timestamp and the time in which the file is accepted by PDA is beyond the range of the delay tolerance, the file will be identified as an invalid one. If the attacker tampered with the timestamp, he could not calculate the F matched with the new timestamp because the random number f is unknown.

7. Performance Analysis

The performance of the proposed scheme is discussed in this section. The computational cost of different entities in the proposed scheme is shown in Table 1. We take into consideration the computational costs of SMD and PDA. We consider the cost of collision-resistant hash function, bilinear pairing, and scalar multiplication [33]. In Table 1, M represents scalar multiplication, P denotes bilinear pairing,

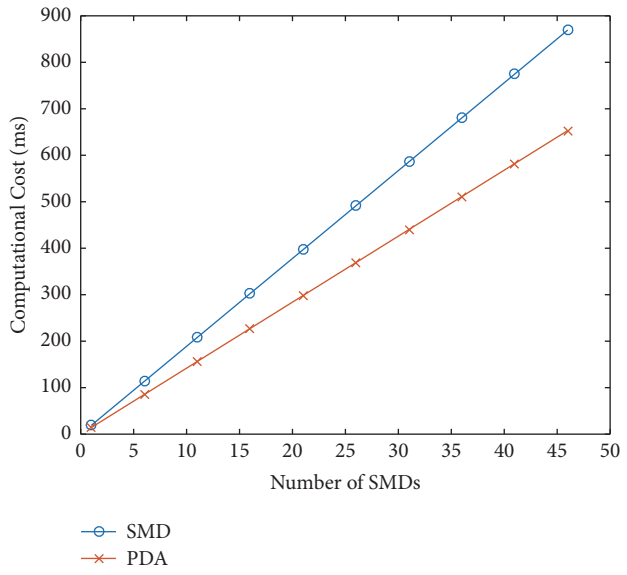


Figure 4: The time cost of SMD and PDA in the emergency medical system.

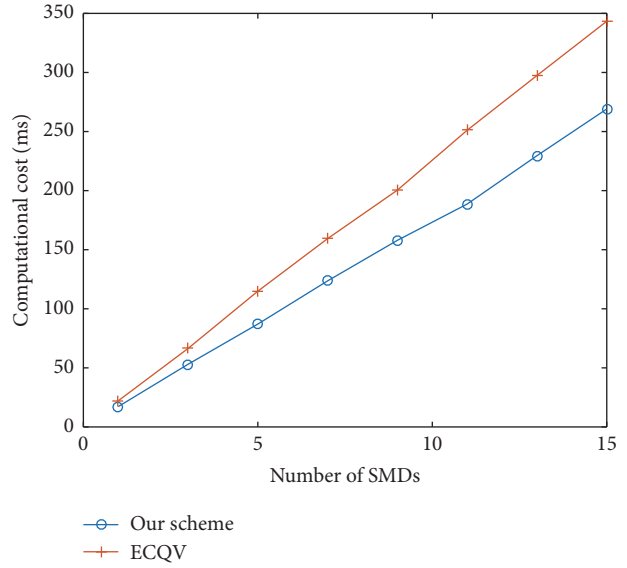


Figure 5: The time cost of an SMD in our scheme and ECQV.

and H represents collision-resistant hash function operation. The result comes out that the SMD costs 1 scalar multiplication and 1 collision-resistant hash function operation for sending the authentication message to the PDA. The PDA costs 1 scalar multiplication, 1 bilinear pairing, and 1 collision-resistant hash function operation to certificate the identity of SMD.

The efficiency of the proposed scheme is simulated on GNU Multiple Precision Arithmetic (GMP) library and Pairing-Based Cryptography (PBC) library (https://crypto.stanford.edu/pbc/). C language is utilized on a Linux system with Ubuntu 16.04 TLS, a 2.60 GHz Intel(R) Xeon(R) CPU E5-2650 v2, and 8 GB of RAM. The results are shown in Figure 4. Because SMDs and PDA are resources limited devices, controlling their computing resources consumption is very important. Our simulation reflects the time summation of all SMDs and the time cost of PDA when multiple SMDs send authentication requests to PDA. Because the number of devices in WBAN is limited, the simulation results show that our design can effectively reduce the computational cost of PDA. Figure 5 shows the comparison between the novel protocol and ECQV [34]. We can see that when the number of SMDs increase, the computational cost of our novel scheme is lower.

8. Conclusion

In this paper, we discuss the emergency situations in WBANs where the participation of doctors will seriously reduce the efficiency of treatment. In order to solve the problem of emergency treatment, we propose an emergency medical system. Based on the system, an identity-based fast authentication scheme for smart mobile devices in WBANs is proposed. In addition, we also provide a message transmission scheme to improve the system. The authentication scheme is proven to be secure and efficient in our analysis and simulation.

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 work is supported by the National Natural Science Foundation of China under Grant no. 61672295, no. 61672290, no. U1405254, and no. 61772280, Guangxi Key Laboratory of Cryptography and Information Security under Grant no. GCIS201715, the State Key Laboratory of Information Security under Grant no. 2017-MS-10, the 2015 Project of six personnel in Jiangsu Province under Grant no. R2015L06, the CICAEET fund, and the PAPD fund.

References

[1] 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.

[2] X. F. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, "New algorithms for secure outsourcing of modular exponentiations," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 9, pp. 2386–2396, 2014.

[3] J. Shen, A. Wang, C. Wang, J. Li, and Y. Zhang, "Content-centric group user authentication for secure social networks," *IEEE Transactions on Emerging Topics in Computing*, 2017.

[4] T. Jiang, X. Chen, and J. Ma, "Public integrity auditing for shared dynamic cloud data with group user revocation," *IEEE Transactions on Computers*, vol. 65, no. 8, pp. 2363–2373, 2016.

[5] J. Shen, T. Zhou, X. Chen, J. Li, and W. Susilo, "Anonymous and traceable group data sharing in cloud computing," *IEEE*

Transactions on Information Forensics and Security, vol. 13, no. 4, pp. 912–925, 2018.

[6] J. Li, Y. Zhang, X. Chen, and Y. Xiang, "Secure attribute-based data sharing for resource-limited users in cloud computing," *Computers & Security*, vol. 72, pp. 1–12, 2018.

[7] J. Shen, T. Zhou, D. He, Y. Zhang, X. Sun, and Y. Xiang, "Block design-based key agreement for group data sharing in cloud computing," *IEEE Transactions on Dependable and Secure Computing*, pp. 1-1, 2017.

[8] J. Wang, X. Chen, J. Li, J. Zhao, and J. Shen, "Towards achieving flexible and verifiable search for outsourced database in cloud computing," *Future Generation Computer Systems*, vol. 67, pp. 266–275, 2017.

[9] Q. Jiang, M. K. Khan, X. Lu, J. Ma, and D. He, "A privacy preserving three-factor authentication protocol for e-health clouds," *Journal of Supercomputing*, vol. 72, no. 10, pp. 3826–3849, 2016.

[10] J. Shen, C. Wang, and A. Wang, "Intelligent agent-based region division scheme for mobile sensor networks," *Soft Computing*, 2018.

[11] C.-F. Lai, S. Zeadally, J. Shen, and Y.-X. Lai, "A cloud-integrated appliance recognition approach over internet of things," *Journal of Internet Technology*, vol. 16, no. 7, pp. 1157–1168, 2015.

[12] J. Shen, C. Wang, A. Wang, Q. Liu, and Y. Xiang, "Moving centroid based routing protocol for incompletely predictable cyber devices in Cyber-Physical-Social Distributed Systems," *Future Generation Computer Systems*, 2017.

[13] X. Chen, J. Li, X. Huang, J. Ma, and W. Lou, "New publicly verifiable databases with efficient updates," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 5, pp. 546–556, 2015.

[14] J. Shen, D. Liu, J. Shen, Q. Liu, and X. Sun, "A secure cloud-assisted urban data sharing framework for ubiquitous-cities," *Pervasive and Mobile Computing*, vol. 41, pp. 219–230, 2017.

[15] D. Wang and P. Wang, "On the anonymity of two-factor authentication schemes for wireless sensor networks: Attacks, principle and solutions," *Computer Networks*, vol. 73, pp. 41–57, 2014.

[16] 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*, pp. 1-1, 201.

[17] S.-Y. Chiou, Z. Ying, and J. Liu, "Improvement of a privacy authentication scheme based on cloud for medical environment," *Journal of Medical Systems*, vol. 40, no. 4, article 101, 15 pages, 2016.

[18] C.-T. Li, C.-C. Lee, and C.-Y. Weng, "A secure cloud-assisted wireless body area network in mobile emergency medical care system," *Journal of Medical Systems*, vol. 40, no. 5, article 117, 15 pages, 2016.

[19] C.-T. Li, C.-C. Lee, C.-Y. Weng, and S.-J. Chen, "A secure dynamic identity and chaotic maps based user authentication and key agreement scheme for e-healthcare systems," *Journal of Medical Systems*, vol. 40, no. 11, article 233, 10 pages, 2016.

[20] C. Li, T. Wu, C. Chen, C. Lee, and C. Chen, "An efficient user authentication and user anonymity scheme with provably security for IoT-based medical care system," *Sensors*, vol. 17, no. 7, 1482, 18 pages, 2017.

[21] X. Li, M. H. Ibrahim, S. Kumari, A. K. Sangaiah, V. Gupta, and K. R. Choo, "Anonymous mutual authentication and key agreement scheme for wearable sensors in wireless body area networks," *Computer Networks*, vol. 129, no. 2, pp. 429–443, 2017.

[22] A. K. Das, A. K. Sutrala, V. Odelu, and A. Goswami, "A secure smartcard-based anonymous user authentication scheme for healthcare applications using wireless medical sensor networks," *Wireless Personal Communications*, vol. 94, no. 3, pp. 1899–1933, 2017.

[23] J. Srinivas, D. Mishra, and S. Mukhopadhyay, "A mutual authentication framework for wireless medical sensor networks," *Journal of Medical Systems*, vol. 41, no. 5, article 80, 19 pages, 2017.

[24] T. Hayajneh, B. J. Mohd, M. Imran, G. Almashaqbeh, and A. V. Vasilakos, "Secure authentication for remote patient monitoring withwireless medical sensor networks," *Sensors*, vol. 16, no. 4, article 424, 25 pages, 2016.

[25] Y. Park and Y. Park, "A selective group authentication scheme for IoT-based medical information system," *Journal of Medical Systems*, vol. 41, no. 4, article 48, 8 pages, 2017.

[26] P. Mohit, R. Amin, A. Karati, G. P. Biswas, and M. K. Khan, "A standard mutual authentication protocol for cloud computing based health care system," *Journal of Medical Systems*, vol. 41, no. 4, article 50, 13 pages, 2017.

[27] X. Li, J. Niu, M. Karuppiah, S. Kumari, and F. Wu, "Secure and efficient two-factor user authentication scheme with user anonymity for network based e-health care applications," *Journal of Medical Systems*, vol. 40, no. 12, article 268, 12 pages, 2016.

[28] X. Li, M. H. Ibrahim, S. Kumari, and R. Kumar, "Secure and efficient anonymous authentication scheme for three-tier mobile healthcare systems with wearable sensors," *Telecommunication Systems*, vol. 67, no. 3, pp. 1–26, 2017.

[29] Y. Xiang, W. Zhou, and M. Guo, "Flexible deterministic packet marking: An IP traceback system to find the real source of attacks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 4, pp. 567–580, 2009.

[30] D. Wang, H. Cheng, D. He, and P. Wang, "On the challenges in designing identity-based privacy-preserving authentication schemes for mobile devices," *IEEE Systems Journal*, vol. 12, no. 1, pp. 916–925, 2018.

[31] D. He, S. Zeadally, N. Kumar, and J. H. Lee, "Anonymous authentication for wireless body area networks with provable security," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2590–2601, 2017.

[32] D. Wang and P. Wang, "Understanding security failures of two-factor authentication schemes for real-time applications in hierarchical wireless sensor networks," *Ad Hoc Networks*, vol. 20, no. 2, pp. 1–15, 2014.

[33] X. Chen, F. Zhang, W. Susilo, H. Tian, J. Li, and K. Kim, "Identity-based chameleon hashing and signatures without key exposure," *Information Sciences*, vol. 265, pp. 198–210, 2014.

[34] J. Shen, S. Chang, Q. Liu, and Y. Ren, "Implicit authentication protocol and self-healing key management for WBANs," *Multimedia Tools & Applications*, vol. 77, no. 9, pp. 11381–11401, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 7579161, 16 pages
https://doi.org/10.1155/2018/7579161

Research Article

Trusted Authority Assisted Three-Factor Authentication and Key Agreement Protocol for the Implantable Medical System

Deming Mao,[1] **Ling Zhang**,[2] **Xiaoyu Li**,[1] **and Dejun Mu**[1]

[1]*School of Automation, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China*
[2]*Southwest Institute of Telecommunication, Chengdu 610041, Sichuan, China*

Correspondence should be addressed to Deming Mao; maodmnwpu@163.com

Received 5 April 2018; Accepted 27 June 2018; Published 29 July 2018

Academic Editor: Joseph Liu

The application of implantable medical devices (IMDs), which solves the problems of geographical distance limitation and real-time health monitoring that plague patients and doctors, has caused great repercussions in the medical community. Despite the great potential of wide application, it also brings some security and privacy issues, such as the leakage of health data and unauthorized access to IMDs. Although a number of authentication and key agreement (AKA) schemes have been developed, we find that some subtle attacks still remain to be addressed. Then we propose an improved AKA scheme which achieves strong security features including user anonymity and known key security. It is formally proved to be secure under the Real-or-Random model. Moreover, a comprehensive security analysis shows that our scheme can resist various attacks and satisfy the desired requirements. Finally, the performance analysis shows the superiority of our protocol which is suitable for the implantable medical system.

1. Introduction

With the improvement of wireless communication technologies, the implantable medical devices (IMDs), such as pacemakers, cranial nerve stimulators, and cochlear implants, have been widely used in the medical services field [1, 2]. All these micro devices implanted in patients' body can continuously monitor and collect data to reflect the patient's health. Through controller node (CN), implantable medical devices are able to transmit the data to the remote attending physician or the medical institution, which greatly simplifies the treatment process of patients and breaks the limitation of region. Generally speaking, the combination of these advanced technologies improves health care practices, urgent care, and preventive health [3].

A typical architecture of implantable medical system is shown in Figure 1. CN and IMDs firstly register to the trusted authority (TA) before they are deployed into the system. Then, IMDs collect data such as body temperature, heart beats, and blood pressure, which can be derived by CN via wireless communication technologies, such as Bluetooth or ZigBee [4]. After the collection process, the CN needs to be plugged into the Internet via an access point to be accessible by the attending physician or the medical institution. In the meantime, cloud servers may be used for storing collected health data to ease the storage burden on mobile devices [5, 6].

However, it is the application of wireless communication that makes the transmission of medical data face the potential security risks [7–9]. According to the Dolev-Yao threat model [10], the implantable medical system is facing a wide range of malicious attacks which may cause the leakage of health data and unauthorized access to IMDs. In response to the serious security threats, it is imperative to design a mutual authentication and key agreement (AKA) mechanism which can ensure the confidentiality of the transmitted sensor data and resist malicious attacks.

1.1. Related Work. With the wireless interface enabled, IMDs can be accessed by an authorized operator in physical proximity via the IMDs programmer. However, the wireless communication and networking capabilities of IMDs turn out to be the major sources of security vulnerabilities [11, 12]. For this purpose, access control for implantable medical

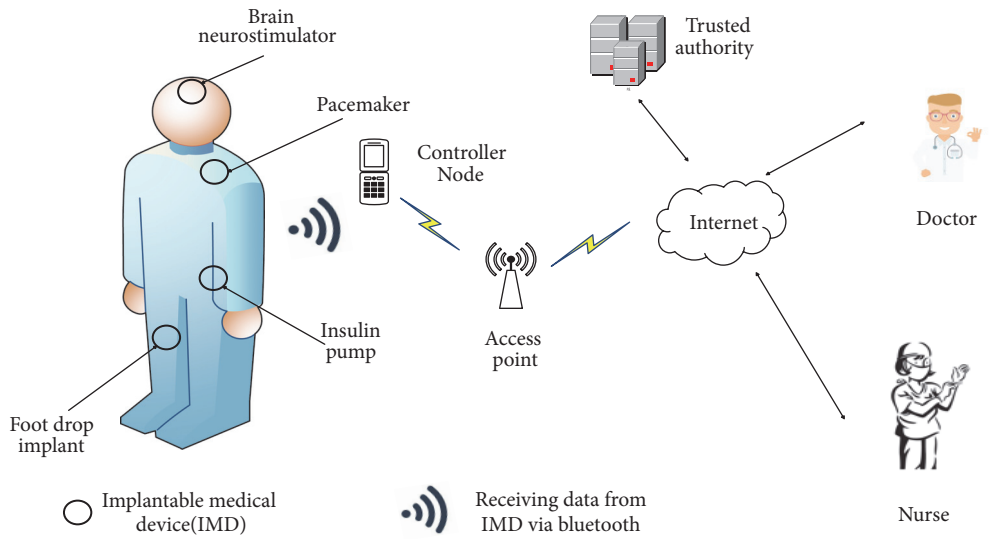


Figure 1: The network model of the implantable medical system.

system is highly desired and many schemes have also been put forward in this field.

Initially, considering the scarce energy reserves and limited communication capacity of IMDs, some schemes based on symmetric key cryptography [15–19] were proposed, they realized high encryption speed and efficiency at the same time but showed weaknesses of resisting against certain attacks, and the complexity of key management will introduce large memory and communication overhead which contradicts their original intentions. This means that the symmetric key cryptography based schemes are difficult to provide a complete security guarantee for implantable medical system.

Then, traditional public key cryptography (TPKC) based authentication schemes [20, 21] were implemented in IMDs. Unfortunately, the limited computing capability and battery capacity of the mobile device hinders the application of TPKC in implantable medical system. The concept of ECC (Elliptic curve cryptosystem) was then put forward [22] which provided the same security with a much smaller key size compare to the TPKC [23] so that many ECC-based protocols were proposed subsequently [13, 24]. In 2013, Liu et al. [25] put forward a scheme in which they used the bilinear pairing defined on the elliptic curve to design a new certificateless signature scheme, but later in 2014, Xiong [26] analyzed the Liu et al.'s authentication protocol and concluded that their scheme was prone to a kind of attack by a key replacement adversary [27]. In 2016, He et al. [28] also claimed that the Liu et al.'s scheme cannot resist the impersonation attack; meanwhile they put forward their own improved protocol. In 2018, Li et al. [29] analyzed the loopholes in each layer of the current implantable medical system and put forward a complete three-layer scheme.

As we know, each authentication factor has its own advantages and disadvantages. Passwords are prone to dictionary attacks while smart cards may be lost. A number of two-factor protocols [30–38] have been put forward. In these schemes, two kinds of factors, i.e., passwords and smart cards, are combined to achieve user authentication. In 2015, He et al. came up with a scheme [35] where the smart card is used to store some private parameters about healthcare applications using wireless medical sensor networks. Wei et al. proposed an anonymous authentication scheme [33] for wireless body area networks in 2017 as well as gave a formal security analysis of the protocol.

To further enhance the security strength of two-factor protocols, three-factor authentication (3FA) schemes which consolidate all three factors (i.e., passwords, smart cards, and biometrics) have attracted more and more attentions [14, 39–44]. In 2017, Wei applied the fuzzy extractor scheme into his newly proposed protocol [39] to handle the biometrics. Meanwhile Jiang et al. presented a scheme [41] where the biohashing is used to protect the biometrics. In 2016, Wu et al. proposed a 3FA scheme [43] aiming at summarizing the flaws that existed in previous typical protocols and came up with a more complete solution. In 2017, Li et al. [40] remedied flaws in Jiang et al.'s scheme [32] in which fuzzy commitment is used to protect biometrics. In 2017, Wazid et al. provided a 3FA scheme [14] for IMDs and claimed that their protocol could meet the known security, but we reveal that the protocol cannot achieve complete security.

1.2. Motivations and Contributions. With the popularity of the IMD, its safety and privacy protection have attracted great attention and a large number protocols in this field have emerged, but few of them can achieve the desired security guarantee. In such a situation, it is imperative to sum up the defects in previous protocols and propose new schemes to make the implantable medical system more secure and reliable. Among these protocols, we pick Wazid et al.'s scheme [14] as a typical case study to analyze some defects of the scheme. Then we propose a trusted authority assisted 3FA protocol which effectively solves the security vulnerabilities in the original protocol. Our contributions are summarized as follows:

(i) First, we find out three drawbacks of the most recent 3FA protocol of Wazid et al. To be specific, we find that the scheme cannot withstand offline password guessing attack, the CN impersonation attack, and the authentication phase of the protocol is problematic.

(ii) Second, we propose a trusted authority assisted 3FA protocol. Specifically, we introduce the fuzzy verifier [45] to effectively prevent offline password guessing attack during local login verification phase and adopt the widely used fuzzy vault [46] to protect the biometric template.

(iii) Third, we analyze the security of our protocol both formally and informally. Our protocol not only properly solves the shortcomings in the original scheme, but also achieves perfect forward security, user anonymity, know key security, and so forth. At the same time, our protocol can resist a variety of known attacks.

1.3. Organization of the Paper. The rest of the paper is organized as follows. In Section 2, we briefly review some preliminaries used in this paper, including ECC and the fuzzy vault. Section 3 depicts the details of Wazid et al.'s scheme. Then in Section 4, we present the vulnerabilities in their scheme. In Section 5, we propose an improved scheme. In Section 6, we have an elaborate analysis from both formal and informal point of view. The comparisons of efficiency and features are listed in Section 7. In the end, this paper is concluded in Section 8.

2. Preliminaries

2.1. Fuzzy Vault. The fuzzy vault is a constructor used to protect biometric templates BIO with various built-in algorithms. Its security relies on the secret key K and BIO. It works in key binding mode where the biometric and the key are monolithically bound within a binding mechanism. Compared with fuzzy extractor [47], the Euclidean distance measurement used in fuzzy vault has been widely accepted in most biostatistical applications [48]. Therefore, in view of the value in practice, we will adopt the fuzzy vault to protect biometric features in our improved scheme.

Specifically, the user selects a polynomial Pol which is used to encode secret key K and be evaluated on all elements in BIO. Then the biometric BIO which is imprinted by user can be converted into a set of L points which lie on the Pol according to $Gen(BIO, K, Pol) = L$. Then, taking L and CP which is a large set of "chaff points" as inputs of $Enc(\cdot)$, we can get the final vault V which equals $CP \cup L$, that is, $Enc(CP, L) = V$. Generally, we put the final vault V in the mobile device.

When the user wants to recover the secret key K, she/he can scan the biometric BIO^* on terminal firstly, then taking the vault V and BIO^* as the inputs of the algorithm $Dec(\cdot)$ which will output the Pol if and only if $|BIO - BIO^*| < \varepsilon$ where ε is the fuzziness parameter. The secret key K can be recovered with the input Pol by the algorithm $Rec(\cdot)$ finally.

2.2. Elliptic Curve Cryptosystem (ECC). Compared with the traditional RSA algorithm, ECC achieves the same security strength with much smaller key size, so ECC is more efficient than RSA. Elliptic curve equation is defined in such a form: nonsingular elliptic curve $E_p(a, b) : y^2 - x^3 + ax + b (\mathrm{mod} p)$ over a prime finite field Z_p, where p is a large prime and $a, b \in Z_p^*$ satisfies $4a^3 + 27b^2 \neq 0 (\mathrm{mod} p)$.

Besides, there are two difficult problems in ECC, namely, Elliptic Curve Discrete Logarithm Problem (ECDLP) and Elliptic Curve Computational Diffie-Hellman Problem (ECCDHP). Specifically, the first one depicts that it is impossible to find an integer $x \in Z_p^*$ that satisfies the formula $Q = x \cdot P$ with two given points P and Q over $E_p(a, b)$. The other one describes that it is hard to calculate the value $xy \cdot P$ with the given points P, $x \cdot P$ and $y \cdot P$, $x, y \in Z_p^*$. These two hard problems guarantee the security of Elliptic Curve primitives, and an adversary still has a great deal of difficulty in getting the secret after obtaining the public values.

Table 1: Notations.

Notations	Description
U_i, MD_i	i_{th} user and his/her mobile device
CN_j	j_{th} controller node
IMD_l	l_{th} implantable medical device
TA	Trusted authority
ID_i, PW_i, BIO_i	U_i's identity, password and biometric information
ID_{TA}, ID_{CN_j}	Identities of TA and controller node
$RID_i, RID_{CN_j}, RTS_{CN_j}$	Pseudo identities of U_i and CN_j, registration timestamp of CN_j
N	1024-bit secret number of TA
T_i	Current timestamp
ΔT	Maximum transmission delay associated with a message
t	Error tolerance threshold used in fuzzy extractor
$k \cdot P$	Elliptic curve point multiplication, $k \in Z_p^*$, $P \in E_p(a, b)$
$h(\cdot)$	Collision-resistant cryptographic hash function
$\|$	Concatenation operation
$\oplus$	Bitwise XOR operation

3. Review of Wazid et al.'s Scheme

In this section, we review the details of Wazid et al.'s scheme, which consists of eight phases, i.e., predeployment, postdeployment, registration, login, authentication and key agreement, password and biometric update, and dynamic control node addition, as well as dynamic IMD addition. The scheme is for the purpose of mutual authentication and key agreement establishment between the mobile device and IMDs. The notations used in this paper are listed in Table 1.

3.1. Predeployment Phase. Before deployment, a trusted authority TA needs to complete the registration for each CN_j

as well as IMD_l. TA first selects a secret 1024-bit number N for CN_j and IMD_l. Then TA picks the identity ID_{CN_j} for CN_j and calculates $RID_{TA} = h(ID_{TA} \| N)$, $RID_{CN_j} = h(ID_{CN_j} \| N)$, $TC_{CN_j} = h(ID_{TA} \| RTS_{CN_j} \| N)$. Meanwhile, TA constructs the univariate polynomial $P(RID_{CN_j}, y)$ according to the polynomial-based key distribution $P(x, y) = \sum_{i=0}^{n} \sum_{j=0}^{n} g_{i,j} x^i y^j \in GF(p)[x, y]$ proposed in [49] where the prime p is chosen as a large number and n is also large to preserve unconditional security and n-collusion resistant property against IMD capture attack. Finally, TA stores $\{RID_{TA}, RID_{CN_j}, TC_{CN_j}, P(RID_{CN_j}, y)\}$ in the memory of CN_j. Similar to the above calculations, TA generates a unique identity ID_{IMD_l} and calculates $RID_{IMD_l} = h(ID_{IMD_l} \| N)$, $P(RID_{IMD_l}, y)$ and then stores the information $\{RID_{IMD_l}, P(RID_{IMD_l}, y)\}$ in the memory of ID_{IMD_l}.

3.2. Postdeployment Phase. After the predeployment phase, CN_j and IMD_l establish a shared key using the information distributed during the predeployment phase. The details of the process are as follows. Firstly, IMD_l sends the message $\langle RID_{IMD_l} \rangle$ to CN_j. Once CN_j receives the message, CN_j responds with the message $\langle RID_{CN_j} \rangle$. Then they calculate the same shared secret key $SK_{IMD_l,CN_j} = P(RID_{IMD_l}, RID_{CN_j})$ and $SK_{CN_j,IMD_l} = P(RID_{IMD_l}, RID_{CN_j})$ on each own for future use.

3.3. Registration Phase. This phase has 4 steps.

Step 1. The user selects his/her identity ID_i at will and forwards it with registration request to TA in a secure channel.

Step 2. After accepting the request, TA computes the pseudo identity of U_i as $RID_i = h(ID_i \| N)$. Then TA continues to compute the value A_i as $A_i = h(RID_{TA} \| ID_i)$. TA sends the message $\langle RID_i, A_i, RID_{TA} \rangle$ to U_i.

Step 3. After receiving registration reply from TA, U_i further selects a private key $k \in Z_p^*$ and computes the corresponding public key $Q = k \cdot P$.

Step 4. U_i inputs his/her password PW_i and imprints fingerprint BIO_i in mobile device MD_i, then MD_i calculates $Gen(BIO_i) = (\sigma_i, \tau_i)$, $RID_i' = RID_i \oplus h(PW_i \| \sigma_i)$, $RPW_i = h(PW_i \| k)$, $D_i = k \oplus h(ID_i \| PW_i \| \sigma_i)$, $RID_{TA}' = RID_{TA} \oplus h(ID_i \| k \| \sigma_i)$, $A_i' = A_i \oplus h(k \| \sigma_i)$, $B_i = h(A_i \| RPW_i)$, and $C_i = h(ID_i \| RID_{TA} \| B_i \| \sigma_i)$. At last, MD_i keeps the data $\{RID_i', RID_{TA}', A_i', C_i, D_i, \tau_i, Gen(\cdot), Rep(\cdot), h(\cdot), t\}$ in its memory.

3.4. Login Phase. As depicted in Figure 2, to login to CN_j, U_i executes the following steps.

Step 1. U_i inputs his/her ID_i, PW_i and BIO_i', then MD_i retrieves the biometric key $\sigma_i' = Rep(BIO_i', \tau_i)$. Then MD_i computes $k' = D_i \oplus h(ID_i \| PW_i \| \sigma_i')$, $RPW_i' = h(PW_i \| k')$, $A_i^* = A_i' \oplus h(k' \| \sigma_i')$, $B_i^* = h(A_i^* \| RPW_i')$, $RID_{TA}^* = RID_{TA}' \oplus h(ID_i \| k' \| \sigma_i')$, $RID_i^* = RID_i' \oplus h(PW_i \| \sigma_i')$, and $C_i^* = h(ID_i \| RID_{TA}^* \| B_I^* \| \sigma_i')$. If C_i^* equals the stored C_i, it means that U_i's inputs are verified as correct; otherwise, the login phase will be terminated immediately.

Step 2. MD_i picks the current timestamp T_1 and a 160-bit random nonce r_i. Then MD_i computes $a_i = h(r_i \| T_1 \| RID_i^* \| RPW_i' \| \sigma_i')$, $b_i = h(RID_{TA}^* \| T_1)$, and $M_1 = a_i \cdot P$ as well as the signature $M_2 = a_i + k' b_i (\mathrm{mod} p)$. At last, MD_i forwards the message $\langle M_1, M_2, T_1 \rangle$ to CN_j via a public channel.

3.5. Authentication and Key Agreement Phase. In this phase, U_i and CN_j need to authenticate each other as well as establish a session key between them for future safe communications; see Figure 2.

Step 1. After obtaining the message $\langle M_1, M_2, T_1 \rangle$, CN_j first checks $|T_1 - T_1^*| ? < \Delta T$, if two values are equal, CN_j calculates $b_i' = h(RID_{TA} \| T_1)$, and then checks $M_2 \cdot P ? = M_1 + b_i' \cdot Q$. Similarly, if verification matches, it indicates that U_i is considered legitimate. Then CN_j chooses T_2 and a random number r_j and continues to compute $c_j = h(r_j \| T_2 \| RID_{CN_j} \| TC_{CN_j})$, $M_4 = c_j \cdot P$, $k_{ij} = c_j \cdot M_1 = (a_i c_j) \cdot P$, session key $SK_{ij} = h(k_{ij} \| RID_{TA} \| T_1 \| T_2)$, and $M_5 = h(SK_{ij} \| T_2)$. Finally, CN_j sends the message $\langle M_4, M_5, T_2 \rangle$ to U_i through the public channel.

Step 2. After receiving the message from CN_j, U_i first judges $|T_2 - T_2^*| ? < \Delta T$, then computes $k_{ij}^* = a_i \cdot M_4 = (a_i c_j) \cdot P$, $SK_{ij}^* = h(k_{ij}^* \| RID_{TA}^* \| T_1 \| T_2)$, and $M_6 = h(SK_{ij}^* \| T_2)$. If $M_6 = M_5$, it indicts that CN_j passes the verification. With that, U_i calculates $M_7 = h(SK_{ij}^* \| T_3)$ and forwards the message $\langle M_7, T_3 \rangle$ to CN_j.

Step 3. CN_j checks $|T_3 - T_3^*| ? < \Delta T$, then computes $M_8 = h(SK_{ij} \| T_3)$, and judges whether $M_8 = M_7$.

Finally, both CN_j and U_i complete the mutual authentication and agree on the same session key which will used for the secure communications in future.

3.6. Password and Biometric Update Phase. If U_i wants to change the password, he/she can execute ensuing procedure.

Step 1. Firstly, U_i inputs ID_i, PW_i^{old}, and BIO_i^{old}. MD_i computes $\sigma_i^{old} = Rep(BIO_i^{old}, \tau_i)$, $k = D_i \oplus h(ID_i \| PW_i^{old} \| \sigma_i^{old})$, $RPW_i^{old} = h(PW_i^{old} \| k)$, $A_i^{old} = A_i' \oplus h(k \| \sigma_i^{old})$, $B_i^{old} = h(A_i^{old} \| RPW_i^{old})$, and $RID_{TA} = RID_{TA}' \oplus h(ID_i \| k \| \sigma_i^{old})$ and checks if C_i^{old} equals $h(ID_i \| RID_{TA} \| B_i^{old} \| \sigma_i^{old})$. If it holds, MD_i asks U_i for the new password PW_i^{new}.

Step 2. After U_i inputs the PW_i^{new} and MD_i calculates $\sigma_i^{new} = Rep(BIO_i^{new}, \tau_i^{new})$, $RPW_i^{new} = h(PW_i^{new} \| k)$, $A_i^{new} = A_i^{old} \oplus h(k \| \sigma_i^{new})$, $B_i^{new} = h(A_i^{new} \| RPW_i^{new})$, $RID_{TA}'' = RID_{TA} \oplus h(ID_i \| k \| \sigma_i^{new})$, $C_i^{new} = h(ID_i \| RID_{TA}'' \| B_i^{new} \| \sigma_i^{new})$, and $D_i'' = k \oplus h(ID_i \| PW_i^{new} \| \sigma_i^{new})$. Finally, MD_i replaces RID_i',

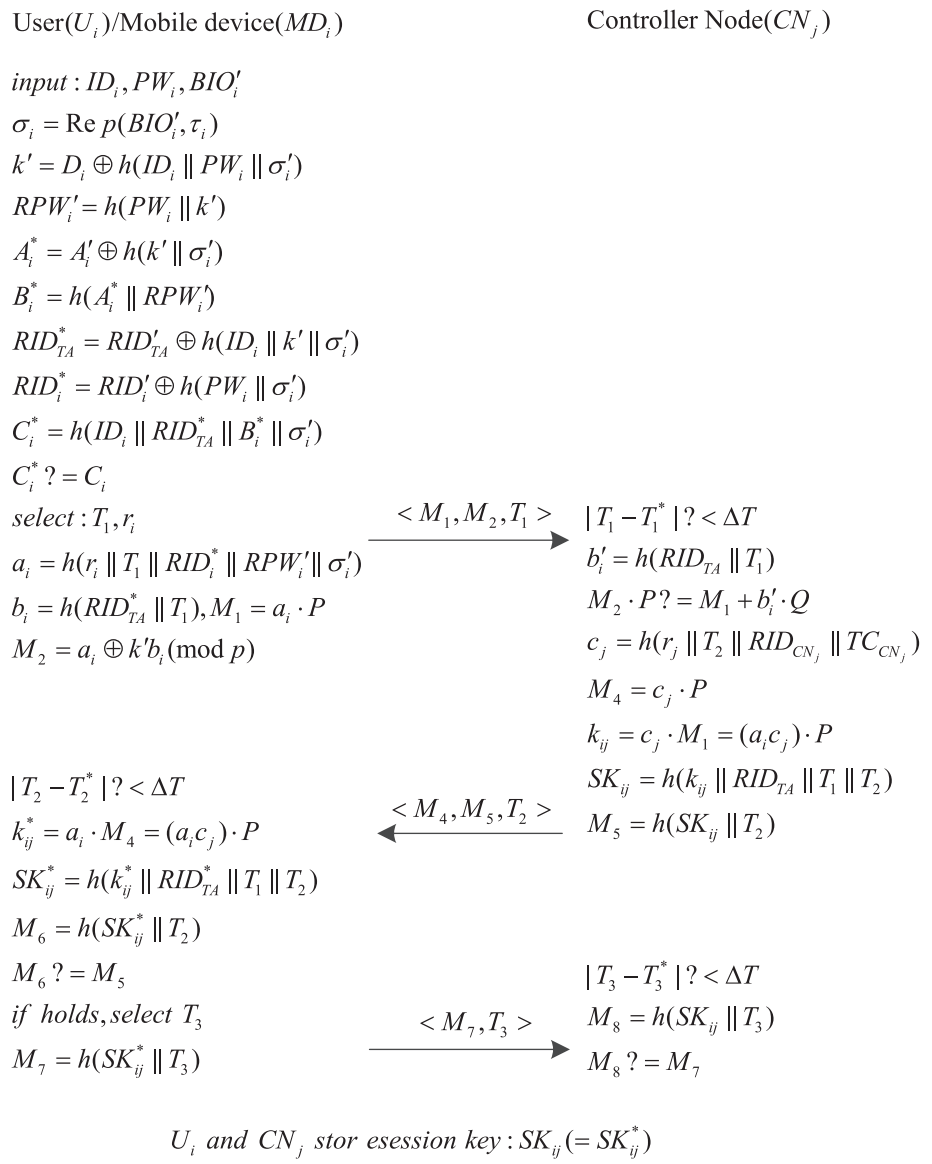


Figure 2: Login and authentication phase of Wazid et al.'s scheme.

RID_{TA}', A_i', C_i, D_i, and τ_i with RID_i'', RID_{TA}'', A_i^{new}, C_i^{new}, D_i'', and τ_i^{new}, respectively.

3.7. Dynamic Controller Node Addition Phase. In this phase, a new controller node CN_j^{new} can be deployed as follows.

First, TA determines a new identity $ID_{CN_j}^{new}$ for CN_j^{new} and calculates $RID_{CN_j}^{new} = h(ID_{CN_j}^{new} \| N)$ and new polynomial $P(RID_{CN_j}^{new}, y)$ as well as $TC_{CN_j}^{new} = h(ID_{TA} \| RTS_{CN_j}^{new} \| N)$ in which the $RTS_{CN_j}^{new}$ is the newly generated registration timestamp. Finally, TA stores the parameters $\{RID_{CN_j}^{new}, TC_{CN_j}^{new}, RID_{TA}, P(RID_{CN_j}^{new}, y)\}$ into the memory of CN_j^{new} before it is deployed into the system.

3.8. Dynamic IMD Addition Phase. In this phase, we can deploy a new IMD (IMD_l^{new}). Specifically, TA computes $RID_{IMD_l}' = h(ID_{IMD_l}' \| N)$ and $P(RID_{IMD_l}', y)$ and then stores $\{RID_{IMD_l}'^{new}, P(RID_{IMD_l}'^{new}, y)\}$ in the memory of IMD_l^{new}.

4. Weakness of the Wazid et al.'s Scheme

The widely accepted Dolev-Yao threat model (DY model) [10] demonstrates that the adversary A can fully control the public channel between communicators. That is, A is capable of eavesdropping, stealing, inserting, deleting, and modifying the messages in the open channel. Most recently, Wang et al. [45] have provided a complete summary of the adversary's capabilities and present twelve evaluation criteria for a secure protocol, i.e., no password verifier-table, no smart card loss attack, mutual authentication, and so forth. According to above evaluation criteria, we make a reasonable analysis of Wazid et al.'s scheme and find that the protocol

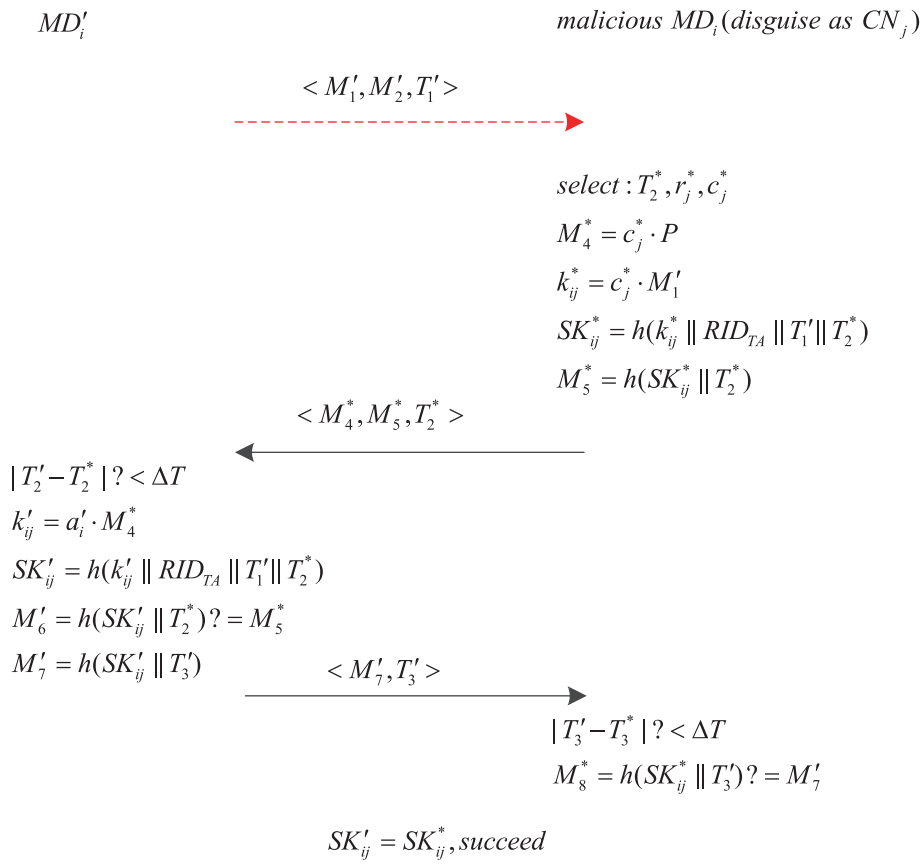


Figure 3: The controller node impersonation attack in Wazid et al.'s scheme.

has the following three flaws, i.e., offline password guessing attack, controller node impersonation attack, and Incorrect authentication process. As a result, it cannot achieve mutual authentication; that is, the scheme fails to meet the security claimed by the authors.

4.1. Offline Password Guessing Attack. To achieve user friendliness, in registration phase, users are allowed to choose their own identities and passwords at will; the majority of users will choose easy-to-recall ID and PW; the combination of these low entropy ID and PW are likely to be vulnerable to offline guessing attack. A probabilistic polynomial time (PPT) adversary can offline enumerate all (ID, PW) pairs in Cartesian product $D_{id} * D_{pw}$, where D_{id} and D_{pw} represent ID space and PW space, respectively. In a 3FA protocol, we should ensure that even the MD_i and biometric have been corrupted, and the whole scheme can still resist this type attack to protect the security of user's secrets. Based on all above assumptions, the adversary can launch an offline password guessing attack through the following processes.

Step 1. We assume that the adversary A has acquired MD_i and biometric BIO_i of the user and then obtains the secret parameters $\{RID_i', RID_{TA}', A_i', C_i, D_i, \tau_i, Gen(\cdot), Rep(\cdot), h(\cdot), t\}$ stored in the MD_i.

Step 2. The adversary A picks a (ID_i', PW_i') pair and calculates $\sigma_i = Rep(BIO_i, \tau_i)$, $k^* = D_i \oplus h(ID_i^* \| PW_i^* \| \sigma_i)$, $RPW_i^* = h(PW_i^* \| k^*)$, $A_i^* = A_i' \oplus h(k^* \| \sigma_i)$, $B_i^* = h(A_i^* \| RPW_i^*)$, $RID_{TA}^* = RID_{TA}' \oplus h(ID_i^* \| k^* \| \sigma_i)$, $RID_i^* = RID_i' \oplus h(PW_i^* \| \sigma_i)$, and $C_i^* = h(ID_i' \| RID_{TA}' \| B_i^* \| \sigma_i)$.

Step 3. Finally, A checks whether $C_i^* = C_i$, and if it holds, we can say that the (ID_i', PW_i') selected by the adversary is a legal one. Otherwise, A can choose another (ID_i, PW_i) pair to continue implementing above steps until success.

4.2. The Controller Node Impersonation Attack. In registration phase, TA picks a secret number N and calculates TA's pseudo identifier $RID_{TA} = h(ID_{TA} \| N)$ which is a fixed value. What is more, in predeployment phase, both MD_i and CN_j have obtained RID_{TA}; for a malicious MD_i, he/she can disguise himself/herself as CN_j to communicate with another MD_i' as shown in Figure 3.

Step 1. The malicious MD_i intercepts the first authentication message $\langle M_1', M_2', T_1' \rangle$ sent by MD_i' which is ought to have been received by CN_j.

Step 2. Then MD_i can impersonate CN_j to communicate with MD_i', MD_i selects time stamp T_2^*, random value r_j^*,

and c_j^*, Then MD_i computes $M_4^* = c_j^* \cdot P$, $k_{ij}^* = c_j^* \cdot M_1'$, session key $SK_{ij}^* = h(k_{ij}^* \parallel RID_{TA} \parallel T_1' \parallel T_2^*)$, and $M_5^* = h(SK_{ij}^* \parallel T_2^*)$. Finally, MD_i forwards the constructed false message $\langle M_4^*, M_5^*, T_2^* \rangle$ to MD_i'.

Step 3. After receiving the message from MD_i, MD_i' will check $|T_2' - T_2^*|? < \Delta T$ and then calculate $k_{ij}' = a_i' \cdot M_4^*$, session key $SK_{ij}' = h(k_{ij}' \parallel RID_{TA} \parallel T_1' \parallel T_2^*)$ and $M_6' = h(SK_{ij}' \parallel T_2^*)$, and obviously M_6' equals M_5^* which means that MD_i passes the verification of MD_i'. Then MD_i' computes $M_7' = h(SK_{ij}' \parallel T_3')$ and sends the message $\langle M_7', T_3' \rangle$ to MD_i.

Step 4. Once MD_i receives the message, MD_i checks $|T_3' - T_3^*|? < \Delta T$ and computes $M_8^* = h(SK_{ij}^* \parallel T_3')$, then he/she will successfully verify that M_8^* equals the received message M_7'.

At this point, MD_i and MD_i' have completed mutual authentication and negotiated the same session key ($SK_{ij}' = SK_{ij}^*$) used in future sessions. In real life, this situation is manifested as the adversary (MD_i, e.g., a doctor) successfully disguises as another patient and sends false health information to his/her attending doctor, which is easy to cause medical accident as well as being extremely harmful to the patient.

4.3. Incorrect Authentication Process. In authentication phase, U_i computes $M_1 = a_i \cdot P$ and $M_2 = a_i + k' b_i (\mathrm{mod} p)$ and then sends the message $\langle M_1, M_2, T_1 \rangle$ to CN_j. Normally, after CN_j receiving the message, she/he computes $b_i' = h(RID_{TA} \parallel T_1)$ and then judges the legality of M_2 via checking $M_2 \cdot P? = M_1 + b_i' \cdot Q$. But it is not hard to notice that the message $\langle M_1, M_2, T_1 \rangle$ does not contain the public key Q. Without knowledge of Q, CN_j cannot complete the judgement of signature, so that CN_j fails to authenticate U_i.

5. The Proposed Scheme

To correct these shortcomings in Section 4, we remedy the protocol of Wazid et al. from the following aspects. (1) In the predeployment phase, TA chooses a random value $x \in Z_p$ as the private key and computes the corresponding public key $Q_{TA} = x \cdot P$. (2) We add the fuzzy verifier to prevent the offline password guessing attack in login phase. (3) We adopt the more widely used fuzzy vault to protect biometric templates instead of fuzzy extractor.

There are also eight phases in our proposed scheme: predeployment, postdeployment, registration, login, authentication and key agreement, password and biometric update, and dynamic control node addition as well as dynamic IMD addition.

5.1. Predeployment Phase. TA first selects a secret 1024-bit number N and chooses the finite cyclic additional group G generated by a point P with a large prime order n over a finite field Z_p on an elliptic curve. Then TA selects the private key $x \in Z_p$ only known to itself, whose corresponding public key is $Q_{TA} = x \cdot P$ which is made public.

TA computes the value $TC_{CN_j} = h(ID_{CN_j} \parallel RTS_{CN_j} \parallel N)$ and stores $\{TC_{CN_j}, ID_{CN_j}\}$ in the memory of TA as well as CN_j and then adds the univariate polynomial $P(TC_{CN_j}, y)$ to the memory of CN_j.

The computing processes in predeployment phase of the IMD_l is the same as that of Wazid et al.'s scheme, so the details are omitted.

5.2. Postdeployment Phase. The specific process of this phase is as follows.

Firstly, IMD_l sends the message $\langle RID_{IMD_l} \rangle$ to CN_j; once CN_j receives the message, CN_j responds with the message $\langle TC_{CN_j} \rangle$. At the same time, they calculate the same shared secret key $SK_{IMD_l,CN_j} = P(RID_{IMD_l}, TC_{CN_j})$ and $SK_{CN_j,IMD_l} = P(RID_{IMD_l}, TC_{CN_j})$ on each own for future use.

5.3. User Registration Phase. In this phase, U_i registers with TA by executing ensuing procedure as shown in Figure 4.

Step 1. U_i inputs the selected ID_i and password PW_i and imprints the biometric BIO_i into the MD_i. MD_i chooses the private key $k \in Z_p$ and computes the corresponding public key $Q_u = k \cdot P$, as well as keeping the both secret. Finally, U_i submits the ID_i and Q_u to TA via the secure channel.

Step 2. After receiving the registration request from U_i, TA calculates $RID_i = h(ID_i \parallel x \parallel N)$ and stores specific $\{ID_i, Q_u\}$ of U_i in the memory. Then TA forwards the value RID_i to U_i.

Step 3. Upon receiving the message, MD_i chooses a random number K and calculates fuzzy vault parameters $Gen(Pol, BIO_i, K) = L$ and $Enc(CP, L) = V$ as well as $RPW_i = h(PW_i \parallel k)$ and $D_i = k \oplus h(ID_i \parallel PW_i \parallel K)$. Then, MD_i computes the verification value $T_i = h(h(ID_i \parallel RPW_i \parallel K) \mathrm{mod} l)$ where $2^8 \le l \le 2^{16}$ is a medium integer which represents the capacity of the pool of the $\langle ID_i, PW_i \rangle$ pair against the offline password guessing attack in the Wazid et al.'s scheme. After the calculation of $RID_i' = RID_i \oplus h(PW_i \parallel K \parallel k)$, MD_i stores the parameters $\{T_i, D_i, RID_i', V, Q_u, l, h(\cdot), Dec(\cdot), Rec(\cdot), Gen(\cdot), Enc(\cdot)\}$.

5.4. Login Phase. As showed in Figure 5, in this phase, U_i inputs ID_i, PW_i, and the biometric BIO_i' on the MD_i. Then U_i regains the fuzzy vault parameter K' by computing the value $Dec(BIO_i', V) = Pol'$ and $Rec(Pol') = K'$. With K', MD_i continues to calculate $k = D_i \oplus h(ID_i \parallel PW_i \parallel K')$ and $RPW_i' = h(PW_i \parallel k)$ and checks $T_i? = h(h(ID_i \parallel RPW_i' \parallel K') \mathrm{mod} l)$. If two values are not equal, MD_i refuses the login request; otherwise, MD_i believes that ID_i, PW_i, and BIO_i' are legitimate and continues to compute $RID_i = RID_i' \oplus h(PW_i \parallel K \parallel k)$. Then, MD_i generates the current timestamp T_1 and random numbers a_i and c_i. With these numbers, MD_i continues to calculate $b_i = h(RID_i \parallel T_1 \parallel ID_i \parallel ID_{CN_j})$, $M_1 = a_i \cdot P$, $M_2 = a_i + k b_i (\mathrm{mod} p)$, $M_3 = a_i \cdot Q_{TA}$, $UID_i = (ID_i \parallel ID_{CN_j}) \oplus h(M_3)$, and $PKS_i = c_i \oplus h(RID_i \parallel M_3 \parallel T_1)$.

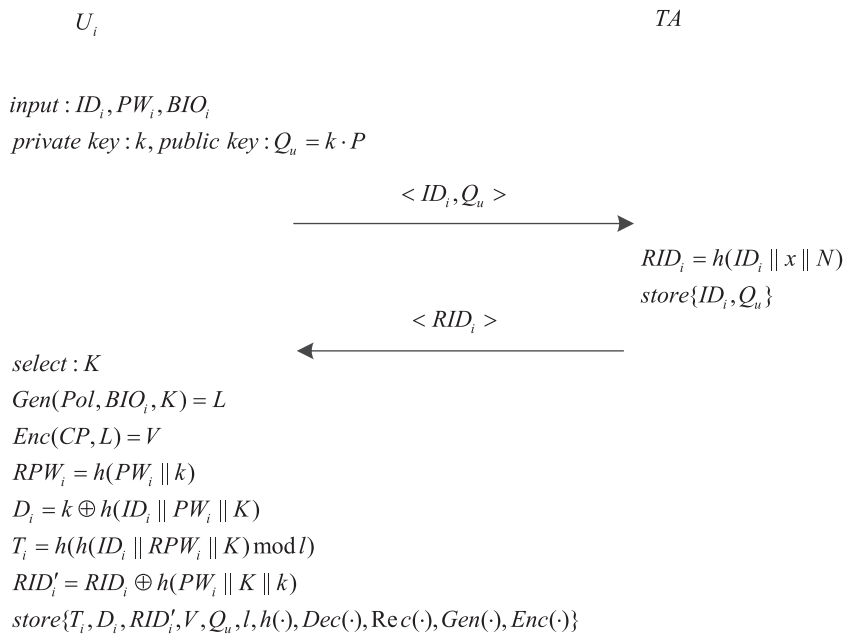


Figure 4: User registration phase of our scheme.

Finally, MD_i sends the message $\{M_1, M_2, PKS_i, T_1, UID_i\}$ to TA via a public channel.

5.5. Authentication and Key Agreement Phase. By executing following procedures, mutual authentication is established among U_i, TA, and CN_j, and a secure session key is negotiated between U_i and CN_j.

Step 1. After receiving the login request $\{M_1, M_2, PKS_i, T_1, UID_i\}$, TA first judges if $|T_1 - T_2| \leq \Delta T$ holds, where T_2 is the current timestamp and ΔT is the maximum transmission delay. If it is invalid, TA terminates the session; otherwise, TA computes the value $ID_i^* \| ID_{CN_j}^* = UID_i \oplus h(x \cdot M_1)$ and retrieves Q_u^* (i.e., the public key of U_i) corresponding to ID_i^*. Then TA computes $RID_i^* = h(ID_i^* \| x \| N)$ and $b_i^* = h(RID_i^* \| T_1 \| ID_i^* \| ID_{CN_j}^*)$ and checks the validation of the signature by checking if the equation $M_2 \cdot P = M_1 + b_i^* \cdot Q_u^*$ holds. Specifically, the equality means that TA certifies U_i's legitimacy; otherwise, TA terminates the session. Then, TA continues to calculate $c_i = PKS_i \oplus h(RID_i \| x \cdot M_1 \| T_1)$, $M_4 = h(ID_i \| TC_{CN_j} \| c_i)$, $DID_{TA} = ID_i \oplus h(UID_i \| TC_{CN_j} \| T_2)$, and $PKS_{TA} = c_i \oplus h(TC_{CN_j} \| ID_i \| T_2)$. Finally, TA sends the message $\{M_1, DID_{TA}, PKS_{TA}, UID_i, T_2, M_4\}$ to CN_j via the public channel.

Step 2. After receiving the message from TA, CN_j first checks the validation of the condition $|T_2 - T_3| \leq \Delta T$ where T_3 is the current timestamp. If it does not hold, the session is terminated here; otherwise, CN_j regains the value of ID_i and c_i by computing $ID_i^* = DID_{TA} \oplus h(UID_i \| TC_{CN_j} \| T_2)$ as well as $c_i^* = PKS_{TA} \oplus h(TC_{CN_j} \| ID_i^* \| T_2)$. Then, CN_j checks if M_4 equals the result of the computation of $h(ID_i^* \| TC_{CN_j} \| c_i^*)$. If it does not hold, CN_j terminates the session; otherwise, it means that CN_j verifies TA's legality. Then CN_j selects a random number c_j and goes on with the computation of $M_5 = c_j \cdot P$, $M_6 = c_j \cdot M_1$, the session key $SK_{ji} = h(ID_i \| ID_{CN_j} \| M_6)$, and $M_7 = h(SK_{ji} \| c_i \| T_3)$. Finally, the massage $\{M_5, T_3, M_7\}$ will be sent to U_i for authentication.

Step 3. When receiving the massage $\{M_5, T_3, M_7\}$ from CN_j, U_i will first check the validation of condition $|T_3 - T_4|? \leq \Delta T$; if it holds, U_i continues to calculate the session key $SK_{ij} = h(ID_i \| ID_{CN_j} \| a_i \cdot M_5)$ and judge if the value M_7 equals $h(SK_{ij} \| c_i \| T_3)$. The final verification shows that the mutual authentication among the U_i, TA, and CN_j is accomplished and the session key $SK_{ij} = h(ID_i \| ID_{CN_j} \| a_i \cdot M_5) = h(ID_i \| ID_{CN_j} \| a_j \cdot M_1) = SK_{ji}$ is established for future sessions.

5.6. Password and Biometric Update Phase. In this phase, we allow U_i to update the password at will by the following process, which is executed locally without involving TA for security reasons.

Step 1. First, U_i inputs her/his ID_i, PW_i^{old}, and BIO_i' on the terminal. Then MD_i calculates fuzzy vault parameters $Dec(BIO_i', V) = Pol'$ and $Rec(Pol') = K'$ and regains the private key $k' = D_i \oplus h(ID_i \| PW_i^{old} \| K')$ and $RPW_i' = h(PW_i^{old} \| k')$. MD_i checks whether T_i equals $h(h(ID_i \| RPW_i' \| K') \bmod l)$ or not. If it does not hold, MD_i rejects the request; otherwise, MD_i claims for the new PW_i^{new}.

Step 2. When U_i inputs the new password PW_i^{new}, MD_i computes $RPW_i^{new} = h(PW_i^{new} \| k)$, $D_i^{new} = k \oplus h(ID_i \| PW_i^{new} \| K)$, $T_i^{new} = h(h(ID_i \| RPW_i^{new} \| K) \bmod l)$, and $RID_i'^{new} = RID_i' \oplus h(PW_i^{new} \| K \| k) \oplus h(PW_i^{new} \| K \| k)$.

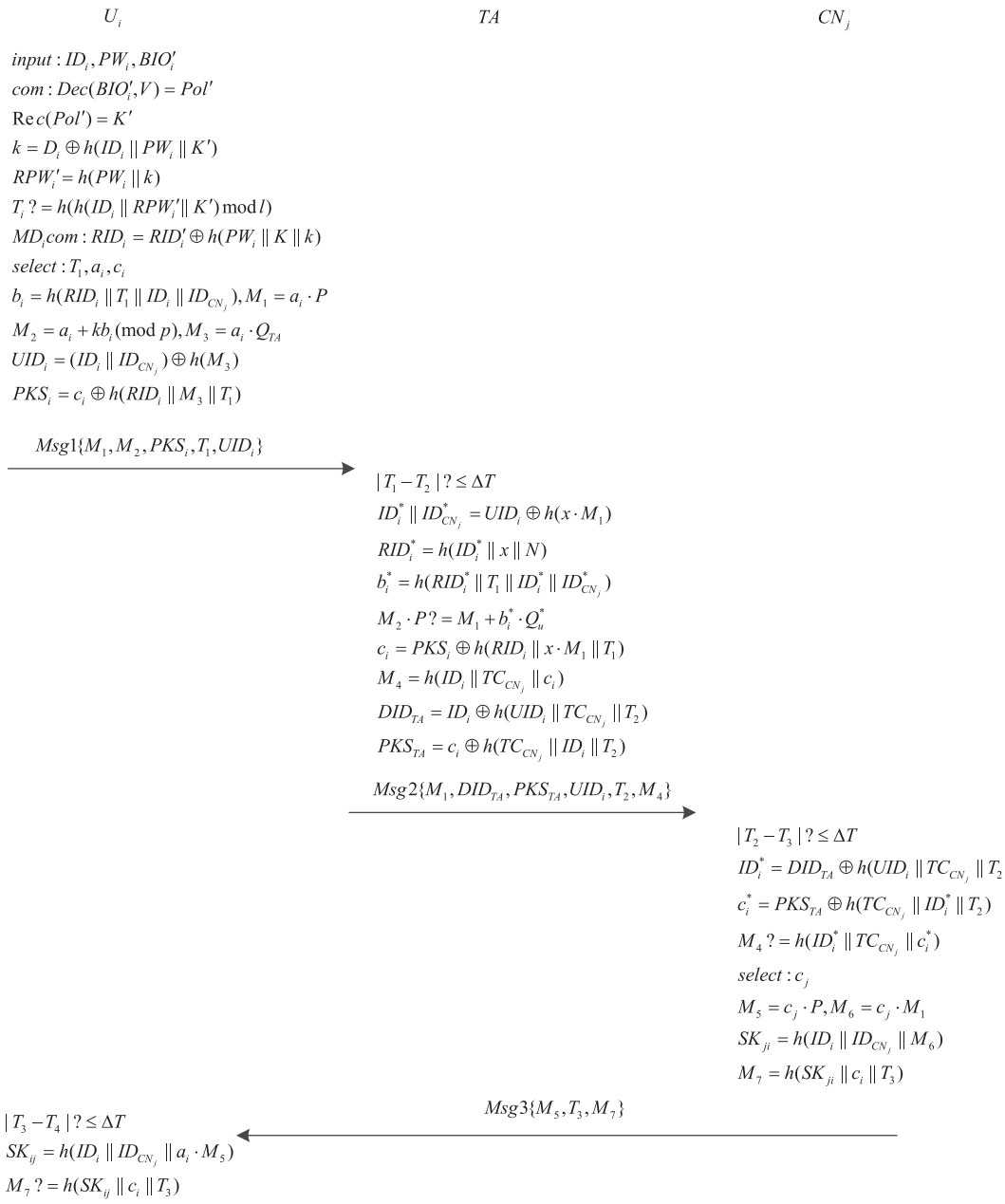


Figure 5: Login and authentication phase of our scheme.

Step 3. After the computation, MD_i updates the value of D_i^{new}, T_i^{new}, and RID_i^{lnew} in the list. Above processes simulate the situation that user only wants to update the password and maintains original biometric where $BIO_i^{new} = BIO_i$. The password and biometric update phase are summarized in Figure 6.

5.7. Dynamic Controller Node Addition Phase. In this phase, we can deploy a new control node as follows.

Step 1. TA first picks a new identity for CN_j^{new}, called $ID_{CN_j}^{new}$, then TA repeats the calculation $TC_{CN_j}^{new} = h(ID_{CN_j}^{new} \| RTS_{CN_j}^{new} \| N)$ of CN_j^{new} in the predeployment phase where $RTS_{CN_j}^{new}$ is newly generated registration timestamp. Next, TA calculates the univariate polynomial $P(TC_{CN_j}^{new}, y)$.

Step 2. Finally, TA stores the parameters $\{TC_{CN_j}^{new}, ID_{CN_j}^{new}\}$ into its memory and stores the credentials $\{TC_{CN_j}^{new}, ID_{CN_j}^{new}, P(TC_{CN_j}^{new}, y)\}$ into the memory of CN_j^{new} prior to its deployment.

5.8. Dynamic IMD Addition Phase. Depending on the real situation, the patient needs to check the state of the implantable device in time to ensure that accurate health data

User(U_i)	Mobile device(MD_i)
$input: ID_i, PW_i^{old}, BIO_i'$	$MD_i com: Dec(BIO_i', V) = Pol'$
	$\mathrm{Re}\, c(Pol') = K'$
	$k' = D_i \oplus h(ID_i \| PW_i^{old} \| K')$
	$RPW_i' = h(PW_i^{old} \| k')$
	$T_i = h(h(ID_i \| RPW_i' \| K') \bmod l)$
$input: PW_i^{new}$	$RPW_i^{new} = h(PW_i^{new} \| k)$
	$D_i^{new} = k \oplus h(ID_i \| PW_i^{new} \| K)$
	$T_i^{new} = h(h(ID_i \| RPW_i^{new} \| K) \bmod l)$
	$RID_i'^{new} = RID_i' \oplus h(PW_i^{new} \| K \| k) \oplus h(PW_i^{new} \| K \| k)$
	replace D_i, T_i and RID_i' with D_i^{new}, T_i^{new} and $RID_i'^{new}$

Figure 6: Password and biometric update phase of our scheme.

is conveyed, so we often need to replace an old IMD or add a new IMD. In the case that we use a new IMD_l^{new} to replace the existing one, please refer to Wazid et al.'s scheme for the details.

6. Security Analysis

We analyze the security of our proposed scheme in this section; it fully proves that our scheme can solve the shortcomings of Wazid et al.'s scheme and resist all kinds of known attacks. The security features such as user anonymity and forward secrecy are guaranteed in our protocol.

6.1. Security Model. Our scheme involves three interacting entities, such as U_i with $\{PW_i, BIO_i, MD_i, k\}$, CN_j with TC_{CN_j}, and TA which keeps his/her private key x. Each participant can activate multiple protocol instances and run multiple session instances in parallel. The U_i^{th} is defined as the ith instance of U_i, and the same rules apply to CN_j^{th} and TA^a. All of these instances can be seen as oracles which have three states below.

(i) Accept state: when the oracle has received the last valid message of the protocol, we can say the oracle accepts the message.

(ii) Reject state: when the oracle has received any incorrect message, the oracle will reject the received message.

(iii) $\perp$ state: when the oracle outputs no answer of the queries, we say that the oracle is in an unresponsive state which is defined as $\perp$ state.

We give the security model of our scheme, which combines the security models of [33, 45].

Definition 1 (partnering). If the instances of U_i^{th} and CN_j^{th} satisfy the following three conditions meanwhile, we determine that they are partnered to each other. (1) One of the instances is the target object of session for the other instances in the protocol, that is, the partner identification of U_i^{th} is CN_j^{th} and vice versa. (2) Both instances accept the messages mutually and negotiate the same secure session key. (3) Both instances share the same session identifier.

Definition 2 (freshness). An instance called fresh must meet the following conditions. (1) Before the instance U_i^{th} accepts the protocol run and generates the session key, neither the participants U_i^{th} nor the partners of the instance U_i^{th} are completely corrupted. (2) Neither U_i^{th} nor his/her partner instances are queried of Reveal(U_i^{th}/CN_j^{th}) by the adversary or disclose the session key.

Definition 3 (correctness). When U_i^{th} and CN_j^{th} are partnered as well as accepted, they will agree on the same session key.

Definition 4 (adversary capabilities). Interaction between the adversary A and participants in the protocol is implemented via oracle queries to simulate the abilities of attackers in reality. All oracle queries are listed as follows.

(i) Execute($U_i^{th}, CN_j^{th}, TA^a$): this oracle simulates the passive attacks (such as eavesdropping, tracking) where the adversary can get all response messages $\langle Msg1, Msg2, Msg3\rangle$ exchanged during the honest execution of authentication process.

(ii) Send($U_i^{th}/CN_j^{th}/TA^a, m$): this oracle models the active attacks where the adversary can forward a modified message m to $U_i^{th}/CN_j^{th}/TA^a$. Then he/she will get the response generated from $U_i^{th}/CN_j^{th}/TA^a$ who executes the procedure of honest protocol after receiving m. Additionally, the query Send(U_i^{th}, start) initials the protocol.

(iii) Test(U_i^{th}/CN_j^{th}): this query does not model the actual attack capabilities of adversary A but rather measures the semantic security of the session key SK. For a

participant instance U_i^{th}/CN_j^{th}, if the instance does not generate the session key, an undefined symbol $\perp$ will be returned. Otherwise, a uniform coin is thrown, if the result is 1, the true session key of the instance U_i^{th}/CN_j^{th} is returned; otherwise, a random number of the same length as the session key is returned. The adversary needs to guess the result of the toss to see whether he/she gets a real session key or a random number. Notice that the Test(U_i^{th}/CN_j^{th}) oracle query can only be used for fresh instance and up to once.

(iv) Reveal(U_i^{th}/CN_j^{th}): this oracle simulates the reveal of session key SK to adversary if U_i^{th}/CN_j^{th} really holds SK and has not been queried by a Test(U_i^{th}/CN_j^{th}) before. Otherwise the $\perp$ will be returned.

(v) Corrupt(U_i^{th}, a): this oracle query is used to model the corruption ability of the adversary; we assume A can get any one factor of U_i^{th} but not all.

If $a = 1$, it responses A with the password PW_i of U_i^{th}.

If $a = 2$, it responses A with all the security parameters stored in the MD_i of U_i^{th}.

If $a = 3$, it responses A with the biometric BIO_i of U_i^{th}.

If $a = 4$, it responses A with the private key k of U_i^{th}.

(vi) Corrupt(CN_j^{th}/TA^a): the adversary can get the long-term secret values of CN_j^{th}/TA^a, such as TC_{CN_j} of CN_j^{th} or the private key x of TA^a.

Definition 5 (random oracle). We determine the cryptographic one-way hash function H which can be accessed by all participants including A as a random oracle.

A 3FA protocol should guarantee the semantic security which is defined from Test-query. In the process run of the protocol P, A can ask the Test-query just once while other queries; i.e., Execute-query, Reveal-query, or Send-query can be asked multiple times in polynomial time. Besides, A can only make Test-query on a fresh instance. The adversary's operation is to guess the result of the coin toss in the Test-query, then we treat the event in which the adversary correctly guesses the result as a successful attack, credited as Succ(A). Only after the participants have completed the strict mutual authentication can a common session key be negotiated. The advantage of an adversary$A$ breaking the session key security of protocol P is defined as $\mathrm{Adv}_{P,D}^{ake}(A) = 2\Pr[\mathrm{Succ}(A)] - 1$ where D denotes the password space whose distribution follows a Zipf's law [50].

Theorem 6 (semantic security). *Given a 3FA protocol P, if the advantage $\mathrm{Adv}_{P,D}^{ake}(A)$ of an arbitrary PPT adversary breaking the session key security of the protocol is at most a negligible amount $n(l)$ larger than $C' \cdot q_{send}^{s'}$, then we believe that the P satisfies the semantic security, where the q_{send} denotes the number of active attacks by the PPT adversary and $n(l)$ represents a negligible function for the security parameter l.*

$$\mathrm{Adv}_{P,D}^{ake}(A) \le C' \cdot q_{send}^{s'} + n(l) \tag{1}$$

As shown above, $C' = 0.062239$ and $s' = 0.155478$ represent the Zipf parameters put forward by Wang et al. [50].

6.2. Security Proof. Assuming that DDH holds in a cyclic group, the public key encryption algorithm used in the protocol is CCA secure, and the signature algorithm is unforgeable for adaptively chosen messages. Here we prove Theorem 6 by simulating several mixing games. The mixing games start with a real attack game, and then we gradually modify the simulation rules in each game until the adversary's attack advantage to distinguish the correct session key from a random key of the same length becomes zero and then the game ends. For two adjacent mixing games, we will calculate the upper bound of the attacker's advantage gap and finally calculate the upper bound of adversary's attack on this 3FA protocol. We use Δ_i to indicate the difference between mixing games G_i and G_{i+1} and use $Adv_i(A)$ to denote the advantage of A in hybrid games G_i.

(i) G_0: this experiment is the start game which simulates the real attack mode of the adversary we demonstrate in Section 6. So, we can get

$$Adv_{P,D}^{ake}(A) = Adv_0(A) \tag{2}$$

(ii) G_1: in this game, we simulate all random oracles H in the protocol by maintaining a hash query list l_{hash}. Besides, we also simulate a private hash oracle H' by holding another list l'_{hash} which records the Hash-query directly implemented by the adversary. Obviously, the game is indistinguishable from a real one, so we have

$$\Delta_1 = \left|Adv_1(A) - Adv_1(A)\right| \le n(l) \tag{3}$$

(iii) G_2: we exclude some impossible collisions in the G_2, i.e., the collisions of messages $\langle Msg1, Msg2, Msg3\rangle$ in sessions and the collisions in the outputs of Hash-query. According to the birthday paradox, we have

$$\Delta_2 = \left|Adv_2(A) - Adv_1(A)\right| \le n(l) \tag{4}$$

(iv) G_3: we will revise the session simulation rules for the passive attacks that the adversary asks through the Execute-query. We suppose that U_i constructs the $Msg1$ using another (ID_i^*, PW_i^*) pair chosen from Cartesian product $D_{id} * D_{pw}$ instead of the real one. That is, parameters $k^* = D_i \oplus h(ID_i^* \| PW_i^* \| K)$, $RID_i^* = RID_i' \oplus h(PW_i^* \| K \| k)$, and $b_i = h(RID_i \| T_1 \| ID_i^* \| ID_{CN_j})$ are calculated and so that the signature can be calculated as $M_2 = a_i + k^* b_i^* \pmod p$. Upon receiving the message $Msg1$, TA continues to simulate session with the false identity. If TA is

lucky enough to guess the real (ID_i, PW_i), the game is terminated. The real (ID_i, PW_i) and the pseudo (ID_i^*, PW_i^*) can be seen as two challenge messages for the encryption algorithm, so the difference between the games G_3 and G_2 is at most the advantage of A breaking the encryption algorithm's CPA security of the signature. And the CPA security of the signature can be reduced to the DDH hypothesis. So, we can conclude

$$\Delta_3 = \left|Adv_3(A) - Adv_2(A)\right| \leq n(l) \tag{5}$$

(v) G_4: in this game, we continue to revise the simulation session rules in passive attacks. We use the private hashing function H' to compute the session key SK_{ij} without the Diffie-Hellman parameters a_i and c_j, that is, $SK_{ij} = H'(ID_i \parallel ID_{CN_j})$. Since we have excluded the collisions in the previous game, only A computes the valid Diffie-Hellman parameters $a_i c_j \cdot P$ and sends the query $(ID_i, ID_{CN_j}, a_i c_j \cdot P)$ to H and can A distinguish the difference between G_4 and the previous one. But the capability of A is limited by the hardness of DDH security where given g^a, g^b, g^{ab} and g^a, g^b, g^c, A cannot tell g^{ab} from g^c. Based on the intractability of the DDH problem, we have

$$\Delta_4 = \left|Adv_4(A) - Adv_3(A)\right| \leq n(l) \tag{6}$$

(vi) G_5: in this game, we start to revise the simulation session rules by active attacks. We take the Send$(TA, (Msg1))$ as the example, and if U_i is not corrupted and A correctly constructs the signature, then we say that A wins the game and terminate the simulation. Based on the unforgeability security of the signature, then we have

$$\Delta_5 = \left|Adv_5(A) - Adv_4(A)\right| \leq n(l) \tag{7}$$

(vii) G_6: we continue to revise the simulation session rules in active sessions. We acknowledge that A wins the game when A has successfully fabricated the message $Msg\{M_1, DID_{TA}, PKS_{TA}, UID_{TA}, M_4\}$ and sent it to TA. We use the private hash function H' to simulate the active sessions. The authenticator M_4 is calculated as $M_4 = H'(ID_i \parallel TC_{CN_j} \parallel c_i)$ where the c_i is randomly selected from a cyclic group. When the c_i corresponds to a fake PW_i^*, the distribution of c_i is indistinguishable from the uniform distribution on a cyclic group. Then we have

$$\Delta_6 = \left|Adv_6(A) - Adv_5(A)\right| \leq n(l) \tag{8}$$

(viii) G_7: we change the simulation rules in active sessions for the last time in this game. If A correctly forge the message $Msg3\{M_5, M_7, T_3\}$, then we say A wins the game and terminate the game. The authenticator M_7 contains the random number c_i which is unknown to A. We have eliminated this situation in previous game. So, we have

$$\Delta_7 = \left|Adv_7(A) - Adv_6(A)\right| \leq n(l) \tag{9}$$

The only way to succeed in this game is to obtain the parameters in MD_i and guess U_i's real password. A is unable to get any information of PW_i from simulation, according to the Zipf law, we get

$$Adv_8(A) \leq n(l) \leq C' \cdot q_{send}^{s'} \tag{10}$$

Therefore, Theorem 6 is proved.

6.3. Other Discussions. In this aspect, we demonstrate that our protocol can resist various known attacks as well as achieve security characteristics such as user anonymity, forward security, and key security.

6.3.1. Privileged Insider Attack. In the registration phase of our protocol, U_i sends the message consisting of the identity ID_i and corresponding public key Q_u without any knowledge of the password PW_i, so that TA has no approach to derive PW_i. Obviously, our scheme can withstand the privileged insider attack.

6.3.2. Stolen-Verifier Attack. In this attack mode, an attacker can steal the verification parameters stored by TA to cheat U_i, while we just put ID_i and Q_u in the verification table which contains no knowledge about password PW_i. Therefore, our scheme is immune to the stolen-verifier attack.

6.3.3. Offline Password Guessing Attack with Stolen Mobile Device. For this situation, we usually suppose that the A has gained the security parameters $\{T_i, D_i, RID_i', Q_u, V, l\}$ stored in the MD_i and the biometric BIO_i simultaneously; A can eavesdrop authentication messages $\langle Msg1, Msg2, Msg3\rangle$ transmitted via the public channel.

A picks a candidate $\langle ID_i^*, PW_i^*\rangle$ pair in the Cartesian product $D_{id} * D_{pw}$ and computes $Dec(BIO_i, V) = Pol$, $Rec(Pol) = K$, $k^* = D_i \oplus h(ID_i^* \parallel PW_i^* \parallel K)$, and $RPW_i^* = h(PW_i^* \parallel k)$ as well as the verification value $T_i^* = h(h(ID_i^* \parallel RPW_i^* \parallel K) \bmod l)$. In general, A can determine the chosen $\langle ID_i^*, PW_i^*\rangle$ pair's validation by checking if T_i^* equals the stored value T_i. If it holds, it means that A has guessed the correct $\langle ID_i^*, PW_i^*\rangle$ of U_i successfully; otherwise, he/she can pick another $\langle ID_i^*, PW_i^*\rangle$ pair continuing to attack. However, we introduce the fuzzy-verifier $T_i = h(h(ID_i \parallel RPW_i \parallel K) \bmod l)$ which is effective in leaving adequate candidates for A to identify and thus making it impossible for a PPT adversary to successfully guess the password.

Hence, the offline password guessing attack can not damage U_i's security.

6.3.4. Undetectable Online Password Guessing Attack. In the proposed scheme, once A tries initialing the protocol, he/she needs to make sure that the chosen password PW_i^* is valid to construct the verification signature $M_2 = a_i + kb_i \pmod p$ which will pass authentication of TA. Otherwise, the wrong PW_i^* will be observed easily by TA. So, our scheme can withstand the undetectable online password guessing attack.

6.3.5. Modification Attack. In our protocol, even A intercepts the messages transmitted in the channel, it is still impossible

for A to construct $Msg1\{M_1, M_2, PKS_i, T_1, UID_i\}$, $Msg2\{M_1, DID_{TA}, PKS_{TA}, UID_i, T_2, M_4\}$, and $Msg3\{M_5, T_3, M_7\}$ which are protected by the secret value, private key or hash functions to pass the message verification. For example, in $Msg1$ A is unable to calculate the value $M_2 = a_i + kb_i(\mathrm{mod}p)$, since $b_i = h(RID_i \parallel T_1 \parallel ID_i \parallel ID_{CN_j})$ where $RID_i = RID_i' \oplus h(PW_i \parallel K \parallel k) = h(ID_i \parallel x \parallel N)$ consists of secret values only known to U_i or TA such as PW_i, private key k, and x, so that A's login request will be rejected by TA. Similarly, A cannot construct the valid verification parameters M_4 without knowledge of TC_{CN_j} or M_7 due to the hardness of ECCDH problem introduced in Section 2.2. Thus, all modified messages will be detected and rejected by receiver simultaneously.

In conclusion, modification attack is impossible in our scheme.

6.3.6. User Impersonation Attack. We suppose that A plans to impersonate as a legitimate user U_i to interact with TA. The key step is to construct a valid value M_2 to pass the verification of TA. However, A is unable to calculate $M_2 = a_i + kb_i(\mathrm{mod}p)$ without b_i. To get $b_i = h(RID_i \parallel T_1 \parallel ID_i \parallel ID_{CN_j})$, he/she needs to know the most of long-term values. Therefore, our proposed scheme is immune to the user impersonation attack.

6.3.7. Control Node Impersonation Attack. We have analyzed that the malicious MD_i may successfully impersonate CN_j to cheat another MD_i^* in Wazid et al.'s scheme. On the one hand, both MD_i and CN_j hold the same parameter RID_{TA} which composes the correct verification value $M_5 = h(SK_{ij} \parallel T_2)$ and $SK_{ij} = h(k_{ij} \parallel RID_{TA} \parallel T_1 \parallel T_2)$. On the other hand, in Wazid et al.'s scheme, the essential parameter c_j is not verified when it is sent to MD_i. But in our scheme, this attack mode cannot be implemented, and the malicious MD_i is unable to fabricate M_7 without knowing c_i of MD_i^*, so we solve the potential pitfall in Wazid et al.'s scheme.

From another point of view, an adversary A cannot construct the verification value M_7 due to the hardness of ECCDH, so A fails to impersonate a CN_j. In a word, the control node impersonation attack has no threat to our protocol.

6.3.8. TA Impersonation Attack. For A, it is computationally infeasible to get the value $M_4 = h(ID_i \parallel TC_{CN_j} \parallel c_i)$ which is protected by hash function and critical parameters TC_{CN_j} as well as nonce c_i. The c_i can be derived from two functions as $c_i = PKS_{TA} \oplus h(TC_{CN_j} \parallel ID_i \parallel T_2) = PKS_i \oplus h(RID_i \parallel M_5 \parallel T_1)$, but even A has intercepted the parameters PKS_{TA}, UID_i, and PKS_i; he/she still cannot calculate c_i without RID_{CN_j}, RID_i, or M_3, and then M_4 cannot be computed. In short, our scheme is immune to the TA impersonation attack.

6.3.9. Denial-of-Service (DoS) Attack. Before U_i's login request is sent to TA, the password PW_i^*, identity ID_i^*, and biometric BIO_i^* input in the terminal by U_i will be determined locally by verifying the value of T_i^*. According to the protocol, only when $T_i^* = T_i$, the process will continue. Hence, our protocol can withstand such an attack.

6.3.10. Replay Attack. When an adversary A wants to send the intercepted messages $\langle Msg1, Msg2, Msg3\rangle$ to receiver again, it will fail to pass the protection of timestamp $\langle T_1, T_2, T_3, T_4\rangle$. All these intercepted messages will be seen overdue. So, our scheme can withstand this attack effectively.

6.3.11. Mutual Authentication. Mutual authentication means that before the doctor gets health information from CN_j, U_i, TA, and CN_j have confirmed the legitimacy of the other two parties. In our protocol, TA holds the public key Q_u to verify the signature M_2, and then U_i is authenticated. In the same way, we take the verification values M_4 and M_7 which consist of some parameters only known to them just like private key or nonce to accomplish mutual authentication. That is, when they affirm that each other is legal, a secure session key is negotiated between U_i and CN_j.

6.3.12. Known Key Security. Our entire protocol's purpose is to ensure the safety of subsequent medical information delivery after mutual authentication is completed. The session key $SK_{ij} = h(ID_i \parallel ID_{CN_j} \parallel a_i c_j \cdot P)$ which depends on random numbers a_i and c_j can be different and independent in every key agreement phase. Even some session keys are disclosed, in the next session, the SK_{ij} will maintain secure. Hence, our protocol guarantees the security of the session key.

6.3.13. Perfect forward Secrecy. At the final step of authentication phase, U_i and CN_j negotiate a session key $SK_{ij} = h(ID_i \parallel ID_{CN_j} \parallel a_i c_j \cdot P) = h(ID_i \parallel ID_{CN_j} \parallel a_i \cdot M_5) = h(ID_i \parallel ID_{CN_j} \parallel c_j \cdot M_1)$. To calculate the session key with $M_1 = a_i \cdot P$, A has to solve the ECCDH problem as we showed before. It follows that even long-term keys of U_i and CN_j are disclosed, the session key still maintains secure. Hence, the proposed protocol achieves perfect forward secrecy.

6.3.14. User Anonymity. In the proposed protocol, we conceal the identity ID_i in the $b_i = h(RID_i \parallel T_1 \parallel ID_i \parallel ID_{CN_j})$, $UID_i = (ID_i \parallel ID_{CN_j}) \oplus h(M_3)$, and $DID_{TA} = ID_i \oplus h(UID_i \parallel TC_{CN_j} \parallel T_2)$. It shows that ID_i is protected by private key x in $RID_i = h(ID_i \parallel x \parallel N)$, nonce a_i in $M_3 = a_i \cdot Q_{TA}$. That means in addition to the U_i, TA, and CN_j, no one knows the ID_i. So, our scheme achieves user anonymity.

6.3.15. User Untraceability. In the proposed protocol, messages $Msg1\{M_1, M_2, PKS_i, T_1, UID_i\}$, $Msg2\{M_1, DID_{TA}, PKS_{TA}, UID_i, T_2, M_4\}$, and $Msg3\{M_5, T_3, M_7\}$ transmitted among U_i, TA, and CN_j are dynamic and different from before ones because the sender randomly selects a number to compose messages. For instance, in $Msg1$, the introductions of a_i and c_j make the parameters different for each login phase to prevent A from using static values to track user. In short, it is impossible for A to track U_i in our scheme.

6.3.16. Biometric Template Privacy. Our scheme can effectively maintain the privacy of biometric BIO_i. On the one hand, user does not offer CN_j the biometric template, and there is no knowledge about U_i's biometric template in the

Table 2: Comparison of security features.

	Wang et al.'s scheme [13]	Wazid et al.'s scheme [14]	Our scheme
Mutual authentication	×	√	√
Known key security	√	√	√
Perfect forward secrecy	√	√	√
User anonymity	√	√	√
Biometric template privacy	√	√	√
Resisting modification attack	√	√	√
Resisting user impersonation attack	×	√	√
Resisting server(CN) impersonation attack	√	×	√
Resisting man-in-the-middle attack	√	√	√
Resisting stolen-verifier attack	√	√	√
Resisting privileged insider attack	√	√	√
Resisting replay attack	√	√	√
Resisting modification attack	√	√	√
Resisting password guessing attack	√	×	√
Resisting secure key agreement	×	×	√

Table 3: Comparison of computation cost.

scheme	Wang et al.'s [13]	Wazid et al.'s [14]	Ours
$U_i(MD_i)$	$T_{bp} + 3T_{em} + 5T_H + T_{se}$	$3T_{em} + 12T_H + T_{as} + T_{me}$	$3T_{em} + 10T_H + T_{as} + T_{me}$
TA	-	-	$4T_{em} + 7T_H + T_{as}$
CN_j	$T_{bp} + 2T_{em} + 5T_H + T_{se}$	$4T_{em} + 5T_H + T_{as}$	$2T_{em} + 5T_H$
Overall	$2T_{bp} + 5T_{em} + 10T_H + 2T_{se}$	$7T_{em} + 17T_H + 2T_{as} + T_{me}$	$9T_{em} + 22T_H + 2T_{as} + T_{me}$

memory of CN_j. On the other hand, we firstly use fuzzy vault to convert the form of biometric template to V. Even A obtains the V form MD_i, he/she still cannot recover the biometric template because the algorithms of fuzzy vault are one-way operations. Moreover, the biometric template itself is difficult to lose or falsify. In short, our protocol guarantees the privacy of biometric template.

7. Features and Efficiency Comparison

This section shows the comparisons of our scheme and other two related works (Wang et al. [13], Wazid et al. [14]) in efficiency and the advantages/disadvantages showed in Tables 3 and 2, respectively. Specifically, we analyze the computation cost from the point of time complexity to compare the efficiency. What needs to be explained is that we only focus on the login and authentication phases and ignore the bit-XOR operation due to its low computation consumption. Besides, we use the symbols of T_{em}, T_H, T_{bp}, T_{se}, T_{me}, and T_{as} to represent the time cost of elliptic curve point multiplication, hash function, bilinear pairing, symmetric key encryption/decryption, modular exponentiation, and asymmetric key encryption/decryption, respectively.

From Tables 2 and 3, it could be seen that although the calculation cost of our scheme is a little higher than the other two solutions, we have greatly satisfied various security standards in terms of security, which is superior to Wang et al.'s protocol [13] in resisting impersonation attack and achieving mutual authentication. And our scheme makes up for the flaws we analyzed in Wazid et al.'s protocol [14]. In general, our protocol is more suitable for use in implantable medical system, within the acceptable computational energy consumption of the devices.

8. Conclusion

We take the most recent scheme of Wazid et al. as a typical example to show the subtlety of the design of 3FA for the implantable medical system. We have found that the scheme cannot resist three types of drawbacks, i.e., password guessing attack, controller node impersonation attack, and the incorrect authentication process. Then we have presented a trusted authority assisted 3FA protocol for the implantable medical system. Specifically, we have made the following amendments. TA is introduced in the authentication phase of the newly proposed solution. We have also replaced fuzzy extractor with the more widely applied fuzzy vault to the biometrics. The new protocol is provably secure under DDH assumption; the efficiency comparison and features analysis indicate that while a little efficiency is sacrificed, our protocol satisfies all the required security features. Overall, our new protocol is suitable for use in the implantable medical system.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

All the authors have contributed equally to this work.

Acknowledgments

This work is supported by National Natural Science Foundation of China (no. 61672433) and Basic Research Project from Science and Innovation Council of Shenzhen (nos. 201703063000511 and 201703063000517).

References

[1] R. Altawy and A. M. Youssef, "Security tradeoffs in cyber physical systems: a case study survey on implantable medical devices," *IEEE Access*, vol. 4, pp. 959–979, 2016.

[2] M. Rushanan, A. D. Rubin, D. F. Kune, and C. M. Swanson, "SoK: security and privacy in implantable medical devices and body area networks," in *Proceedings of the 35th IEEE Symposium on Security and Privacy (SP '14)*, pp. 524–539, San Jose, Calif , USA, May 2014.

[3] P. K. Sahoo, "Efficient security mechanisms for mhealth applications using wireless body sensor networks," *Sensors*, vol. 12, no. 9, pp. 12606–12633, 2012.

[4] D. Wu, J. Yan, H. Wang, D. Wu, and R. Wang, "Social Attribute Aware Incentive Mechanism for Device-to-Device Video Distribution," *IEEE Transactions on Multimedia*, vol. 19, no. 8, pp. 1908–1920, 2017.

[5] J. Xiong, Y. Zhang, L. Lin et al., "ms-PoSW: A multi-server aided proof of shared ownership scheme for secure deduplication in cloud," *Concurrency & Computation Practice & Experience*, no. 5, Article ID e4252, 2017.

[6] S. Kumari, X. Li, F. Wu, A. K. Das, K.-K. R. Choo, and J. Shen, "Design of a provably secure biometrics-based multi-cloud-server authentication scheme," *Future Generation Computer Systems*, vol. 68, pp. 320–330, 2017.

[7] G. Zheng, R. Shankaran, M. A. Orgun, L. Qiao, and K. Saleem, "Ideas and Challenges for Securing Wireless Implantable Medical Devices: A Review," *IEEE Sensors Journal*, vol. 17, no. 3, pp. 562–576, 2017.

[8] D. Wu, S. Si, S. Wu, and R. Wang, "Dynamic trust relationships aware data privacy protection in mobile crowd-sensing," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1-1, 2017.

[9] D. Wu, F. Zhang, H. Wang, and R. Wang, "Security-oriented opportunistic data forwarding in Mobile Social Networks," *Future Generation Computer Systems*, vol. 87, pp. 803–815, 2018.

[10] D. Dolev and A. C.-C. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 198–208, 1983.

[11] L. Wu, X. Du, M. Guizani, and A. Mohamed, "Access Control Schemes for Implantable Medical Devices: A Survey," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1272–1283, 2017.

[12] E. Marin, D. Singelée, F. D. Garcia, T. Chothia, R. Willems, and B. Preneel, "On the (in)security of the latest generation implantable cardiac defibrillators and how to secure them," in *Proceedings of the 32nd Annual Computer Security Applications Conference, ACSAC 2016*, pp. 226–236, Los Angeles, Calif, USA, December 2016.

[13] C. Wang and Y. Zhang, "New authentication scheme for wireless body area networks using the bilinear pairing," *Journal of Medical Systems*, vol. 39, no. 11, article 136, 2015.

[14] M. Wazid, A. K. Das, N. Kumar, M. Conti, and A. V. Vasilakos, "A novel authentication and key agreement scheme for implantable medical devices deployment," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 4, 2018.

[15] A. Perrig, "The tesla broadcast authentication protocol," *Rsa Cryptobytes*, vol. 20, no. 2, p. 2002, 2005.

[16] C.-C. Lee, C.-W. Hsu, Y.-M. Lai, and A. Vasilakos, "An enhanced mobile-healthcare emergency system based on extended chaotic maps," *Journal of Medical Systems*, vol. 37, no. 5, article 9973, 2013.

[17] F. Wu, X. Li, L. Xu et al., "A lightweight and privacy-preserving mutual authentication scheme for wearable devices assisted by cloud server," *Computers & Electrical Engineering*, vol. 63, pp. 168–181, 2017.

[18] X. Li, M. H. Ibrahim, S. Kumari, A. K. Sangaiah, V. Gupta, and K. R. Choo, "Anonymous mutual authentication and key agreement scheme for wearable sensors in wireless body area networks," *Computer Networks*, 2017.

[19] S. Kumari, X. Li, F. Wu, A. K. Das, H. Arshad, and M. K. Khan, "A user friendly mutual authentication and key agreement scheme for wireless sensor networks using chaotic maps," *Future Generation Computer Systems*, vol. 63, pp. 56–75, 2016.

[20] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE Transactions on Information Theory*, vol. 31, no. 4, pp. 469–472, 1985.

[21] R. L. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," *Communications of the ACM*, vol. 26, no. 1, pp. 96–99, 1983.

[22] V. S. Miller, "Use of elliptic curves in cryptography," in *Advances in Cryptology—CRYPTO'85*, H. C. Williams, Ed., vol. 218 of *Lecture Notes in Computer Science*, pp. 417–426, Springer, 1986.

[23] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 177, pp. 203–209, 1987.

[24] J. Shen, S. Chang, J. Shen, Q. Liu, and X. Sun, "A lightweight multi-layer authentication protocol for wireless body area networks," *Future Generation Computer Systems*, 2016.

[25] J. Liu, Z. Zhang, X. Chen, and K. S. Kwak, "Certificateless remote anonymous authentication schemes for wireless body area networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 332–342, 2014.

[26] H. Xiong, "Cost-effective scalable and anonymous certificateless remote authentication protocol," *IEEE Transactions on Information Forensics & Security*, vol. 9, no. 12, pp. 2327–2339, 2014.

[27] B. Hu, D. Wong, Z. Zhang, and X. Deng, "Key replacement attack against a generic construction of certificateless signature," in *Information Security and Privacy*, vol. 4058 of *Lecture Notes in Computer Science*, pp. 235–246, Springer, Berlin, Germany, 2006.

[28] D. He, S. Zeadally, N. Kumar, and J. H. Lee, "Anonymous authentication for wireless body area networks with provable security," *IEEE Systems Journal*, vol. 11, no. 4, pp. 1–12, 2016.

[29] X. Li, M. H. Ibrahim, S. Kumari, and R. Kumar, "Secure and efficient anonymous authentication scheme for three-tier mobile healthcare systems with wearable sensors," *Telecommunication Systems*, vol. 67, no. 3, pp. 1–26, 2018.

[30] C.-T. Li, C.-C. Lee, C.-Y. Weng, and S.-J. Chen, "A Secure dynamic identity and chaotic maps based user authentication and key agreement scheme for e-Healthcare systems," *Journal of Medical Systems*, vol. 40, no. 11, article 233, 2016.

[31] Q. Jiang, J. Ma, and C. Yang, "Efficient end-to-end authentication protocol for wearable health monitoring systems," *Computers Electrical Engineering*, 2017.

[32] 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.

[33] F. Wei, P. Vijayakumar, J. Shen, R. Zhang, and L. Li, "A provably secure password-based anonymous authentication scheme for wireless body area networks," *Computers and Electrical Engineering*, vol. 65, pp. 322–331, 2018.

[34] 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.

[35] D. He, N. Kumar, J. Chen, C.-C. Lee, N. Chilamkurti, and S.-S. Yeo, "Robust anonymous authentication protocol for health-care applications using wireless medical sensor networks," *Multimedia Systems*, vol. 21, no. 1, pp. 49–60, 2015.

[36] S. Kumari, M. K. Khan, and M. Atiquzzaman, "User authentication schemes for wireless sensor networks: A review," *Ad Hoc Networks*, vol. 27, pp. 159–194, 2015.

[37] T. Chen, C. Lee, M. Hwang et al., "Towards secure and efficient user authentication scheme using smart card for multi-server environments," *The Journal of Supercomputing*, vol. 66, no. 2, pp. 1008–1032, 2013.

[38] C.-T. Li, C.-C. Lee, C.-Y. Weng, and C.-I. Fan, "An extended multi-server-based user authentication and key agreement scheme with user anonymity," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 1, pp. 119–131, 2013.

[39] F.-S. Wei, Q. Jiang, R.-J. Zhang, and C.-G. Ma, "A privacy-preserving multi-factor authenticated key exchange protocol with provable security for cloud computing," *Journal of Information Science and Engineering*, vol. 33, no. 4, pp. 907–921, 2017.

[40] X. Li, J. Niu, S. Kumari et al., "A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments," *Journal of Network & Computer Applications*, vol. 103, no. 1, pp. 194–204, 2018.

[41] Q. Jiang, Z. Chen, B. Li et al., "Security analysis and improvement of bio-hashing based three-factor authentication scheme for telecare medical information systems," *Journal of Ambient Intelligence & Humanized Computing*, pp. 1–13, 2017.

[42] Q. Jiang, S. Zeadally, J. Ma, and D. He, "Lightweight three-factor authentication and key agreement protocol for internet-integrated wireless sensor networks," *IEEE Access*, vol. 5, pp. 3376–3392, 2017.

[43] F. Wu, L. Xu, S. Kumari, and X. Li, "An improved and provably secure three-factor user authentication scheme for wireless sensor networks," *Peer-to-Peer Networking and Applications*, vol. 11, no. 5, pp. 1–20, 2016.

[44] C. Lee, C. Chen, P. Wu, and T. Chen, "Three-factor control protocol based on elliptic curve cryptosystem for universal serial bus mass storage devices," *IET Computers & Digital Techniques*, vol. 7, no. 1, pp. 48–55, 2013.

[45] S. Yin, X. Li, H. Gao, and O. Kaynak, "Data-based techniques focused on modern industry: an overview," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 657–667, 2015.

[46] A. Juels and M. Sudan, "A fuzzy vault scheme," *Designs, Codes and Cryptography. An International Journal*, vol. 38, no. 2, pp. 237–257, 2006.

[47] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: how to generate strong keys from biometrics and other noisy data," *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.

[48] J. Yu, G. Wang, Y. Mu, and W. Gao, "An efficient generic framework for three-factor authentication with provably secure instantiation," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2302–2313, 2014.

[49] C. Blundo, A. D. Santis, A. Herzberg, S. Kutten, U. Vaccaro, and M. Yung, "Perfectly secure key distribution for dynamic conferences," in *Advances in Cryptology —CRYPTO' 92*, vol. 740 of *Lecture Notes in Computer Science*, pp. 471–486, 1993.

[50] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, "Zipfs law in passwords," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 3295148, 12 pages
https://doi.org/10.1155/2018/3295148

Research Article

Multidevice Authentication with Strong Privacy Protection

Jan Hajny, Petr Dzurenda, and Lukas Malina

Brno University of Technology, Czech Republic

Correspondence should be addressed to Jan Hajny; hajny@feec.vutbr.cz

Received 12 April 2018; Revised 20 June 2018; Accepted 3 July 2018; Published 29 July 2018

Academic Editor: Kim-Kwang Raymond Choo

Card-based physical access control systems are used by most people on a daily basis, for example, at work, in public transportation, or at hotels. Yet these systems have often very poor cryptographic protection. User identifiers and keys can be easily eavesdropped on and counterfeited. The privacy-preserving features are almost missing in these systems. To improve this state, we propose a novel cryptographic scheme based on efficient zero-knowledge proofs and Boneh-Boyen signatures. The proposed scheme is provably secure and provides the full set of privacy-enhancing features, that is, the anonymity, untraceability, and unlinkability of users. Furthermore, our scheme supports distributed multidevice authentication with multiple RFID (Radio-Frequency IDentification) user devices. This feature is particularly important in applications for controlling access to dangerous sites where the presence of protective equipment is checked during each access control session. Besides the full cryptographic specification, we also show the results of our implementation on devices commonly used in access control applications, particularly the smart cards and embedded verification terminals. By avoiding costly operations on user devices, such as bilinear pairings, we were able to achieve times comparable to existing systems (around 500 ms), while providing significantly higher security, privacy protection, and features for RFID multidevice authentication.

1. Introduction

Privacy-enhancing technologies constitute a significant part of contemporary cryptography. Modern cryptographic protocols allow privacy-enhanced storing of sensitive data and its processing by cloud services, private information retrieval, or, for example, authentication based on personal attributes, instead of user identifiers. The increasing intensity of research into privacy is supported by national programs and strategies, in particular in USA [1] and EU [2]. While most of the novel schemes are aimed at electronic services, the domain of physical access control is rather neglected. We still use traditional locks, tourniquets, and classical card-based access control mechanisms to manage physical access to our premises. But with the increasing computational power of the programmable smart cards, massive expansion of various personal electronic devices, and the capabilities in RFID communication of our smart phones, we can expect penetration of privacy-enhancing technologies also to the area of physical access control. In particular, in mass applications like public transportation, e-ticketing, e-passports, and eIDs, the benefits of controlling physical access using electronic devices with advanced cryptographic protocols are very appealing.

In this paper, we propose and experimentally evaluate a novel cryptographic scheme that particularly addresses two phenomenons of contemporary cyberspace: lack of user privacy and ubiquitous presence of many personal devices (phones, smart cards, RFID tags, bluetooth dongles, smart watch, etc.) that can be leveraged for stronger authentication and more reliable access control.

In particular, we focus on safety applications in which the users wear multiple safety equipment, such as helmets, harnesses, boots, and protective suits, each with attached programmable RFID tag capable of wireless communication. A user is granted access to (potentially dangerous) premises only if all his equipment is present. In existing systems, the presence of the protective equipment is checked simply by scanning the identifiers using RFID readers. Such an approach is neither secure (identifiers can be counterfeited), nor privacy friendly (identifiers can be traced, behavioral profiles can be created, etc.).

We propose a novel cryptographic scheme for multidevice authentication that is tailored for physical access control systems where the user must prove not only his own identifier, but also *many other auxiliary identifiers* stored on separate devices. In addition, the authentication sessions must support all the key privacy-enhancing features; i.e., the access control process must be *anonymous* (i.e., a user must prove that he belongs to a group of authorized users, but without releasing his concrete identity), *unlinkable* (all the sessions of a single user cannot be linkable to a profile), and *untraceable* (system administrators must be unable to trace honest users in the system). On the other side, the scheme must provide efficient means for revocation and identification of malicious users. In our cryptographic system, we provide all the required features that are often contradictory and completely unavailable in existing schemes (in particular, the presence of many identifiers versus anonymity; the untraceability and strong cryptographic security versus efficiency on RFID tags and stickers).

In our scheme, users can be granted access to premises upon proving the presence of particular devices in their proximity (e.g., the safety equipment) or personal attributes (age, membership, citizenship, etc.). The access control process may (the extent of privacy-enhancing features can be initially set by the administrator; if required, identification or user tracing may be enforced by the access control system) proceed in a fully private manner, without disclosing user identity or being traceable in the system.

2. Related Work

Most of the existing practical physical access control systems are based on the following technologies: NXP's Mifare and DESfire; HID's Prox and iClass; and Legic Prime and Advant. NXP's Mifare Classic, introduced in 1994, is a very popular technology used in physical access control systems. Although very old and insecure, the technology is still used in many applications, even those security sensitive. The authentication protocol is based on a unique 4B card identifier UID. In some implementations, the card just reveals UID to the terminal without any authentication protocol. In that case, UID can be easily eavesdropped on and used by an attacker for impersonation. In other implementations, a simple authentication protocol is used but is considered insecure due to many existing practical attacks [3–5] on the encryption algorithm CRYPTO1. The insufficient security of the CRYPTO1 algorithm used in the Mifare Classic made NXP improve the cryptographic protection and release Mifare DESFire. The old encryption algorithm was replaced by 3DES algorithm. The authentication protocol was further improved in Mifare DESFire EV1 which supports the AES encryption algorithm [6]. The protocol itself remained without any major changes. However, even Mifare DESFire was successfully attacked, although the attacks [7, 8] were aimed on the implementation, not cryptographic weaknesses. The HID Prox technology contains no cryptographic protection. HID iClass employs an authentication protocol based on the 3DES algorithm, but attacks on this protocol are available [9]. Legic Prime has weak proprietary cryptographic protection [10]. Legic Advant is protected by symmetric block algorithms (DES [11], 3DES, and AES). None of the major commercial technologies provide any protection of privacy.

With the introduction of the first attribute-based credential schemes, such as the Idemix [12], U-Prove [13], and HM12 [14], the variants for physical access control systems also started to appear. The U-Prove scheme was implemented on Multos smart cards [15]. The user is able to prove his attribute in less than 1 s using this implementation. However, the unlinkability property cannot be provided by the cryptographic design of the protocol. The Idemix was also implemented on the Multos smart card platform, with cca 1 s needed to generate the attribute proof. The pilot implementation of the HM12 scheme using Multos ML3 smart cards [16] required around 2.4 s in total to generate and verify the proof, including the communication overhead. No testing was done on multiple devices because the distributed proof is not supported by these schemes.

Many types of personal and wearable devices forming the so-called Internet of Things (IoT) have appeared recently. Authentication issues have been solved by different techniques on these devises. Xu and Weitao propose biometric authentication using wearables with face recognition using smart-glass and gait recognition using smart watch. Riva et al. [17] combine multiple sources of authentication data, which is close to our approach. However, all these schemes are using mainly biometric authentication factors. Cha et al. [18] present a simple model for two-device authentication for micropayment systems using mobile and wearable devices. Nevertheless, their proposal lacks details and concrete cryptographic functions. Butun et al. [19] address multilevel authentication issue in cloud computing. Gonzalez-Manzano et al. [20] present an access control mechanism for cloud-based storage service access by using a set of devices. However, their scheme is based on symmetric cryptography and thus does not provide nonrepudiation. Hajny et al. [21] use many wearable and IoT devices to do the authentication process. However, the scheme misses privacy-enhancing properties, because each user is uniquely represented by his/her public key.

In summary, there are several authentication solutions that involve IoT devices. However, there are only very few papers focusing on multidevice authentication. Currently, none of the proposals is provably secure and supports the privacy-enhancing features. Furthermore, most of the schemes remain only theoretic.

2.1. Our Contribution. The cryptographic scheme presented in this paper takes a novel approach for the access control based on rather the presence of multiple devices in user's proximity than the direct verification of user identifiers. The novel approach has two key benefits: it significantly improves the privacy protection of users and allows the authentication based on the presence of many low-performance devices. Our scheme is the first practical proposal with implementation results that combines strong security, all standard privacy-enhancing features, and efficiency.

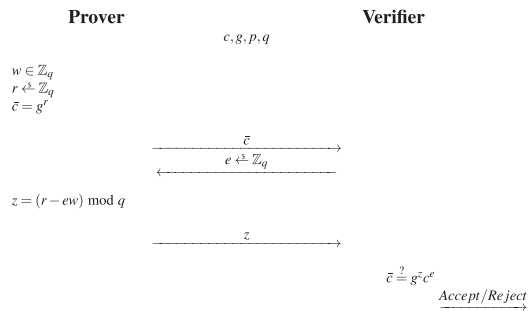


FIGURE 1: Schnorr's proof of knowledge of discrete logarithm PK $\{w : c = g^w\}$ in $\mathbb{Z}_p^*$.

(i) *Provable security*: all algorithms are provably secure, based on primitives with rigorous formal proofs.

(ii) *Multidevice authentication*: the scheme allows user authentication based on the presence of many personal devices.

(iii) *Anonymity*: the scheme allows authentication based on anonymous proofs of knowledge of private user and/or device identifiers.

(iv) *Unlinkability*: the scheme prevents creating user behavior profiles based on the authentication sessions linking.

(v) *Untraceability*: the scheme prevents any entity from tracing users (or their devices).

(vi) *Efficiency*: the authentication protocol is fast on constrained user devices (i.e., smart cards) and embedded verification terminals.

(vii) *Revocation and identification*: the proposed scheme is compatible with major revocation and identification schemes [22–24] for attribute-based credentials.

We not only provide the cryptographic description and security proofs of our scheme, but also provide practical implementation results based on benchmarks on RFID devices and an embedded hardware terminal. These results prove that the scheme can be practically implemented on existing off-the-shelf devices.

3. Preliminaries

3.1. Notation. We describe proof of knowledge protocols (PK) using the efficient notation introduced by Camenisch and Stadler [25]. The protocol for proving the knowledge of discrete logarithm of c with respect to g is denoted as $\mathrm{PK}\{\alpha : c = g^\alpha\}$. The symbol ":" means "such that" and "$|x|$" is the bit length of x. We write $a \xleftarrow{\$} A$ when a is sampled uniformly at random from A. A secure hash function is denoted as $\mathscr{H}$.

3.2. Proofs of Knowledge. The statements about discrete logarithms in prime order groups can be easily proven using the Σ-protocols [26].

A simple yet very useful protocol for proving the discrete logarithm knowledge is based on the Schnorr signature scheme [27]. Using this protocol, the prover proves his knowledge of a discrete logarithm with respect to public parameters c, g, p, q; i.e., he proves the knowledge of $w : c = g^w \bmod p$, where p is prime modulus, q is group order, and g is $\mathbb{Z}_p^*$ generator. The protocol is depicted in Figure 1.

The proof of discrete logarithm knowledge is a simple 3-way protocol where the prover commits to a random number r in the first step, receives a challenge e in the second step, and responds by z to the challenge in the third step. The protocol is Honest Verifier Zero-Knowledge (HVZK). Note that the verifier does not have to know the private input w of the prover to be able to verify its knowledge. We recall the properties of the protocol below.

Proof. Completeness: prover who knows w is always accepted: $\bar{c} = g^z c^e = g^{r-ew} g^{ew} = g^r g^{-ew} g^{ew} = g^r = \bar{c}$. □

Proof. Soundness: let us assume a cheating prover is ready to answer at least 2 random challenges e, e' after committing to r without knowing w. Then, his responses z, z' must be accepted in verifier's checks:

$$\bar{c} = g^z c^e, \tag{1}$$

$$\bar{c} = g^{z'} c^{e'}, \tag{2}$$

we divide (1) and (2) and get

$$\frac{\bar{c}}{\bar{c}} = \frac{g^z c^e}{g^{z'} c^{e'}} \Longleftrightarrow \quad 1 = g^{z-z'} c^{e-e'}, \tag{3}$$

after multiplying both sides of (3) by $g^{-(z-z')}$ and raising to the power of $(e - e')^{-1}$, we get

$$g^{(z'-z)(e-e')^{-1}} = c \tag{4}$$

and we get the discrete logarithm $w = (z' - z)(e - e')^{-1}$ that is easy to efficiently compute for the dishonest prover; thus we reached the contradiction because the cheating prover unaware of w was assumed. □

Proof. Honest Verifier Zero-Knowledge: the ZK property is proven by proving the existence of the following ZK simulator M_V^*:

(1) the simulator randomly chooses the response $z' \overset{\$}{\leftarrow} \mathbb{Z}_q$.

(2) the simulator randomly chooses $e' \overset{\$}{\leftarrow} \mathbb{Z}_q$.

(3) the simulator computes $\overline{c'} = g^{z'} c^{e'}$.

□

The M_V^*'s output $\overline{c}', e', z'$ is computationally indistinguishable from the real protocol output c, e, z.

The protocol for proving the knowledge of a discrete logarithm described above can be extended to the discrete logarithm representation proof and discrete logarithm equivalence proof [25]. The principles and security proofs remain the same.

3.3. Bilinear Pairing. Let $\mathbb{G}_1$, $\mathbb{G}_2$, and $\mathbb{G}_T$ be groups of prime order q. A bilinear map $\mathbf{e} : \mathbb{G}_1 \times \mathbb{G}_2 \longrightarrow \mathbb{G}_T$ is a map which satisfies bilinearity, i.e., $\mathbf{e}(g_1^x, g_2^y) = \mathbf{e}(g_1, g_2)^{xy}$ for all $x, y \in \mathbb{Z}_q$; nondegeneracy, i.e., for all generators $g_1 \in \mathbb{G}_1$ and $g_2 \in \mathbb{G}_2$, $\mathbf{e}(g_1, g_2)$ generates $\mathbb{G}_T$; and efficiency, i.e., there exists an efficient algorithm $\mathcal{G}(1^k)$ that outputs the bilinear group $(q, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, \mathbf{e}, g_1, g_2)$. The pairing is a bilinear map and it is symmetric if $\mathbb{G}_1 = \mathbb{G}_2$. There are many types of pairings; however only Weil, Tate, Ate, and Eta parings are mainly used in cryptography due to their efficient computation. Our scheme makes use of the Tate pairing [28], since it is the fastest among them. Most of the known parings uses Millers algorithm [29] to do computations on elliptic curves.

3.4. Weak Boneh-Boyen Signature. The weak Boneh-Boyen (wBB) signature scheme [30] can be used to efficiently sign (blocks of) messages. Furthermore, the signature scheme can be easily integrated with the zero-knowledge proofs so that the knowledge of signed messages (and signatures themselves) can be proven anonymously, unlinkably, and untraceably. We recall the signing and verification algorithms below; the efficient proofs of knowledge are described, e.g., in [24].

`Setup`: On input security parameter k, generate a bilinear group $(q, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, \mathbf{e}, g_1, g_2) \longleftarrow \mathcal{G}(1^k)$. Take $sk \overset{\$}{\leftarrow} \mathbb{Z}_q$, compute $pk = g_2^{sk}$, and output sk as private key and $pk = (q, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, g_1, g_2, \mathbf{e}, pk)$ as public key.

`Sign`: On input message $m \in \mathbb{Z}_q$ and secret key sk, output $\sigma = g_1^{1/(sk+m)}$.

`Verify`: On input the signature σ, message m, and public key pk, output 1 iff $\mathbf{e}(\sigma, pk) \cdot \mathbf{e}(\sigma^m, g_2) = \mathbf{e}(g_1, g_2)$ holds.

Showing the constant signature σ multiple times would make the authentication protocol linkable. All user sessions would be linkable to a single profile, which would make the resulting scheme very privacy unfriendly. To avoid linkability of signatures, users can only prove the knowledge of a valid signature by using the proof defined in [24]. In this proof, the user chooses a random value $r \overset{\$}{\leftarrow} \mathbb{Z}_q$ and computes randomized auxiliary values $\sigma' = \sigma^r$ and $\overline{\sigma} = \sigma'^{-m} g_1^r$. Then, the knowledge of a signature is proven by constructing the zero-knowledge proof $\pi = PK\{(m, r) : \overline{\sigma} = \sigma'^{-m} g_1^r\}$ and verifying $\mathbf{e}(\overline{\sigma}, g_2) = \mathbf{e}(\sigma', pk)$. The verifier is convinced that the user indeed knows a valid signature on a known message, although the proof does not release any of these values. That construction is perfect for our scheme, because the users want to convince verifiers that they know (device) identifiers signed by registrars, in an anonymous, untraceable, and unlinkable manner.

The wBB signatures were proven existentially unforgeable against a weak (nonadaptive) chosen message attack under the q-SDH assumption [30].

4. Multidevice Authentication with Privacy Protection

First, we define the formal requirements on the authentication scheme. Next, we define the algorithms and entities in the scheme. Finally, we present the concrete instantiation of the privacy-enhanced multidevice authentication scheme based on the wBB signatures described in the previous section.

4.1. Requirements. We require the scheme to be secure, private, and efficient.

Security Requirements

Completeness: registered users must be accepted by the `Authenticate` protocol.

Soundness: unregistered users must be rejected by the `Authenticate` protocol.

Zero-Knowledge: the `Authenticate` protocol transcript must be simulatable without the knowledge of identifiers, thus provably releasing no sensitive information.

Privacy Requirements

Anonymity: users must be able to prove the knowledge of their identifiers anonymously, without disclosing them.

Untraceability: user authentication sessions must be untraceable by all system entities, including registrars.

Unlinkability: all single user's authentication sessions must be mutually unlinkable.

Efficiency Requirements

Readiness for RFID devices: the scheme must be fast on constrained devices, in particular smart cards. No operations, that are unavailable on RFID devices (such as bilinear pairings), can be used in user's algorithms.

4.2. Definition of Algorithms. We define the algorithms and protocols of our scheme in this section. The communication pattern is depicted in Figure 2 and employs the registrar (i.e., a central server that manages users and their equipment), users (i.e., user devices such as smart cards or smart phones), terminals (i.e., embedded devices with RFID readers typically attached next to doors), and tags (i.e., devices that need to be present during authentication and access control, typically

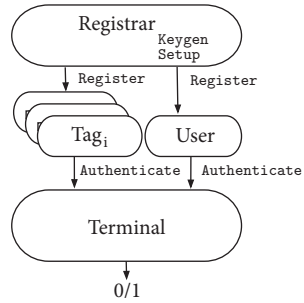


Figure 2: Architecture of multidevice authentication with privacy protection.

safety equipment with programmable RFID sticks, such as the helmet, respirator, or harness).

$(par) \longleftarrow \texttt{Setup}(1^k, n)$: the algorithm is run by the registrar. It inputs the security parameter k and the maximum number of tag classes (i.e., helmets, harnesses, boots, etc.). The algorithm outputs the public system parameters par.

$(sk_r, pk_r) \longleftarrow \texttt{Keygen}(par)$: the algorithm is run by the registrar. On the input of public system parameters par, it generates its private key sk_r and public key pk_r. The registrar distributes the public key to all other entities.

$(\langle ID_i, \sigma_i \rangle_{i=1}^n, ID_u, \sigma_u) \longleftarrow \texttt{Register}(par, sk_r, pk_r)$: the algorithm is run by the registrar. On the input of system parameters and its keypair, the registrar generates the tags' identifiers ID_i with corresponding signatures σ_i and user's identifier ID_u with a corresponding signature σ_u. The tag identifiers and signatures are securely delivered to tags and the user identifier and signature are delivered securely to the user device.

$(0/1) \longleftarrow \texttt{Authenticate}(par, \langle ID_i, \sigma_i \rangle_{i=1}^n, ID_u, \sigma_u, pk_r)$: the cryptographic protocol is run jointly by the user device, tags, and the terminal. It inputs system parameters, registrar's public key, private identifiers, and corresponding signatures and returns 1 iff signatures and IDs are valid, or 0 otherwise.

4.3. Instantiation Using wBB Signatures. In this section, we present the concrete instantiations of cryptographic algorithms defined in Section 4.2. We use the wBB signature scheme to certify the identifiers of tags and users in the `Register` algorithm and interactive proofs of knowledge to prove the knowledge of respective signatures and identifiers in the `Authenticate` protocol. We use the Camenisch-Stadler notation [25] to describe the proof of knowledge protocols.

`Setup`. The algorithm inputs the security parameter k and the maximum number of tag classes n. It generates the bilinear group with parameters $par = (q, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, \mathbf{e}, g_1, \ldots, g_n, g_u \in \mathbb{G}_1, g_2 \in \mathbb{G}_2)$ satisfying $|q| = k$.

`Keygen`. The algorithm inputs the public parameters par, selects random registrar's private keys $sk_r = (sk_0, sk_1, \ldots, sk_n, sk_u) \overset{\$}{\leftarrow} \mathbb{Z}_q^*$, and computes the public keys $pk_r = (pk_0 = g_2^{sk_0}, pk_1 = g_2^{sk_1}, \ldots, pk_n = g_2^{sk_n}, pk_u = g_2^{sk_u})$. It outputs the private keys as registrar's private output and the public key as the public output.

`Register`. The algorithm inputs the registrar's keys and public parameters, randomly selects tag and user identifiers $(ID_1, \ldots, ID_n, ID_u) \overset{\$}{\leftarrow} \mathbb{Z}_q$, and computes the wBB signatures $(\sigma_1, \ldots, \sigma_n)$ on tag identifiers $(ID_1, \ldots, ID_n)$ and the aggregated user signature σ_u and auxiliary values $\langle \sigma_{u_i}, \sigma_{u_i}^{-ID_i} \rangle_{i=1}^n$, $\sigma_{u_u}, \sigma_{u_u}^{-ID_u}$ that allow the construction of efficient proofs of knowledge in the `Authenticate` protocol. The algorithm outputs the tag identifiers and corresponding signatures as a private output to tags. The user identifier, the aggregated signature, and auxiliary values are outputted to the user as a private output. Both tags and the user receive the initial *seed* required for the synchronization of the zero-knowledge proofs as a private input. The algorithm is depicted in Figure 3.

`Authenticate`. `Authenticate` is an algorithm distributed among the user, terminal, and tags that inputs the identifiers and respective signatures and outputs 1 iff (1) all signatures are valid and created by the registrar and (2) all identifiers of the user are present and signed. Otherwise it outputs 0. The protocol is a distributed proof of knowledge of wBB signatures where the tags prove that they know their identifiers and corresponding signatures (without actually revealing them) and, at the same time, the user proves that he has an aggregated signature on all his tag identifiers, plus his own identifier. As the user does not know the tag identifiers, all tags must be present and participate on the proof construction. As a result, the user is able to anonymously, untraceably, and unlinkably prove his valid registration by the registrar and the presence of all his tags, i.e., the safety equipment. The protocol is depicted in abstract CS notation in Figure 4. We also provide the full description in Figure 5 in Section 6 focused on implementation.

5. Security Analysis

The registrar issues the wBB signatures to tags and users in the `Register` algorithm. Then, the user and tags prove

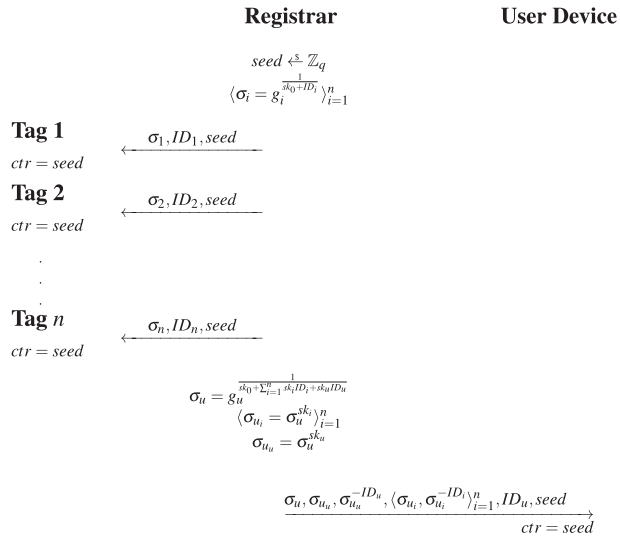


FIGURE 3: `Register` protocol.

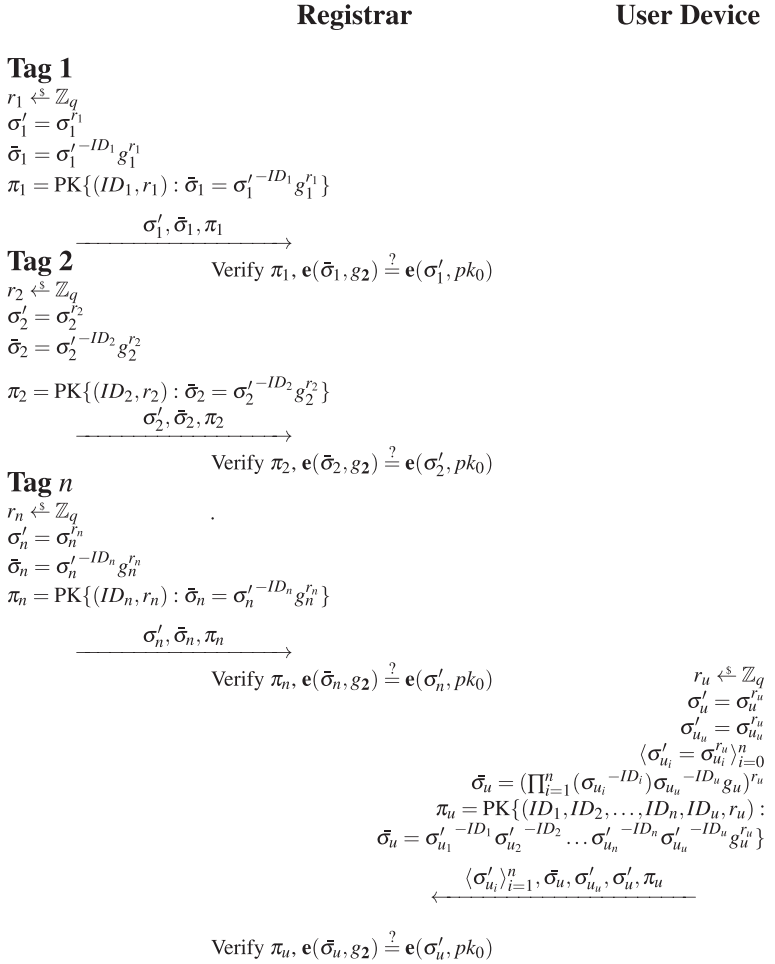


FIGURE 4: `Authenticate` protocol in CS notation.

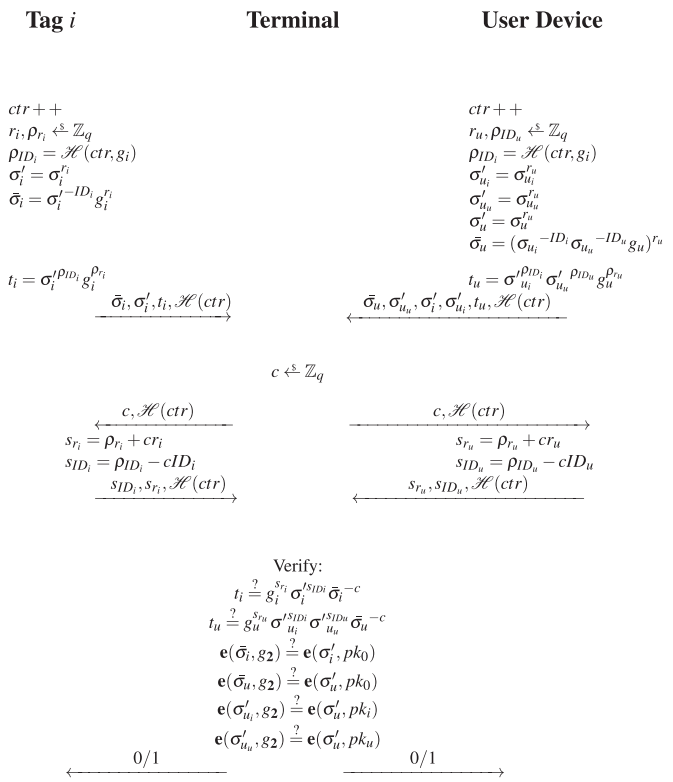


FIGURE 5: `Authenticate` protocol in full notation for i^{th} tag.

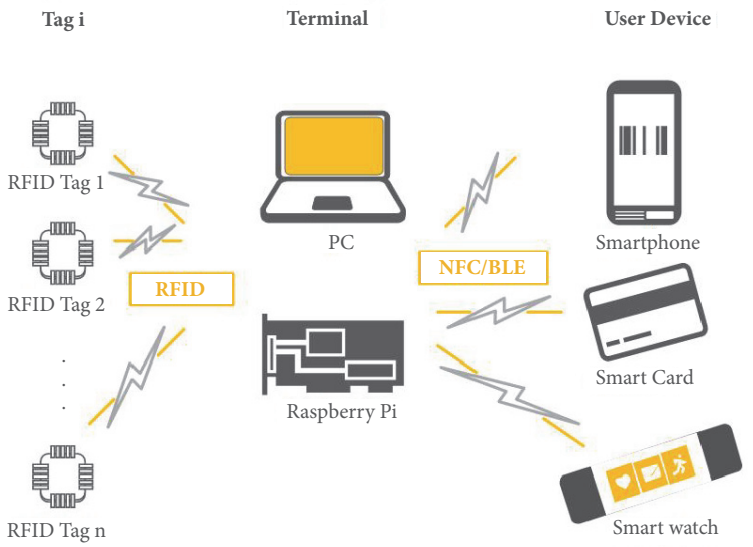


FIGURE 6: Tested scenario.

the knowledge of such signatures to the terminal using the distributed zero-knowledge proofs in the `Authenticate` protocol.

Lemma 1. *The weak Boneh-Boyen signatures are unforgeable against a weak chosen message attack under the q-Static Diffie-Hellman assumption [30].*

Lemma 1 is proven in [30].

Lemma 2. *The protocol presented in Figure 5 is complete, sound, and zero-knowledge.*

We construct the proof for a tag i using the standard proving technique for zero-knowledge protocols. For other

devices and the user, the proof is constructed analogically.

Proof. Completeness: honest users pass the terminal's check.

$$t_i = g_i^{s_{r_i}} \sigma_i'^{s_{IDi}} \overline{\sigma_i}^{-c} \tag{5}$$

$$= g_i^{\rho_{r_i}} g_i^{cr_i} \sigma_i^{r_i s_{ID_i}} \sigma_i'^{ID_i c} g_i^{-cr_i} \tag{6}$$

$$= g_i^{\rho_{r_i}} g_i^{cr_i} \sigma_i^{r_i \rho_{ID_i}} \sigma_i^{-r_i c ID_i} \sigma_i^{r_i ID_i c} g_i^{-cr_i} \tag{7}$$

$$= g_i^{\rho_{r_i}} \sigma_i'^{\rho_{ID_i}} \tag{8}$$

$$\mathbf{e}\left(\overline{\sigma_i}, g_2\right) = \mathbf{e}\left(\sigma_i', pk_0\right) \tag{9}$$

$$\mathbf{e}\left(\sigma_i^{-ID_i r_i} g_i^{r_i}, g_2\right) = \mathbf{e}\left(\sigma_i', g_2^{sk_0}\right) \tag{10}$$

$$\mathbf{e}\left(g_i^{(-ID_i r_i + r_i(sk_o + ID_i))/(sk_0 + ID_i)}, g_2\right) = \mathbf{e}\left(\sigma_i', g_2^{sk_0}\right) \tag{11}$$

$$\mathbf{e}\left(g_i^{r_i sk_o/(sk_0 + ID_i)}, g_2\right) = \mathbf{e}\left(\sigma_i', g_2^{sk_0}\right) \tag{12}$$

$$\mathbf{e}\left(\sigma_i'^{sk_0}, g_2\right) = \mathbf{e}\left(\sigma_i', g_2^{sk_0}\right) \tag{13}$$

□

Error Probability. If implemented correctly, the user will be always accepted.

Proof. Soundness: only registered users pass terminal's check. □

Assume a user who is not registered (i.e., does not know the identifier ID_i) and passes the terminal's check for two different challenges c and c' with two different responses s and s':

$$t_i = g_i^{s_{r_i}} \sigma_i'^{s_{IDi}} \overline{\sigma_i}^{-c} \tag{14}$$

$$t_i = g_i^{s'_{r_i}} \sigma_i'^{s'_{IDi}} \overline{\sigma_i}^{-c'} \tag{15}$$

and we get

$$\overline{\sigma_i}^{c-c'} = g_i^{s_{r_i} - s'_{r_i}} \sigma_i'^{s_{IDi} - s'_{IDi}} \tag{16}$$

and therefore

$$\overline{\sigma_i} = g_i^{(s_{r_i} - s'_{r_i})/(c-c')} \sigma_i'^{(s_{IDi} - s'_{IDi})/(c-c')} \tag{17}$$

Thus the user can efficiently compute both the randomizer $r_i = (s_{r_i} - s'_{r_i})/(c - c')$ and the identifier $ID_i = (s_{IDi} - s'_{IDi})/(c - c')$ and we reached the contradiction to our original assumption.

Error Probability. The attacker will pass the verification check if he can predict the challenge c. The probability of soundness error is thus $P = 2^{-|c|} = 2^{-q} = 2^{-224}$, which is negligible. With an expected rate of 100 ms per challenge, the expected time of breach is 4×10^{58} years.

Proof. Zero-Knowledge: the protocol releases no private information, i.e., there exists a zero-knowledge simulator M_V^*. Using the public parameters and the public key $(\overline{g}, \overline{g}^x)$ (we follow the proof presented in [24] that allows the simulator to input an auxiliary public key $(\overline{g}, \overline{g}^x)$: $\overline{g} \overset{\$}{\leftarrow} \mathbb{Z}_q$ from the registrar), the simulator chooses randomly and uniformly $(s_{r_i}, s_{ID_i}, r, c) \overset{\$}{\leftarrow} \mathbb{Z}_q$, computes $\sigma_i' = \overline{g}^r$, $\overline{\sigma_i} = (\overline{g}^x)^r$, $t_i = g_i^{s_{r_i}} \sigma_i'^{s_{IDi}} \overline{\sigma_i}^{-c}$, and outputs the proof $\pi = (\overline{\sigma_i}, \sigma_i', t_i, \mathcal{H}(r), c, (s_{r_i}, s_{ID_i}))$. The simulated transcript is computationally indistinguishable from the real run of the protocol. □

Error Probability. The attacker can try to guess the randomizers $r_i, \rho_{r_i}, r_u, \rho_{r_u}$ and break the discrete logarithm assumption. The probability is $P = 2^{-q} = 2^{-224}$ for each device, which is negligible. With an expected rate of 10 ms per computing the guess (the exponentiation), the expected time of breach is 4×10^{57} years.

As a result of the zero-knowledge property and randomization of all signatures, the protocol is also *anonymous*, *untraceable,* and *unlinkable*.

6. Implementation Aspects

The `Authenticate` protocol has been implemented as a standard 3-way interactive zero-knowledge proof of knowledge protocol described in Section 3.2. We use a parallel composition with one challenge and one response for all tags of a user to construct an AND proof for both tag and user signatures. The `Authenticate` protocol for i^{th} tag is fully specified in Figure 5.

To keep user devices synchronized, we use a counter that is initialized by a seed generated by the registrar. In the beginning of each session, the counter increments. To avoid losing synchronization, the hashed counter is broadcasted by the terminal so that the devices can compare it with their actual counter value (and with, e.g., 10 next precomputed values) and sync in case their counter is behind. The hashed counter also serves as the session identifier and thus is present in all three steps of the protocol.

In the first step of the protocol, the tag generates randomizers $r_i, \rho_{r_i}, \rho_{ID_i}$, computes randomized signatures $\sigma_i', \overline{\sigma_i}$, and computes the commitment to randomizers t_i. The randomized signatures, commitment to randomizers, and hashed randomizers are sent to the terminal.

In the second step, the terminal randomly selects its challenge c and sends it to all tags and devices, together with the obtained hash.

In the third step, the tag computes their answers s_{r_i}, s_{ID_i} of the zero-knowledge protocol.

After receiving the answers, the terminal is able to verify that the tag knows a valid signature and a corresponding tag identifier with respect to registrar's public key pk, without actually learning any user- or tag-identifying values.

The proof construction for the user is the same with the exception that the answers containing tag IDs are omitted, because the terminal makes use of the values received by

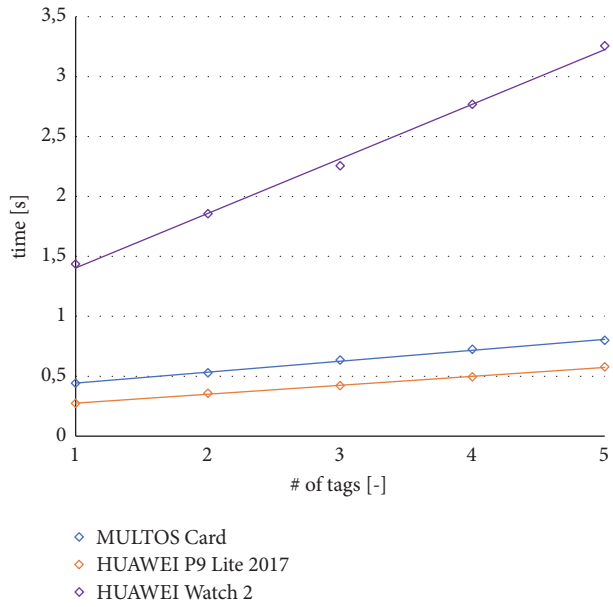


Figure 7: Dependence of the proof construction time on the number of user device.

the devices. Instead of proving tag IDs, the user proves the knowledge of his own user ID.

6.1. Performance Analysis. The scheme was designed to be practical and fast on constrained RFID devices, such as smart cards and programmable RFID tags. Therefore, the bilinear pairings, which are the most computationally complex operations in our algorithm, are only computed in the terminal which normally has more resources than user device. The second most complex operation is the exponentiation (implemented as scalar multiplication of an elliptic-curve point) and it is reduced to a minimum. The user device needs $(5 + 2d)$ exponentiations to construct a "user proof" with d personal tags. Each tag must compute 5 exponentiations to generate a "tag proof". However, our implementation uses only 4 exponentiations, since the value $(\sigma_i^{-ID_i} g_i)$ is precomputed within a `Register` protocol and is used for the randomized signature $\overline{\sigma_i} = \sigma_i'^{-ID_i} g_i'^{r_i} = (\sigma_i^{-ID_i} g_i)^{r_i}$ construction. The complexity of the other operations (random number generation, addition, and multiplication) are only minor, compared to pairings and exponentiations. In order to verify the proof, the terminal must compute $(4 + 4d)$ bilinear pairings and $(4 + 3d)$ exponentiations.

We provide performance measurement of crucial operations on common devices, which are widely used in the access control applications, i.e., a smart card, smart phone, smart watch (as user devices), and a custom-built RFID terminal with ARM or Intel CPU and programmable RFID tags (as RFID tags attached to safety equipment). The hardware and software specification of all the devices is presented in Table 1.

The testing scenario is depicted in Figure 6. The user needs to hold a wearable device, such as a smart phone (HUAWEI P9 Lite 2017), a smart card (Multos Card), or smart watch (HUAWEI Watch 2) and some safety equipment, such as helmets, harnesses, boots, and protective suits, each of them with a programmable RFID tag attached. The tag is equipped with a programmable chip SC23Z018 with Multos 4.2 operation system. The proofs are collected and verified by a terminal. We use Raspberry Pi 3 to represent the terminal. In another scenario, PC (Intel i7-7700 CPU, 16 GB RAM) acts as a central authentication server representing the case of a centralized access control system. The system uses RFID communication between tags and a terminal, and NFC or BLE communication between a terminal and a user device.

The performance of critical operations and the estimation of the running time of the `Authenticate` protocol with one RFID tag and one user device are presented in Table 2. In addition, we provide measurement of the selected devices where we consider different elliptic curves types, in particular type A and D. Both curves satisfy the NIST key recommendation for 80-bit security strength [31]. The performance is measured in milliseconds (the measurement of clock cycles is unavailable on the smart card platform) and the values are an average of 10 measurements, excluding communication overhead. For the implementation of EC operations, the PBC library was used [32] on the terminal and jPBC [33] library on Android devices. Native assembler code was used to perform operations on the Multos smart card.

The proposed authentication scheme can be used in many types of access control scenarios and for different types of devices. Therefore, we provide the results of each protocol using one RFID tag. Furthermore, we present the crucial EC operations' benchmarks on a wide range of devices in Table 3. The time is measured in milliseconds and the values are an average of 10 measurements, as in the previous case. All measurements were performed by using the elliptic curve d159 from the PBC library. We did not consider Android devices as a terminal device, since the pairing operation requires too much time and therefore it is not usable in practice.

Figure 7 depicts the time required for a proof construction on different devices (Multos smart card, Android smart

Table 1: Specification of tested devices.

	Type	CPU/MCU	OS	RAM
Tag	SC	SC23Z018	Multos 4.2	2 KB
User	SC	SC23Z018	Multos 4.2	2 KB
User	Phone	Kirin 655	Android 7.0	3 GB
User	Watch	ARM Cortex-A7	Android 7.0	768 MB
Terminal	Pi 3	ARM Cortex-A53	Raspbian 9.3	1 GB
Terminal	PC	Intel i7-7700	Debian 8.6	16 GB

Tag: programmable RFID stick, User: user device, Terminal: terminal, SC: smart card, Phone: HUAWEI P9 Lite 2017, Pi 3: Raspberry Pi 3 Model B, and Watch: HUAWEI Watch 2.

Table 2: Benchmark results based on EC type.

	Terminal [ms]	User Device [ms]	Tag [ms]
Elliptic Curve Type A			
Exponentiation	10	67	81
Pairing	15	125	-
Verification	192	-	-
Tag Proof Generation	-	-	444
User Proof Generation	-	448	-
Elliptic Curve Type D			
Exponentiation	4	38	40
Pairing	31	1050	-
Verification	271	-	-
Tag Proof Generation	-	-	277
User Proof Generation	-	273	-

Table 3: Benchmark results of all tested devices.

	SC [ms]	Phone [ms]	Watch [ms]	Pi 3 [ms]	PC [ms]
Exponentiation	40	38	207	3.3	0.4
Pairing	-	1050	6571	31	2.4
Tag Proof Generation	277	154	900	18	4
User Proof Generation	441	273	1502	24	5
Verification	-	-	-	271	21

phone, and smart watch for various number of tags). These devices act as a user device.

6.2. Revocation and Identification. Besides strong privacy-enhancing features, there must be also mechanisms to revoke and/or identify malicious users. All users are theoretically identifiable and traceable by their user IDs. However, these IDs are "hidden" in the signatures as the exponents. Due to the discrete logarithm problem assumption, one cannot easily get the identifiers and do the revocation and identification. However, our scheme is compatible with the major revocation schemes that are already available for cryptographic anonymous credential schemes [22–24]. In these revocation schemes, the hidden exponent (the user ID) is used as a revocation handle and can be disclosed only by designated authorities. Additionally, valid users remain anonymous while malicious users are identifiable and traceable by a designated authority, such as a court. Such schemes are provably secure, efficient, and compatible without any modification; thus we refer to their specification (e.g., the scheme designed directly for smart cards [24]) in case revocation is needed.

7. Conclusions

We presented a cryptographic scheme that allows a novel approach for controlling physical access. Instead of the verification of fixed user or device identifiers, the terminals can check only the knowledge of such identifiers in a private manner, without explicitly exposing any personal information or the identifiers themselves. Furthermore, the presence of other RFID devices, possibly the safety equipment, can be enforced. Our protocols are based on proven cryptographic

algorithms and are very practical—the proofs can be generated in under 500 ms on constrained devices, such as smart cards. We provided the full cryptographic description of all algorithms, the security and efficiency analysis, and the implementation results on constrained devices. We find the scheme especially useful in applications where the physical access to dangerous environment is granted upon proving the presence of required safety equipment and where the strong privacy-protection regulation is enforced by law.

As for the future work, we will focus on the optimization of the verification algorithm, since the current verification time grows linearly with the number of tags involved in the authentication protocol. In particular, we would like to reduce the number of bilinear pairings which is the most time-consuming operation in the protocol.

Data Availability

All necessary information is provided in the paper and in cited literature.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Research described in this paper was financed by the National Sustainability Program under Grant LO1401. For the research, infrastructure of the SIX Center was used.

References

[1] W. Xu, "Mobile applications based on smart wearable devices," in *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems, SenSys 2015*, pp. 505-506, Republic of Korea, November 2015.

[2] NIST. Federal information processing standards publication (FIPS 197). Advanced Encryption Standard (AES), 2001.

[3] N. Courtois, K. Nohl, and S. ONeil, *Algebraic attacks on the crypto-1 stream cipher in mifare classic and oyster cards*, IACR Cryptology ePrint Archive, 2008.

[4] N. T. Courtois, "The dark side of security by obscurity: And cloning MiFare classic rail and building passes, anywhere, anytime," in *Proceedings of the SECRYPT 2009 - International Conference on Security and Cryptography*, pp. 331–338, Italy, October 2009.

[5] F. D. Garcia, P. Van Rossum, R. Verdult, and R. W. Schreur, "Wirelessly pickpocketing a Mifare Classic card," in *Proceedings of the 2009 30th IEEE Symposium on Security and Privacy*, pp. 3–15, USA, May 2009.

[6] C. Paquin, "U-prove cryptographic specification v1.1," Technical report, 2011.

[7] O. Riva, C. Qin, K. Strauss, and D. Lymberopoulos, "Progressive authentication: Deciding when to authenticate on mobile phones," in *In USENIX Security Symposium*, pp. 301–316, 2012.

[8] V. S. Miller, "The Weil pairing, and its efficient calculation," *Journal of Cryptology. The Journal of the International Association for Cryptologic Research*, vol. 17, no. 4, pp. 235–261, 2004.

[9] W. Mostowski and P. Vullers, "Efficient U-prove implementation for anonymous credentials on smart cards," *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, vol. 96, pp. 243–260, 2012.

[10] P. Vullers and G. Alpár, "Efficient Selective Disclosure on Smart Cards Using Idemix," in *Policies and Research in Identity Management*, vol. 396 of *IFIP Advances in Information and Communication Technology*, pp. 53–67, Springer, Berlin, Heidelberg, 2013.

[11] D. Oswald and C. Paar, "Breaking Mifare DESFire MF3ICD40: Power analysis and templates in the real world," in *Cryptographic Hardware and Embedded Systems–CHES 2011*, vol. 6917, pp. 207–222, 2011.

[12] J. Camenisch and E. V. Herreweghen, "Design and implementation of the idemix anonymous credential system," in *Proceedings of the 9th ACM Conference on Computer and Communications Security CCS '02*, pp. 21–30, ACM, New York, NY, USA, November 2002.

[13] S. Micali and C.-P. Schnorr, "Efficient, perfect polynomial random number generators," *Journal of Cryptology. The Journal of the International Association for Cryptologic Research*, vol. 3, no. 3, pp. 157–172, 1991.

[14] J. Hajny and L. Malina, "Unlinkable attribute-based credentials with practical revocation on smart-cards," in *Smart Card Research and Advanced Applications - CARDIS*, S. Mangard, Ed., vol. 7771 of *Lecture Notes in Computer Science*, pp. 62–76, Springer, Berlin, Heidelberg, 2013.

[15] NIST. Federal information processing standards publication (FIPS 46-3). Data Encryption Standard (DES), 1999.

[16] J. Hajny, L. Malina, and O. Tethal, "Privacy-Friendly Access Control Based on Personal Attributes," in *Proceedings of the The 9th International Workshop on Security*, vol. 8639 of *Lecture Notes in Computer Science*, pp. 1–16, Springer International Publishing.

[17] https://srlabs.de/analyzing-legic-prime-rfids/.

[18] B.-R. Cha, S.-H. Lee, S.-B. Park, and Y.-K. Ji, "Design of micropayment to strengthen security by 2 factor authentication with mobile wearable devices," *Advanced Science and Technology Letters (ASTL)*, vol. 109, no. 7, pp. 28–32, 2015.

[19] I. Butun, M. Erol-Kantarci, B. Kantarci, and H. Song, "Cloud-centric multi-level authentication as a service for secure public safety device networks," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 47–53, 2016.

[20] L. Gonzalez-Manzano, J. d. Fuentes, and A. Orfila, "Access Control for the Cloud Based on Multi-device Authentication," in *Proceedings of the 2015 IEEE Trustcom/BigDataSE/ISPA*, pp. 856–863, Helsinki, Finland, August 2015.

[21] J. Hajny, P. Dzurenda, and L. Malina, "Multi-device authentication using wearables and iot," in *In Proceedings of the 13th International Joint Conference on e-Business and Telecommunications, ICETE 2016*, pp. 483–488, SCITEPRESS - Science and Technology Publications, Lda, Portugal, 2016.

[22] E. Brickell, J. Camenisch, and L. Chen, "Direct anonymous attestation," in *Proceedings of the 11th ACM Conference on Computer and Communications Security (CCS '04)*, pp. 132–145, ACM, New York, NY, USA, 2004.

[23] P. Bichsel, J. Camenisch, T. Groß, and V. Shoup, "Anonymous credentials on a standard Java card," in *Proceedings of the 16th ACM Conference on Computer and Communications Security, CCS'09*, pp. 600–610, New York, NY, USA, November 2009.

[24] J. Camenisch, M. Drijvers, and J. Hajny, "Scalable revocation scheme for anonymous credentials based on n-times unlinkable

proofs," in *Proceedings of the 15th ACM Workshop on Privacy in the Electronic Society, WPES 2016*, pp. 123–133, New York, NY, USA, 2016.

[25] J. Camenisch and M. Stadler, "Efficient group signature schemes for large groups," in *Advances in Cryptology — CRYPTO '97*, vol. 1294 of *Lecture Notes in Computer Science*, pp. 410–424, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.

[26] R. Cramer, *Modular Design of Secure Yet Practical Cryptographic Protocols*, Universiteit van Amsterdam, 1997.

[27] The White House. National strategy for trusted identities in cyberspace, 2011. http://www.whitehouse.gov/sites/default/files/rss_viewer/NSTICstrategy_041511.pdf .

[28] S. D. Galbraith, K. Harrison, and D. Soldera, "Implementing the tate pairing," in *Algorithmic Number Theory*, C. Fieker and D. R. Kohel, Eds., pp. 324–337, Springer, Berlin, Heidelberg, 2002.

[29] I. Naumann and G. Hogben, "Enisa: Privacy features of eid cards," *Network Security Newslette*, vol. 2008, no. 8, pp. 9–13, 2008.

[30] D. Boneh and X. Boyen, "Short signatures without random oracles," in *Advances in Cryptology—EUROCRYPT 2004*, vol. 3027 of *Lecture Notes in Computer Science*, pp. 56–73, Springer, Berlin, Germany, 2004.

[31] E. Barker, "Recommendation for key management part 1: General (revision 4)," in *NIST Special Publication Part 1*, vol. 800 (57), pp. 1–147, 2016.

[32] M. Meriac, "Heart of darkness-exploring the uncharted backwaters of hid iclasstm security," *Heart*, 2010.

[33] A. De Caro and V. Iovino, "jPBC: Java pairing based cryptography," in *Proceedings of the 16th IEEE Symposium on Computers and Communications (ISCC '11)*, IEEE, pp. 850–855, Kerkyra, Corfu, Greece, July 2011.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 9359878, 10 pages
https://doi.org/10.1155/2018/9359878

Research Article

An SDN-Based Connectivity Control System for Wi-Fi Devices

Duc-Thang Nguyen and Taehong Kim

School of Information and Communication Engineering, Chungbuk National University, Cheongju, Republic of Korea

Correspondence should be addressed to Taehong Kim; taehongkim@cbnu.ac.kr

Received 27 March 2018; Revised 20 June 2018; Accepted 11 July 2018; Published 24 July 2018

Academic Editor: Kim-Kwang Raymond Choo

In recent years, the prevalence of Wi-Fi-enabled devices such as smartphones, smart appliances, and various sensors has increased. As most IoT devices lack a display or a keypad owing to their tiny size, it is difficult to set connectivity information such as service set identifier (SSID) and password without any help from external devices such as smartphones. Moreover, it is much more complex to apply advanced connectivity options such as SSID hiding, MAC ID filtering, and Wi-Fi Protected Access (WPA) to these devices. Thus, we need a new Wi-Fi network management system which not only facilitates client access operations but also provides a high-level authentication procedure. In this paper, we introduce a remote connectivity control system for Wi-Fi devices based on software-defined networking (SDN) in a wireless environment. The main contributions of the proposed system are twofold: (i) it enables network owner/administrator to manage and approve connection request from Wi-Fi devices through remote services, which is essential for easy connection management across diverse IoT devices; (ii) it also allows fine-grained access control at the device level through remote control. We describe the architecture of SDN-based remote connectivity control of Wi-Fi devices. While verifying the feasibility and performance of the proposed system, we discuss how the proposed system can benefit both service providers and users.

1. Introduction

The popularity of wireless devices is increasing rapidly, and an increasing number of end-devices connect to the Internet through Wi-Fi—anything from home appliances to factory facilities. It is not hard to find a Wi-Fi signal in places such as offices, cafes, or even across entire cities. In addition, most essential appliances such as smartphones, laptops, speakers, and wearable gadgets support the Wi-Fi protocol.

Basically, most home or office devices need a service set identifier (SSID), which is broadcasted by typical Wi-Fi routers. In order to enhance wireless network security, a number of authentication procedures such as SSID hiding, media access control identifier (MAC ID) filtering, or Wi-Fi Protected Access 2 (WPA2) are applied. For instance, in the case of the SSID hiding method, only users who know the SSID can access the Wi-Fi network. In the most popular authentication Wi-Fi mechanism—WPA2—the router works in a secure mode with an SSID and password pair. Anyone who provides the correct pair is granted network access. Besides, with MAC ID filtering, only devices with allowed MAC addresses can join the network or use the Internet. Moreover, with the appearance of advanced security mechanisms such as Extensible Authentication Protocol (EAP) and Lightweight Extensible Authentication Protocol (LEAP), users' data and information can be protected more efficiently while surfing the Internet.

Nevertheless, as a side effect of the many security advantages, the complex authentication procedure could affect user experience when connecting to the network. For instance, when a stranger visits the user's home and wants to use their Wi-Fi network, they must be told the SSID and the password. In the case of MAC-based authentication, an entry for the guest device's MAC address should be added to the authentication service in advance by the network administrator for setting up a connection. This means that the more complex the authentication setting, the more difficult it becomes to set up the authentication procedure for adding or replacing new devices. In particular, nontechnical elderly people who are not familiar with IT devices may find it tedious to use complicated authentication procedures.

Besides, with IoT trending, the number of smart appliances using home network protocols such as ZigBee [1], Z-WAVE [2], and Bluetooth [3] has been increasing explosively. Even though these kinds of protocols are suitable for lightweight or low-energy devices, IoT devices are still equipped with Wi-Fi functionality for easy management. Moreover, the "things" now being added to homes, offices, and industrial facilities are quite small and do not have a display or convenient data entry capability to allow users to configure them. As a result, the devices need to be configured by a smarter device that can transmit the necessary configuration information to join the network. For example, in order to make a smart switch join a home network, customers have to install a separate application on their smartphone or tablet and then the app can transmit Wi-Fi access information to the IoT devices via near-field communication (NFC) or Bluetooth. Then, the gadget can connect to the current network and perform related setup to use its features. As with the previous examples, instructions for such a configuration can sometimes be difficult for a novice user unfamiliar with IT.

Although a remote connectivity control scheme for Wi-Fi-enabled equipment is essential for both customers and network owner/administrator, it is hard to implement one through just minor enhancements to traditional network systems. Note that software-defined networking (SDN) provides an architecture enabling programmatic enhancements to network management and configuration [4–9], which form the base architecture for us to achieve our goals.

Thus, in this paper, we propose a new SDN-based mechanism that not only enhances the connection-establishment process for new wireless devices but also improves network management functionalities. More specifically, our system comes with interesting use cases. (1) It enables automatic network configuration and allows administrators or users to set up new Wi-Fi device joining their network process remotely. These events are announced to owner/administrator through remote notification services and displayed on a web-based user interface. (2) Using its associated station's information and per-client virtual access points, the connectivity control system supports flexible network functions (e.g., network access rules for individual devices). To demonstrate the feasibility and benefits of our approach, we evaluate it on a real prototype. We believe that our system can bring innovation to network features and additionally facilitate user-network interactions.

This paper is organized as follows. Section 2 reviews some related studies and Section 3 provides preliminaries for the proposed system. With the use cases of our system in Section 4, Section 5 describes the system architecture in detail. Section 6 provides performance evaluations from experimental testbed. And Section 7 discusses the security enhancements for the connectivity control system. Finally, we conclude this paper in Section 8.

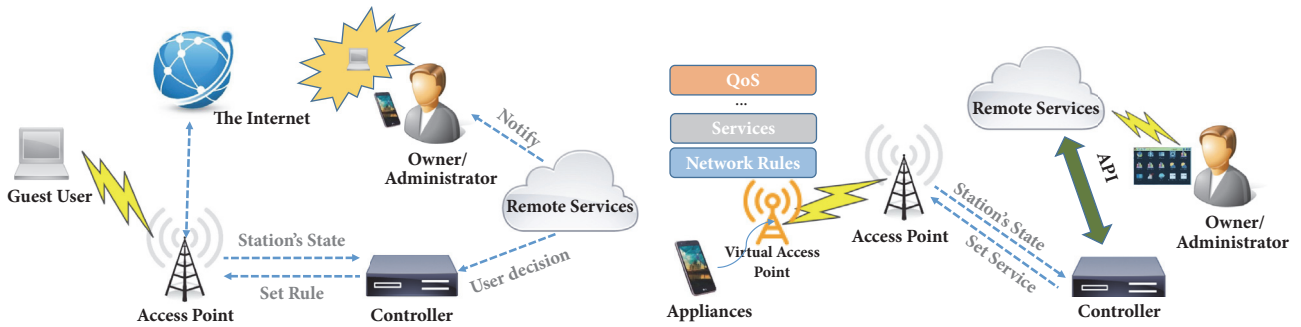
2. Related Works

Connectivity control of smart devices, especially wireless-enabled home or office devices, is one of the important issues in the IoT paradigm. Numerous manufacturers such as Samsung [10], Apple [11], and Philips [12] are developing their own smart device ecosystems, which users who buy their equipment can benefit from. For example, customers can configure and manage their devices through cloud services supported by the manufacturer. However, these kinds of services are specific to each manufacturer. This indicates that if customers own gadgets of different brands, they have to use different services for each device. Thus, this operation may require complex steps, inconveniencing customers.

In order to solve these connectivity control issues, M. Lee et al. in [13] proposed a new autoconfiguration method of home network with SDN controller support. Using Floodlight—SDN controller acting as a cloud-based home network controller—the system enables automatic recognition and management of home devices without requiring specific equipment. There are diverse home devices such as notebooks, smartphones, and light sensors and their MAC addresses as an identifier are stored in a database. Besides, SDN enables allocating bandwidth for quality of service (QoS) for each device. However, this system has not considered wireless environment. Thus, it lacks a number of functionalities for WLANs such as access point, association, and roaming. On the contrary, our paper targets WLANs and improves traditional access points with LVAP abstraction, which we will explain comprehensively in the next section, for connectivity control of home devices.

Yet another approach to IoT device management in smart homes was proposed by Vijay in [14]. In this study, the authors built and demonstrated the advantages of using SDN in IoT security and network services from an Internet service provider. Specifically, the implemented design showed the user interface, which contains the subscribers' household devices and allows update or deletion of the flow rules of access control. This also means that customers have the ability to remotely block/quarantine devices based on network activity via a web-based portal. Nevertheless, the author did not clarify the process of establishing a connection for each device and focused on security after the successful setup of all equipment. On the other hand, this paper specifies the detail procedure for connection setup as well as proving feasible through a testbed. In addition, the network owner/administrator has the right to allow/reject network access to clients without using MAC filtering or blocking device IPs.

The study in [15] suggests a simple mechanism that allows small network owner/administrator to provide Internet access to guests using OpenID or certificates issued by third-party services. It has been suggested that this system seems to be a useful approach since it reduces authentication operations. Similar to this work, authors in [16, 17] introduced a solution which enhances the wireless access services for enterprise WLAN by combining SDN and an authentication server. However, these mechanisms require display-enabled devices, which small IoT devices such as switches, plugs, and sensors are not. Hence, we extend this proposal in order to apply to any type of IoT devices regardless of their size, providing easy device configuration and management through the remote system. More technically, as mentioned before, our connectivity control system handles the client's primitive state and moves the network decision module to the



(a) Remote connectivity control of a guest's Wi-Fi device: owner allows/rejects a guest's device through rule set up to the Wi-Fi network

(b) Remote connectivity control and management of IoT devices: owner controls IoT devices through remote services provided by the controller

Figure 1: SDN-based connectivity system use cases.

remote controller. Thus, customers can install and manage their network automatically without complex steps.

3. Preliminaries

So far, a concept of SDN for wireless network was introduced in [18–20] and built upon Odin's architecture [21]. These systems also inherit Light Virtual Access Point (LVAP) abstraction, master-agent paired modules, and the traditional OpenFlow controller [22]. While the OpenFlow controller determines the best path for network traffic, the master module makes decisions about allocation of network resources. In addition, the client state is recorded and forwarded by the agent module, which is integrated into the wireless access points.

LVAPs are key components of the Odin framework. Basically, each client will be attached to the image of the access point they connect to. Specifically, in physical wireless access points, whenever the Wi-Fi card receives a valid 802.11 frame, it immediately spawns an LVAP associated with the client. Hence, the LVAP becomes a potential per-client virtual access point for the client to perform association. Besides, since this data is transmitted to the controller, the LVAP manager module built into the controller will hold the unique virtual access point information such as BSSID, MAC address, IP address, and SSID as a representation of the client. For instance, there are two particular access points which are both managed by the same controller. A client is previously associated with the system and attached an LVAP. When the client moves out of the coverage area of the first access point while entering the second areas, the handoff occurs without requiring a reassociation and exchanging additional layer 2 or 3 messages. This mechanism effectively enhances the handoff performance and distributed-client problems [23–25].

However, the above-mentioned architecture is implemented as an "open" system; any devices can join and access the Internet without permission. Thus, this system lacks a controllable connectivity module that deals with device authentication and management. In order to enable the connectivity control system, we propose an approach that involves using the SDN controller as a "decision center" that collaborates with other business services. Through this system, the controller holds all the information about associated clients and shares these resources with other network services. Thus, it can help the owners/administrators to manage and take control of their own network system.

Furthermore, LVAP can take advantage of the ability to support multiple services through each virtual access point. For example, the network administrator can control each client individually with respect to connection time and bandwidth limitation without affecting others. In our system, LVAP abstractions are used to mark the appearance and handle the connectivity process of clients. Particularly, since our device manager module holds the data of the attached LVAP in a physical access point, based on user actions, it can make decisions to allow/reject spawning LVAPs or provide network rules for each client. This also means that an administrator who is physically away from the network has full access to take over device connection establishment. We will describe some of these use cases in the next section.

4. Use Cases

Our solution is based on programmable network devices in the spirit of SDN. In this section, we discuss some use cases for the envisioned system.

4.1. Remote Connectivity Control of a Guest's Wi-Fi Devices. One of the use cases our system supports is sharing Internet connectivity with visitors in home, office, or enterprise. Basically, a visitor should be given the WPA passphrase or the administrator should manually add an entry of his/her MAC address for authentication. The disadvantages of a classic private shared key include the fact that it is impossible to revoke the secret key when he/she leaves as well as the fact that it is relatively easy to crack. The alternative solution, 802.1X, requires the installation of a software client and remote authentication server, thus making it difficult or impossible to use on devices that do not support such protocols. Using our system, we can innovate and simplify authentication. For example, in Figure 1(a), when a visitor wants to use the Internet, he/she will connect to the public SSID broadcasted by the access point. Then, the network owner will be notified of the identity of the newcomer's device

via the remote services working on top of the controller. Finally, he/she will accept or deny the connection and the corresponding rules on the access point will be set by the controller. Thus, the guest does not have to input any password and the network's owner could accept the request remotely. This can, of course, be done at a large-scale offices and enterprises with a supervisor or a network administrator.

4.2. Remote Connectivity Control and Management of IoT Devices. Our connectivity control system also facilitates the connection and management of IoT devices. Note that most IoT devices lack a display or a keypad, requiring the use of other smart devices to establish a connection. As Figure 1(b) illustrates, with our system, the client state will be forwarded immediately to the controller when it interacts with the system. Thus, this information can be used through multiple remote or cloud services to provide network services to customers. For example, a customer buys a smart plug that contains default SSID information from the manufacturer. In addition, an access point in the owner's home also broadcasts the same SSID. This is supported by our concept. As a result, when the plug is powered on, it automatically associates with our system. In addition, owners/administrators will be informed about new devices and can control it through several web or mobile services. Moreover, since each associated client is attached to a virtual access point, it allows fine-grained control of network resources according to user demand and expected QoS level. Note that SSID information can be updated after the IoT device successfully joins the network. Detail procedure will be discussed in the next section.

5. SDN-Based Connectivity Control System

In this section, we describe an SDN-based connectivity control system to achieve the above-mentioned use cases.

5.1. Overview of Connectivity Control System. In this subsection, we describe the procedure of the connectivity control system as shown in Figure 2, together with its software architecture in Figure 3. Figure 2 describes the procedure of granting permission and establishing connection for newcomer. Figure 3 shows that the proposed architecture is composed of diverse software components such as *ClientStateResource* for each access point and *ClientStateHandler*, *DeviceManager*, and *ConnectivityControlApplication* for the controller and remote services. Specifically, when an 802.11 management frame is received by a *ClientStateResource* module which is implemented in the access point, it checks whether the client is assigned an LVAP. Immediately, this process also takes place in the *ClientStateHandler* module in the controller, which assigns an LVAP to each client if they have not been assigned LVAPs or have been previously removed. After that, the client information is checked against the allowed MAC addresses—whether one exists for the client that generated the frame. Registered devices will appear in the *AllowedList* and managed by the *DeviceManager* module, used for device identification and authentication. If so, via client-state information saved in the LVAP, the controller sets up the connection with the station. If there is a newcomer that has not interacted with the system before, the agent informs the controller, and the *ConnectivityControlApplication* on top of the controller framework will start the authentication. Moreover, with the notification system, the network owner/administrator decides to allow/deny the joining requests from devices. Subsequently, the controller attempts to respond to the clients accordingly. The process ends when permission is granted to the new station or the frame is dropped and the joining process fails if the administrator denies the connection. Since all new device-joining events come to the remote owner/administrator, it is possible to achieve the first use case. In other words, it eliminates the complex procedure of modifying configuration such as SSID hiding and MAC filtering at the access point level. It is also possible to make IoT devices without a display joining the access point and then the remote controller just accepts them. Moreover, for the second use case, since each device is attached to an LVAP, the controller has the ability to monitor all Wi-Fi devices status and apply individual device-level control in the form of access control and bandwidth control.

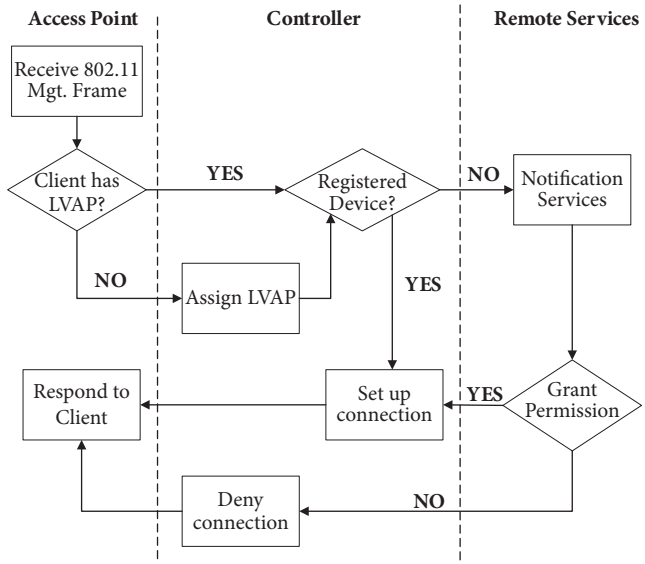


Figure 2: Procedure of SDN-based connectivity control system.

5.2. Software Architecture

5.2.1. Access Point. In our system, an access point contains both OpenFlow-enabled switch instances managing the communication over wired data paths and a *ClientStateResources* holding station states as virtual access points and 802.11 radio signals. Specifically, access points will be equipped with the *ClientStateResources* module, which communicates with the *ClientStateHandler* module in the controller via a persistent TCP connection. Therefore, not only can we perform the normal features of an access point, they are also involved in gathering client states before forwarding it to the controller. In technical detail, when a client is associated with the access point, the access point will hold the SSID, client MAC address, and client IP address information (if granted). Based on this information, a unique

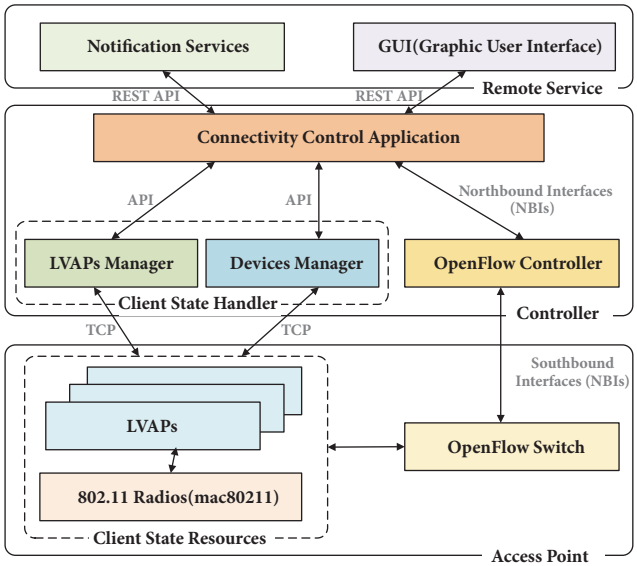


Figure 3: The software architecture of an SDN-based connectivity system.

virtual access point will be attached to the client, which operates a number of services provided by the controller.

In a conventional wireless network, there are numerous access points acting as a central gateway for clients. However, due to restricted programmability levels, current access points are limited in both enhancing network services and simplifying setup operations for users. For example, creating a new SSID or configuring a wireless access point is still complex to nontechnical users. SDN brings an innovation to traditional networks by moving the decision center to the remote controller. Despite the potential of the OpenFlow protocol in SDN controllers, it is still not possible to apply it to a wireless environment effectively. Thus, besides the OpenFlow protocol securing the connection between the switch and controller, a TCP connection is used to transfer the information of client states to the higher layer. Moreover, this connection secures the wireless data path and control path between the controller and the access point.

5.2.2. Controller. The subcomponent of our controller in Figure 3 builds upon the Empower [26] and consists of the following components: (1) *ClientStateHandler* stores and manages LVAP through the LVAP manager and the device manager, (2) *OpenFlowController* is responsible for deciding how a specific packet is handled via a wired control path, and (3) the *ConnectivityControlApplication* on top of the framework. Technically, when a device joins through an access point, its primitive states are immediately transmitted to the *Controller* through the communication protocol between access point and controller. Based on this information, the *LVAPsManager* module can build a logical representation of the client in the controller—an LVAP for handling another process. Furthermore, the *DeviceManager* also recognizes whether devices are registered with the system in advance. If it is registered, it decides which services or rules can be applied to these devices. Finally, *ConnectivityControlApplication*, which is associated with the remote services via an API, will decide whether to keep the connection alive or terminate it based on user demand. After the network application completes its operations, the response state will be sent to the access point and client, and flow rules will be applied to the *OpenFlowController*.

5.2.3. Remote Services. Remote services lay on top of the application layer, which provides a wide range of services to users such as web and cloud services. By passing data through network applications in the controller, a number of associated services can be delivered through the Internet to users. For example, in an SDN-based connectivity control system, a remote server collaborating with the *ConnectivityControlApplication* is responsible for authentication services. Thus, we set up automatic notification services that announce to users which devices are trying to join the network. Additionally, through a RESTful web service, owner/administrator can set up devices connection remotely as well as monitoring the status of active devices.

6. Performance Evaluations

To demonstrate the feasibility and benefits of our approach, we conducted experiments in a test bed. Our controller was evaluated on a nonvirtualized machine with 8 CPU cores supporting hyperthreading and 16 GB of RAM. The host's operating system was Ubuntu 16.04 and it ran modified software named Empower Wi-Fi, introduced by R. Riggio et al. in [19]. We chose this particular apparatus due to the efficiency of developing and maintaining network applications on top of the framework with the support of the Empower software development kit. In addition, a web service contains a graphic user interface (GUI) providing the user with an interactive interface via their personal computer or mobile phone.

All access points run OpenWRT [27] release Chaos Calmer with the ath9k Linux driver, user-level Click modular router [28], and OpenvSwitch (OvS) [29] version 2.39 supporting OpenFlow (OF) [30] version 1.3 and conntrack table management. Besides, each access point also contains the *ClientStateResources* module, collecting client state and communicating with the *ClientStateHandler* module in the higher layer. While OpenvSwitch instances manage the communication via a wired network topology, Click modular router instances implement IEEE 802.11 data paths. Technically, Click is a framework for handling multipurpose packet processing and is used to implement LVAP frame exchange, while all the decision logic is built into the controller. Other gadgets supporting our experiments are Raspberry Pi Model 3B [31] running a FreeRADIUS [32] server and one TP-LINK1043ND router for providing DHCP and Internet access.

In order to support our evaluation, we defined some APIs based on the features of notification systems. Table 1 shows samples of the RESTful API that was used in the network controller. Specifically, when the authentication process takes place in devices control application at the network controller, through this API, the system can capture the decision of

Table 1: REST APIs in the connectivity control module.

Action	API	Description
Allow	/api/device/allow/id	Allow device {id} to join the network
Reject	/api/device/reject/id	Deny device {id} from joining the network
Notify	/api/services/device/id/message	Notify user's device {id} trying to join the network with {message}

Figure 4: Web interface showing information of devices such as stations, SSID, LVAP, access point, station names, and status.

owner/administrator to allow/reject a connection as returned by the notification service. In addition, to control a smart device using remote services, developers can create any kind of web, mobile, or PC application and all of them can easily work with our system.

6.1. Experiments on the SDN-Based Connectivity Control System. Experiments were conducted to verify connectivity control including device recognition and granting of Internet access remotely as in Figure 1(a). Several Wi-Fi devices such as smartphones and notebooks used in our environmental test bed will be identified by their MAC addresses stored in our remote service. The results of the experiment for providing network access to guests are presented in Figure 4. Figure 4 demonstrates the results of the real-time recognition and configuration in our test bed using a Web UI. The process of granting network access is done by attaching LVAP to new devices, checking its information through remote services, and finally notifying the network's owner. When the owner/administrator accepts the guest device's joining request, an IP address is provided by the DHCP service and it successfully connects to the Internet.

In contrast, if the owner/administrator rejects the request, the device will be kept in a "blocked list" and ignored by the system. In addition, the device list is updated in real time with the associated state of each device. The above results demonstrate that our system is possible to use for remote network control based on SDN abstraction, which handles clients' primitive states and programmable access points. Besides, since our access points support multiple SSIDs, we have the ability to create a virtual SSID and multiple networks from the same box. Therefore, each network can work independently for our purpose but still notifies the controller about interactive devices status.

6.2. Evaluation of Round Trip Time for Authentication. This subsection focuses on establishing a connection process for new devices in our system. The round trip time for user operations is measured from when he/she decided to choose an SSID from the device's interface until the connection between the access point and device is established. This procedure consists of two main phases: authentication and DHCP lease. In traditional IEEE 802.11, when an SSID is broadcasted by an access point, there may be a security method such as WPA2-PSK applied to protect the network. This process makes the network more secure but also requires more user operations such as typing passwords. More specifically, in the case of public Wi-Fi, users just choose an SSID and join without typing a secret phrase and the association process starting. But in the case of WPA2-PSK, we measure the time taken for users to provide the password. In our system's case, a virtual SSID without encryption was broadcasted by our access point and a similar process took place as in public Wi-Fi. However, in this case, we built a remote module in the Raspberry Pi acting as a third-party web service containing our test device information such as MAC address and device name. When the authentication happens via the above-mentioned REST API in Table 1, the controller can compare our device's identifier and grant network permission specifically. We performed this evaluation with a Wi-Fi-enabled notebook computer running Ubuntu 16.4 and repeated 10 times for each case.

The result is shown in Figure 5. In the case of the public Wi-Fi, the round trip time from the moment the terminal command for connecting to the network was executed to the time it returned successfully is between 0.79 s to 1.9 s. Besides, the round trip time for WPA2-PSK is from 0.83 s to 2.85 s. In our system, it takes a longer time from 2.15 s to 3.67 s for devices access. Thus, we clearly see that using the LVAP concept in our connectivity control system gives a slightly slower than traditional methods. We can analyze the reason based on Algorithm 1. Instead of immediately responding to the station with an assocResponse frame as usual, our system uses the station state to identify which BSSID belongs to the physical access point and LVAP. After that, the *sendNotificationToRemoteService* containing stations' MAC address calls to ask the system administrator or owner whether to permit the new associated stations or not. The wait time for owner/administrator to reply and execute this function contributes most to latency in authentication time. Besides, the time for OpenFlow rule-matching process and setting up of the Internet connection by *addOpenFlowRule* added to the round trip time. Even though it has a relatively long latency compared to traditional methods, our scheme facilitates users operations. It also means that the guest does not have to ask about the wireless network password or retype it since it can be changed easily. In addition, owners/administrators have the ability to manage

```
1 Function handleAssociationRequest (STA);
  Input: Station Information, STA
  Output: Handle association request frame
2 if APConnectionExists then
3   sta⟵getSTA(currentRequest);
4   bssid⟵getBSSID(STA);
5   lvap⟵getLVAP(STA, BSSID);
6   if lvap exists then
7     physicalAP⟵getPhysicalAP(lvap);
8     switchAP⟵getOFSwitch(physicalAP);
9     if STA exists in allowedList then
10      sendAssoResponse(STA);
11      addOpenFlowRule(switchAP, match, action);
12    else
13      if STA exists in rejectedList then
14        addOpenFlowRule(switchAP,match,action: drop);
15      else
16        //Send Authentication Request to Remote Server;
17        sendNotificationToRemoteService(STA);
18      end
19    end
20    //Updating allowed/rejected list with results returned from remote services.
21  end
22 end
```

Algorithm 1: Pseudo-code of handling association frame from associated station.

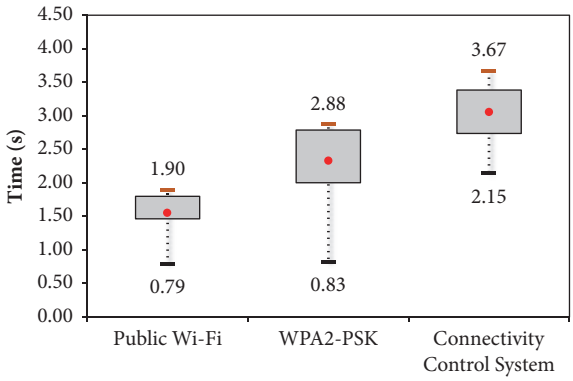


Figure 5: Actual round trip time in authentication phase with notebook.

and configure his/her network in real time remotely with different services supported by the network provider.

Note that, for the measurements in Figure 5, LVAP needs to traverse to the SDN controller, whereas for others it does not need to visit beyond the access points. For a more fair comparison, we compare our evaluation system with IEEE 802.1x Enterprise WPA, which requires traversal to the remote service for authentication and authorization. In order to conduct this experiment, we built a small authentication server running the FreeRADIUS operating system and point to this server through WPA Enterprise security mode on our access points. FreeRADIUS is an open source RADIUS server used by many organizations. It performs authentication, authorization, and accounting (AAA) functions, based on many modularized authentication protocols. Basically, in enterprise WLAN, RADIUS authentication servers are usually used to provide a username/password for each client to join the network. After providing this kind of information, clients will be issued a certificate and use it for authentication. Because only our mobile phone supports 802.1x so we compared two methods on the same device. In this situation, we wrote a simple code that contains device's authentication information and stored it on the mobile. Thus, we do not have to input this information manually on our mobile's screen, which would have taken more time. The result is shown in Figure 6. In this evaluation, while connection establishment in our system takes between 3.82 s and 6.01 s, the IEEE 802.1X authentication time is from 5.25 s to 6.72 s. The results indicate that the 802.1X authentication method takes a slightly longer time than our mechanism. While our method focuses on facilitating devices' joining the network process, 802.1X strengthens the security of the traditional WPA2-PSK authentication mechanism. Therefore, we can conclude that SDN-based connectivity control does not require large latency until it allows new Wi-Fi devices, even though it needs to traverse to the SDN controller through the access point.

Since our system follows the Odin abstract architecture and also inherits its advantages such as client mobility support, multiple logical networks on top of the physical infrastructure, and per-client virtual access points, our connectivity control platform can support these functions including the use cases in Section 4. For example, owner/administrator can apply each device policy or time access rule independently via our provided web GUI. Since devices connect to physical *AccessPoints* which are controlled by the *Controller*, these network services affect the clients. In addition, extensive researches have been conducted on efficient QoS using

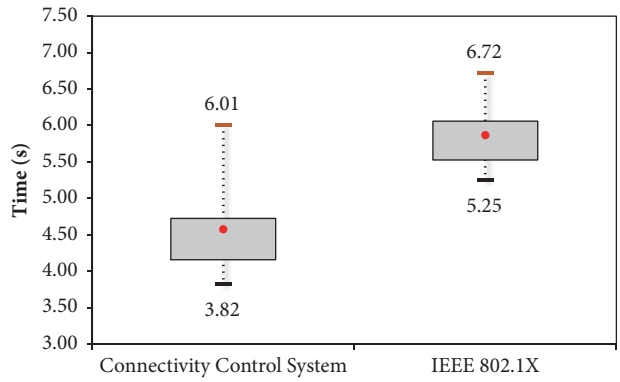


Figure 6: Actual round trip time in authentication phase with smartphone: IEEE 802.1X versus our system.

OpenFlow management in wireless networking environment such as video streaming and VoIP [33, 34]. Other advantages of SDN-based system are building an intelligent WLANs based solution which help handling massive data for future IoT network as [35, 36]. Therefore, our connectivity control system has the potential to provide QoS for applications in easier and more flexible ways compared to traditional wireless network architectures.

7. Security Enhancements

The main goal of this paper is to enable quick and easy connectivity control of Wi-Fi devices including guest devices and IoT devices as discussed in previous sections. In other words, security is not main concern of this paper, but it is an essential requirement since the target system, Wi-Fi, of the proposed solution is commercial. Therefore, this section discusses how the proposed connectivity control system in Section 5 provides security functions such as authentication of devices as well as encryption of packets.

First of all, it is possible to doubt that the connectivity control system is vulnerable to the MAC cloning and spoofing attack, since the MAC address of a device is a unique identifier. Indeed, the device is registered with only MAC address in the proposed system. However, it is interesting to note that the connectivity control system requires intervention from the network owner or administrator, who can check the identity of device through offline directly or indirectly. In detail, the main use cases are guest users and IoT devices in home and office, as discussed in Section 4. In most cases when accepting the new connection from guests or installing new IoT devices in home, the network owner is located in the same place with them. Even though the network owner/administrator delegates its management role to service provider due to any reasons such as lack of knowledge, the owner needs to be connected to the service provider to confirm the identity of a newly connecting device. Therefore, we can prevent spoofing attack from malicious device at the initial connection phase through the identity check in offline. In addition, the proposed system allows us to keep monitoring the status of client through LVAP. Note that our system assigns a unique LVAP to individual client. Moreover, as introduced in Figure 2, the *DeviceManager* module not only verifies whether the connecting device's MAC address is contained in the *AllowedList*, but also checks whether a new spawn LVAP is bound to the current access point or not. This indicates that when a client moves to other locations and disconnects/reconnects to other access points, it immediately triggers the notification system and announces to the system owner. Therefore, we can conclude that the proposed connectivity control system has many options to prevent diverse attacks based on device identification.

The next security issue is about whether the proposed system can establish a secure connection from Wi-Fi devices or not. Note that Section 5 describes the connectivity control in an open network environment, which may result in diverse vulnerabilities from nonencryption of packets. Therefore, we need to discuss the feasibility to apply the existing security protocols such as WPA and WPA2 to our system. Before describing the security enhancement, it is important to remind that the proposed system is designed and targeted to IoT devices which are not equipped with any input and output interfaces. These security protocols such as WPA and WPA2 require a connecting Wi-Fi device to deliver its id and password information, but it is impossible to input this information manually. There exist contradictions between the assumption of the proposed system and the requirement of security enhancement. The only way to fill the gap is to assume that the IoT devices including guest Wi-Fi devices are equipped with preinstalled application, which is used for the connection phase, with help from manufacturers. Since the quick and easy connection management of increasing number of IoT devices is one of the important issues, this is not infeasible assumption.

The proposed security enhancement utilizes that access point can create multiple SSIDs. As shown in Figure 7, the client connects to the open SSID, and it reconnects to the secure SSID after successfully being identified and granted permission by the system. In detail, the client device connects to the access point the same as the procedure in Figure 2. Whereas it finishes the connection establishment once the device is accepted by system administrator in Section 5, the security enhancement requires additional procedures. First, the client device, especially the preinstalled application in the client devices, has to transfer a public key to the access point. Second, the access point replies with information about secure SSID which is encrypted with public key received. Finally, the preinstalled application in client device decrypts the information using its private key, and it changes the connection to the secure SSID. Note that it is highly required to apply lightweight public key cryptography mechanism for IoT devices due to their limited resource, but we can conclude that the proposed security enhancement mechanism is feasible to apply to the proposed connectivity control system and provide the similar level of security with existing solutions.

8. Conclusion

With the explosion of tiny IoT devices and diversification of complicated security options, it becomes hard to configure the connection information of Wi-Fi devices. In this paper,

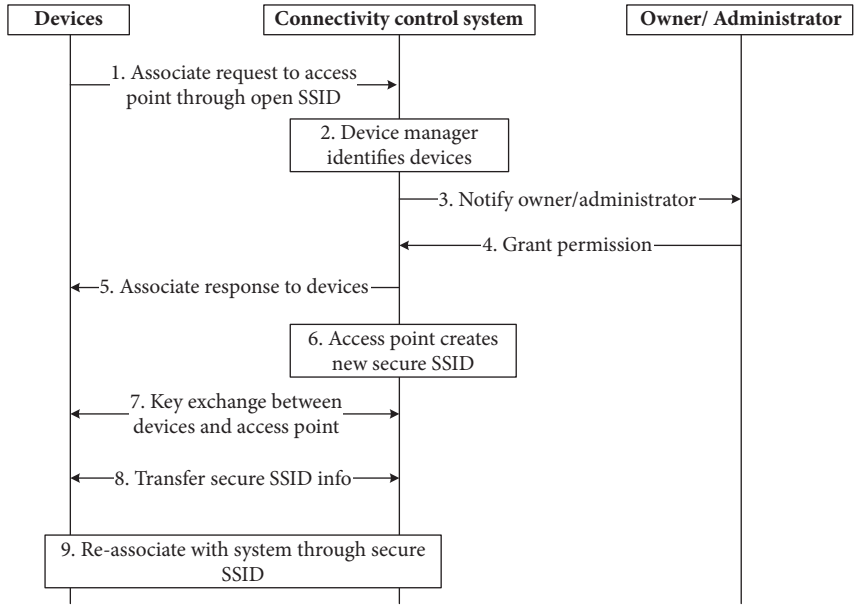


Figure 7: Procedure of secure connection establishment in the connectivity control system.

we have proposed an SDN-based connectivity control system to facilitate connectivity control of Wi-Fi devices as well as providing a high-level authentication. The proposed system allows the network owner or administrator to manage Wi-Fi devices through remote service, and it does not require any configuration information setup to the devices. We have proved the feasibility of the proposed system through the real testbed, and we expect the proposed connectivity control system to be widely utilized to enhance the experience of both users and manufacturers.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1D1A1B03933007).

References

[1] ZigBee, ZigBee Alliance, https://www.zigbee.org/.

[2] Z-Wave, https://www.z-wave.com/.

[3] Bluetooth, The global standard for connection, https://www.bluetooth.com/.

[4] S. Bera, S. Misra, and A. V. Vasilakos, "Software-Defined Networking for Internet of Things: A Survey," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1994–2008, 2017.

[5] D. B. Rawat and S. Reddy, "Recent advances on Software Defined Wireless Networking," in *Proceedings of the Southeast-Con 2016*, Norfolk, VA, USA, April 2016.

[6] A. Gudipati, D. Perry, L. E. Li, and S. Katti, "Softran: software defined radio access network," in *Proceedings of the 2nd ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN '13)*, pp. 25–30, August 2013.

[7] C. Xu, W. Jin, G. Zhao, H. Tianfield, S. Yu, and Y. Qu, "A Novel Multipath-Transmission Supported Software Defined Wireless Network Architecture," *IEEE Access*, vol. 5, pp. 2111–2125, 2017.

[8] M. Casado, N. Foster, and A. Guha, "Abstractions for software-defined networks," *Communications of the ACM*, vol. 57, no. 10, pp. 86–95, 2014.

[9] P. Ruckebusch, S. Giannoulis, D. Garlisi et al., "WiSHFUL: Enabling Coordination Solutions for Managing Heterogeneous Wireless Networks," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 118–125, 2017.

[10] Samsung Smart Home, https://www.samsung.com/smart-home/smartthings/.

[11] Apple Home Kit, https://www.apple.com/shop/accessories/all-accessories/homekit.

[12] Philips, https://www2.meethue.com/.

[13] M. Lee, Y. Kim, and Y. Lee, "A home cloud-based home network auto-configuration using SDN," in *Proceedings of the 12th IEEE International Conference on Networking, Sensing and Control, ICNSC 2015*, pp. 444–449, Taipei, Taiwan, April 2015.

[14] V. Sivaraman, H. H. Gharakheili, A. Vishwanath, R. Boreli, and O. Mehani, "Network-level security and privacy control for smart-home IoT devices," in *Proceedings of the 11th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2015*, pp. 163–167, October 2015.

[15] K.-K. Yap et al., *Separating Authentication, Access and Accounting: A Case Study with OpenWifi*, Stanford University NEC OPENFLOW-TR2011-1, 2011.

[16] Z. Cao, J. Fitschen, and P. Papadimitriou, "FreeSurf: application-centric wireless access with SDN," in *Proceedings of the 17th*

IEEE International Conference on High Performance Switching and Routing, HPSR 2016, pp. 206–212, June 2016.

[17] T. Enghardt, *Authentication, Authorization and Mobility in Openflow-enabled Enterprise Wireless Networks[Master thesis]*, Technische Universitat, Berlin, Germany, 2014.

[18] J. Schulz-Zander et al., "OpenSWDN: Programmatic Control over Home and Enterprise WiF," in *Proceedings of the ACM HotSDN*, pp. 25–30, 2013.

[19] R. Riggio, M. K. Marina, J. Schulz-Zander, S. Kuklinski, and T. Rasheed, "Programming abstractions for software-defined wireless networks," *IEEE Transactions on Network and Service Management*, vol. 12, no. 2, pp. 146–162, 2015.

[20] H. Gacanin and A. Ligata, "Wi-Fi self-organizing networks: Challenges and use cases," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 158–164, 2017.

[21] L. Suresh, J. Schulz-Zander, R. Merz, A. Feldmann, and T. Vazao, "Towards programmable enterprise WLANS with Odin," in *Proceedings of the 1st ACM International Workshop on Hot Topics in Software Defined Networks (HotSDN '12)*, pp. 115–120, Helsinki, Finland, August 2012.

[22] N. McKeown, T. Anderson, H. Balakrishnan et al., "OpenFlow: enabling innovation in campus networks," *Computer Communication Review*, vol. 38, no. 2, pp. 69–74, 2008.

[23] E. Coronado, J. Villalon, and A. Garrido, "Wi-balance: SDN-based load-balancing in enterprise WLANs," in *Proceedings of the 2017 IEEE Conference on Network Softwarization, NetSoft 2017*, Bologna, Italy, July 2017.

[24] Z. Yang, J. Zhang, K. Tan, Q. Zhang, and Y. Zhang, "An adaptive mobility manager for Software-Defined Enterprise WLANs," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '15)*, pp. 1436–1444, Kowloon, Hong Kong, 2016.

[25] F. De Turck, P. Chemouil, W. Kellerer et al., "Guest editors' introduction: special issue on advances in management of softwarized networks," *IEEE Transactions on Network and Service Management*, vol. 13, no. 3, pp. 786–791, 2016.

[26] R. Riggio, T. Rasheed, and F. Granelli, "EmPOWER: a testbed for network function virtualization research and experimentation," in *Proceedings of the Workshop on Software Defined Networks for Future Networks and Services (SDN4FNS '13)*, pp. 1–5, IEEE, Trento, Italy, November 2013.

[27] OpenWrt, http://openwrt.org/.

[28] E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. F. Kaashoek, "The click modular router," *ACM Transactions on Computer Systems*, vol. 18, no. 3, pp. 263–297, 2000.

[29] B. Pfaff, "Extending networking into the virtualization layer," in *Proceedings of the ACM HotNets*, pp. 1–17, 2009.

[30] OpenFlow, OpenFlow Specifications, https://opennetworking.org/.

[31] Raspberrypi, Teach, learn and make with Raspberry Pi, https://www.raspberrypi.org/.

[32] FreeRADIUS, The FreeRADIUS Server Project, https://freeradius.org/.

[33] M. Karakus and A. Durresi, "Quality of Service (QoS) in Software Defined Networking (SDN): A survey," *Journal of Network and Computer Applications*, vol. 80, pp. 200–218, 2017.

[34] D. Das, J. Bapat, and D. Das, "A Dynamic QoS Negotiation Mechanism between Wired and Wireless SDN Domains," *IEEE Transactions on Network and Service Management*, 2017.

[35] D. Tu, Z. Zhao, and H. Zhang, "ISD-WiFi: An intelligent SDN based solution for enterprise WLANs," in *Proceedings of the 8th International Conference on Wireless Communications and Signal Processing, WCSP 2016*, Yangzhou, China, October 2016.

[36] M. S. Carmo, S. Jardim, T. De Souza, A. V. Neto, R. Aguiar, and D. Corujo, "Towards enhanced connectivity through WLAN slicing," in *Proceedings of the 16th Annual Wireless Telecommunications Symposium, WTS 2017*, Chicago, IL, USA, April 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 2484268, 9 pages
https://doi.org/10.1155/2018/2484268

Research Article

A Provably Secure Anonymous Authenticated Key Exchange Protocol Based on ECC for Wireless Sensor Networks

Ke Zhang,[1,2] **Kai Xu**,[3] **and Fushan Wei**[2]

[1]*Network Information Center, Shaanxi Normal University, Xi'an 710062, China*
[2]*State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450002, China*
[3]*School of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, China*

Correspondence should be addressed to Ke Zhang; kezhang2017@163.com

Received 22 January 2018; Revised 7 April 2018; Accepted 19 April 2018; Published 16 July 2018

Academic Editor: Ding Wang

In wireless sensor networks, users sometimes need to retrieve real-time data directly from the sensor nodes. Many authentication protocols are proposed to address the security and privacy aspects of this scenario. However, these protocols still have security loopholes and fail to provide strong user anonymity. In order to overcome these shortcomings, we propose an anonymous authenticated key exchange protocol based on Elliptic Curves Cryptography (ECC). The novel protocol provides strong user anonymity such that even the gateway node and the sensor nodes do not know the real identity of the user. The security of the proposed protocol is conducted in a well-defined security model under the CDH assumption. Compared with other related protocols, our protocol is efficient in terms of communication and enjoys stronger security. The only disadvantage is that our protocol consumes more computation resources due to the usage of asymmetric cryptography mechanisms to realize strong anonymity. Consequently, our protocol is suitable for applications which require strong anonymity and high security in wireless sensor networks.

1. Introduction

1.1. Background. As an important part of Internet of things, wireless sensor networks (WSNs) draw more and more attentions from the industrial and the researchers. Typically, a WSN is composed of numerous tiny sensor nodes. These sensor nodes can be deployed in unattended or hostile environments to collect valuable data of interest. For example, a large amount of visual data such as images and videos can be collected by visual sensor nodes [1]. These nodes have the characteristics of easy deployment, low cost, and high mobility [2]. Due to these merits, WSNs are very useful in many application scenarios, such as natural disaster prevention, machine health monitoring, air temperature monitoring, health care monitoring, and battlefield surveillance.

Usually, the data collected by the sensor nodes will be transmitted to and aggregated by a gateway node periodically. Whenever a user wants to get access to the aggregated data from the gateway node, he will authenticate himself to the gateway node. However, in some application scenarios such as health care monitoring and battlefield surveillance, users have great needs to access the data directly from the sensor nodes. Under such circumstances, the user first sends a request to the gateway node for accessing the real-time data. The gateway will authenticate whether the user is valid or not. If the user is valid, a common session key will be established between the user and the sensor node with the help of the gateway node. The session key can later be used to protect confidentiality and integrity of the data [3].

1.2. Related Work. In order to address the security aspects under the above-mentioned application scenario, many authentication protocols are proposed [4–7]. In 2009, Das [8] proposed a two-factor user authentication protocol, which is claimed to have strong authentication and session key establishment and achieves efficiency. Unfortunately, Khan et al. [9] soon found that Das's protocol is vulnerable to the gateway node bypassing attack and the privileged-insider attack. Besides, Das's protocol also fails to provide password update and mutual authentication. Khan et al. also presented

an improved protocol to fix the susceptibilities of the original protocol. In 2011, Yeh et al. [10] pointed out several weaknesses of Das's protocol. They also design an ECC-Based authentication protocol to meet the needs of applications with higher security requirements. In 2013, Xue et al. [11] proposed a temporal-credential-based mutual authentication scheme among the user, the gateway node, and the sensor node. A credential is issued by the gateway node to each user for authentication. Their protocol only involves lightweight operations such as XOR and hash and is suitable for resource-constrained WSNs. Nevertheless, He et al. [12] soon pointed out that Xue et al.'s protocol is vulnerable to the offline password guessing attack, the user impersonation attack, the sensor node impersonation attack, and the modification attack. They also proposed an improved temporal-credential-based protocol to remedy the weaknesses. Yuan et al. [13] proposed an authentication scheme for WSNs based on pubic key mechanism and biometric characteristics of the user to realize strong authentication. In 2014, Wang et al. [14] analyzed two authentication schemes for WSNs and demonstrated several loopholes. They also investigate the underlying rationale of the security failures and put forward three basic principles for designing secure authentication protocols in WSNs. In 2016, Shen et al. [15] proposed an efficient multilayer authentication protocol and a secure session key generation method for WSNs. They also designed a one-to-many group authentication protocol and a certificate-less authentication protocol, which is of independent interest.

Recently, researchers begin to focus on user's privacy protection in WSNs. Wu et al. [16] proposed an anonymous authentication scheme based on ECC for WSNs with formal security proof. Jiang et al. [17] designed an anonymous lightweight three-factor authentication scheme for WSNs. The security of their protocol is conducted using ProVerif. Wang et al. [20] put forward a new authentication which can resist all known attacks for WSNs. Moreover, they explore the design principle of authentication schemes. They also designed a biometric-based authentication scheme and proved its security using the Burrows-Abadi-Needham (BAN) logic [18]. Li et al. [19] proposed a three-factor anonymous authentication scheme for WSNs. They use fuzzy commitment to deal with the user's biometric template.

1.3. Motivation and Contribution. Until now, there are many authentication protocols designed for protection of security and privacy when accessing real-time data in WSNs. However, there are still some problems which are not solved. Firstly, most of these protocols only have informal heuristic security arguments. It is quite common such that a protocol which is claimed to be secure is soon found to be vulnerable to several attacks. What is worse, the improved protocol still has many vulnerabilities. Secondly, the existing protocol pays little attention to user's privacy. Only few protocols provide user anonymity and these protocols only achieve weak anonymity; i.e., the real identity is hidden to an adversary but is known to the gateway node and sometimes even the sensor node knows the real identity of the user. Last but not least, the existing protocols rely on efficient XOR, symmetric encryption, and hash operations to provide better efficiency. Although these protocols can meet the characteristics of constrained storage, computation, and communication capabilities of sensor nodes, they fail to provide strong security guarantee. For security-critical applications such as battlefield surveillance, security and privacy are more important than the computation and communication efficiency.

In this paper, we investigate the design of anonymous and strongly secure authenticated key exchange protocol in WSNs. We propose an efficient authenticated key exchange protocol for the scenario in which the user wants to access the real-time data directly from the sensor node. The novel protocol has the following advantages. First of all, our protocol enjoys formal security proof in a well-defined security model. The security is conducted in the random oracle model under the CDH assumption. Second, our protocol provides the strongest anonymity in the sense the real identity of the user is only known to himself. Neither the gateway node nor the sensor node can obtain any information of the user's identity, not to mention the adversary. Thirdly, our protocol achieves more security attributes than other related protocols. Consequently, it is more secure than other related protocol and is particularly suitable for security-critical applications in WSNs. The only disadvantage is that it needs more computation resources. However, security and privacy are more important than the computation efficiency in security-critic application. So our protocol is suitable for security-critical applications in WSNs.

The rest of the paper is organized as follows. In Section 2, we present the security model and some preliminaries. We describe the details of the proposed protocol in Section 3. The security proof is given in the random oracle model in Section 4. The performance comparison with other related protocols is summarized in Section 5. We conclude this paper in Section 6.

2. Security Model

In this section, we briefly recall the security model presented in [21, 22]. The security of our protocol will be conducted in this formal security model.

Protocol Participants. The participants of an authentication and key exchange protocol for real-time data retrieval in WSNs involves users U, a gateway node GN, and a sensor node S_i. Each user U registers with the gateway node and each sensor node S_i shares a common secret key with the gateway node.

Protocol Execution. All the participants are modeled as a PPT Turing machine. The i-th instance of a participant P is denoted by P^i. All the communication channels are managed by a probabilistic polynomial time adversary $\mathcal{A}$. The adversary $\mathcal{A}$ can intercept, delay, modify, and even forge a message at will. The capabilities of the adversary are captured through oracle queries. The adversary can make the following oracle queries:

(i) $Execute(U^x, GN^y, S_i^z)$: the execution query captures the passive eavesdrop ability of $\mathcal{A}$. In reply to this oracle query, $\mathcal{A}$ will get all the transcripts of

the authentication instance executed among a user instance U^x, a gateway node instance GN^y, and a sensor node instance S_i^z.

(ii) $Send(P^i, m)$: the send query captures the active attack ability of $\mathscr{A}$. Through the $Send(P^i, m)$ query, $\mathscr{A}$ sends a modified or forged message m to instance P^i in the name of another participant instance. $\mathscr{A}$ will get the message generated by the participant instance P^i upon receiving the message m according to the description of the protocol. The participant P can be a user, a gateway node, or a sensor node.

(iii) $Corrupt(U, PW)$: this query captures the compromise of the user's password. The adversary $\mathscr{A}$ only gets the password of the victim user; it can neither control nor compromise the credential of the user.

(iv) $Corrupt(U, cred)$: this query captures the compromise of the user's terminal. The adversary $\mathscr{A}$ can extract the credential issued by the gateway node and control the victim user's terminal. However, the password of the user is still unknown to $\mathscr{A}$.

(v) $Corrupt(S_i)$: this query captures the compromise of a sensor node S_i. The adversary $\mathscr{A}$ will get the secret key and control the sensor node through this query.

(vi) $Reveal(P^i)$: this query can only be asked to a user instance or a sensor node instance. If the instance P^i accepts the session and generates a session key, $\mathscr{A}$ will get the session key. Otherwise, $\mathscr{A}$ will get the symbol $\perp$ which means the instance P^i does not hold a session key.

(vii) $Test(P^i)$: this query does not capture any real attack ability of $\mathscr{A}$ but is used to measure the security of the session key held by instance P^i. Upon receiving this query, the simulator will flip a coin b. If the result is 1, then it returns the real session key to $\mathscr{A}$. If the result is 0, the simulator will send a random session key of the same length with the real session key to $\mathscr{A}$. $\mathscr{A}$ has to distinguish if the key is real or random. In other words, $\mathscr{A}$ has to guess the coin flip result.

The session identification (sid) is defined as the transcripts shared between a user instance and a sensor node instance. The partner identification (pid) of an instance is defined to be the participant with whom the instance wants to establish a common session key. We say a user instance U^x and a sensor node instance S_i^z are partners if the following conditions are satisfied: (1) these two instances both accept and generate the same session key; (2) these two instances share the same sid; (3) the pid of U^x is S_i and the pid of S_i^z is U; and (4) no other instances accept the same sid with U^x and S_i^z.

If the adversary $\mathscr{A}$ asks both $Corrupt(U, cred)$ and $Corrupt(U, PW)$, the user MU is said to be fully corrupted. When defining the AKE security of the session key, we do not consider the corruption of the gateway node. This is because once the gateway node is corrupted; there is nothing we can do to guarantee the security of the protocol. A user instance or a sensor node instance P^i is said to be fresh if (1) $\mathscr{A}$ does not send *Reveal* queries to the instance or its partner; and (2) the user or the sensor node is not fully corrupted by $\mathscr{A}$.

AKE Security. The security of the session keys is captured by the AKE security. The adversary $\mathscr{A}$ is restricted to ask *Test* queries to fresh instances only; otherwise the adversary $\mathscr{A}$ can trivially win the attack game. The adversary $\mathscr{A}$ is given access to all the oracle queries; the only restriction is that $\mathscr{A}$ only can ask one *Test* query to a fresh instance. The adversary $\mathscr{A}$ needs to guess the hidden bit b used by the simulator when answering the *Test* query. If $\mathscr{A}$ correctly guesses the random bit, then we say $\mathscr{A}$ wins the AKE security game. We denote this event by *Succ*. With respect to the distribution of the passwords, we use the Zipf's law put forward by Wang et al. [21] instead of assuming a uniform distribution. The adversary $\mathscr{A}$'s advantage in attacking the AKE security of a protocol $\mathscr{P}$, when passwords are chosen according to the Zipf's law of a dictionary D, is defined as follows:

$$Adv_{\mathscr{P},\mathscr{D}}^{ake}(\mathscr{A}) = 2 \cdot Pr[Succ] - 1 \quad (1)$$

An authentication and key exchange protocol $\mathscr{P}$ is said to be AKE secure if for all PPT adversary $\mathscr{A}$, the advantage $Adv_{\mathscr{P},\mathscr{D}}^{ake}(\mathscr{A})$ is only negligible larger than $C' \cdot q_{send}^{s'}$, where C' and s' are Zipf parameters and q_{send} is the number of active attack sessions. Moreover, C' and s' are constants depending on the password data set and can be calculated by linear regression.

3. Description of the Protocol

In this section, we describe the proposed anonymous authenticated key exchange protocol based on ECC for WSNs. The most important benefit of ECC is that it provides the same level of security with a smaller key size compared to other cryptography mechanisms such as RSA. So it suits the needs of the resource-constrained nature of the WSN. Our protocol has three phases: the setup phase, the registration phase, the authentication, and key exchange phase. The detailed steps of each phase are described in the following. The symbols used in this paper are summarized in Table 1.

3.1. The Setup Phase. Let p be a large prime and F_p be a finite field of prime order p. Let E be an elliptic curve cryptosystem satisfying the equation $y^2 = (x^3 + ax + b) \bmod p$ such that $a, b \in F_p$ and $4a^3 + 27b^2 \bmod p \neq 0$. The set of rational points in E over finite field F_p is denoted by $E(F_p)$. More precisely, $E(F_p) = \{(x; y) : x, y \in F_p$ such that $y^2 = (x^3 + ax + b) \bmod p\} \cup \{O\}$, where O is the point at infinity. Let G be a cyclic group generated by P, where $P \in E(F_p)$ has a large prime order n. These parameters $(F_p, E, E(F_p), G, P)$ are the system parameters and can be chosen by a trusted third party or the gateway node. The gateway node (GN) chooses a random number $s_{GN} \in Z_n^*$ as his private key and computes the corresponding public key $Q_{GN} = s_{GN}P$. The public key Q_{GN} is published in the whole network. Define six hash functions such that $H_1 : \{0,1\}^* \longrightarrow Z_n^*$,

Table 1: Notations.

notation	meaning	notation	meaning
ID_{GW}	identity of the gateway node	ID_U	identity of the user U
ID_{S_i}	identity of the sensor node S_i	p, n	large prime numbers
F_p	a finite field	E	an elliptic curve defined on F_p
$E(F_p)$	the set of rational points in E	s_{GN}	secret key of the gateway node
PW_U	the password of the user U	$\oplus$	exclusive OR
$\|$	concatenation	$h(m)$	cryptographic hash of message m
$sign_{s_{GN}(m)}$	signature of m singed by s_{GN}	T_{GN/S_i}	timestamp of GN/S_i

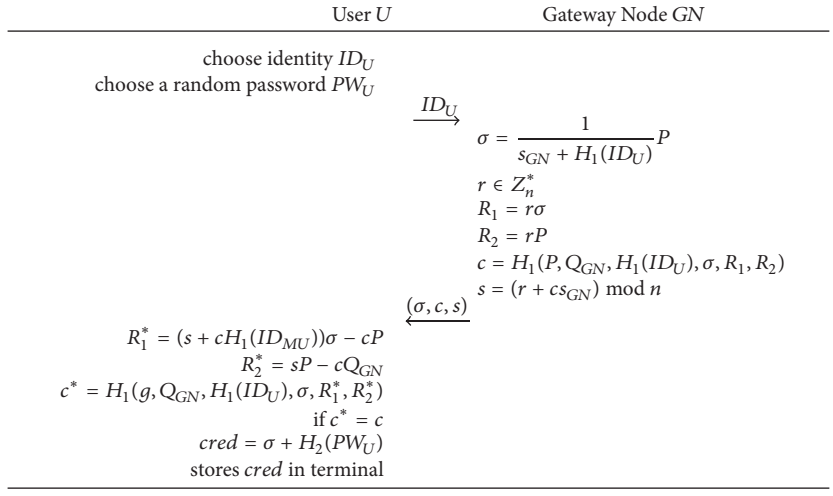


Figure 1: Registration phase of mobile user.

$H_2 : \{0,1\}^* \longrightarrow G^*$, $H_0, H_3, H_4, H_5 : \{0,1\}^* \longrightarrow \{0,1\}^\kappa$, where κ is the security parameter. All these parameters $(F_p, E, E(F_p), G, P, Q_{GN}, H_i(i = 0, 1, \ldots, 5))$ are available to all the entities in the WSN.

3.2. The Registration Phase. If a user U wants to access the data collected by the sensor nodes in the WSN, U has to register himself to the gateway node. For a pictorial illustration of the user registration, please refer to Figure 1. The detailed steps are described in the following.

Step 1. The user U randomly chooses his identity ID_U and his password PW_U from the password dictionary. U sends his identity ID_U to the gateway node GN through a secure channel.

Step 2. When the gateway node GN receives the registration request from the user, GN verifies the validity of the $U's$ identity ID_U. If it is valid and there is no other user in its database registers using the same identity, GN first computes the credential $\sigma = (1/(s_{GN} + H_1(ID_U)))P$. Then GN chooses a random number $r \in Z_n^*$ and computes $c = H_1(P, Q_{GN}, H_1(ID_U), \sigma, R_1, R_2)$ and $s = (r + cs_{GN}) \bmod n$, where $R_1 = r\sigma$, $R_2 = rP$. At last, GN sends the registration message (σ, c, s) to the user U through a secure channel.

Step 3. When the user U receives the registration message (σ, c, s) from GN, U will verify the validity of the message. U computes $R_1^* = (s + cH_1(ID_U))\sigma - cP$, $R_2^* = sP - cQ_{GN}$, and $c^* = H_1(g, Q_{GN}, H_1(ID_U), \sigma, R_1^*, R_2^*)$. U verifies whether c^* is equal to c or not. If the verification is successful, U will accept σ as a valid credential. Finally, U computes $cred = \sigma + H_2(PW_U)$ and then stores his password-protected credential $cred$ in his terminal.

The registration of the sensor node is rather simple compared with the user registration. The sensor node S_i sends the registration request to the gateway node GN through a secure channel. Upon receiving the request, the gateway node GN will compute a symmetric key $K_{(GN,S_i)} = H_3(GN, S_i, s_{GN})$ and send the symmetric key $K_{(GN,S_i)}$ to S_i through a secure channel.

3.3. The Authentication and Key Exchange Phase. Suppose a user U wants to get the real-time data from the sensor node S_i, U has to execute the authentication and key exchange phase with the gateway node GN and the sensor node S_i. During this phase, the user U, the gateway node GN, and the sensor node S_i will authenticate each other. At the end of this phase, a session key will be established between U and S_i to protect the upcoming data transmission. The detailed steps

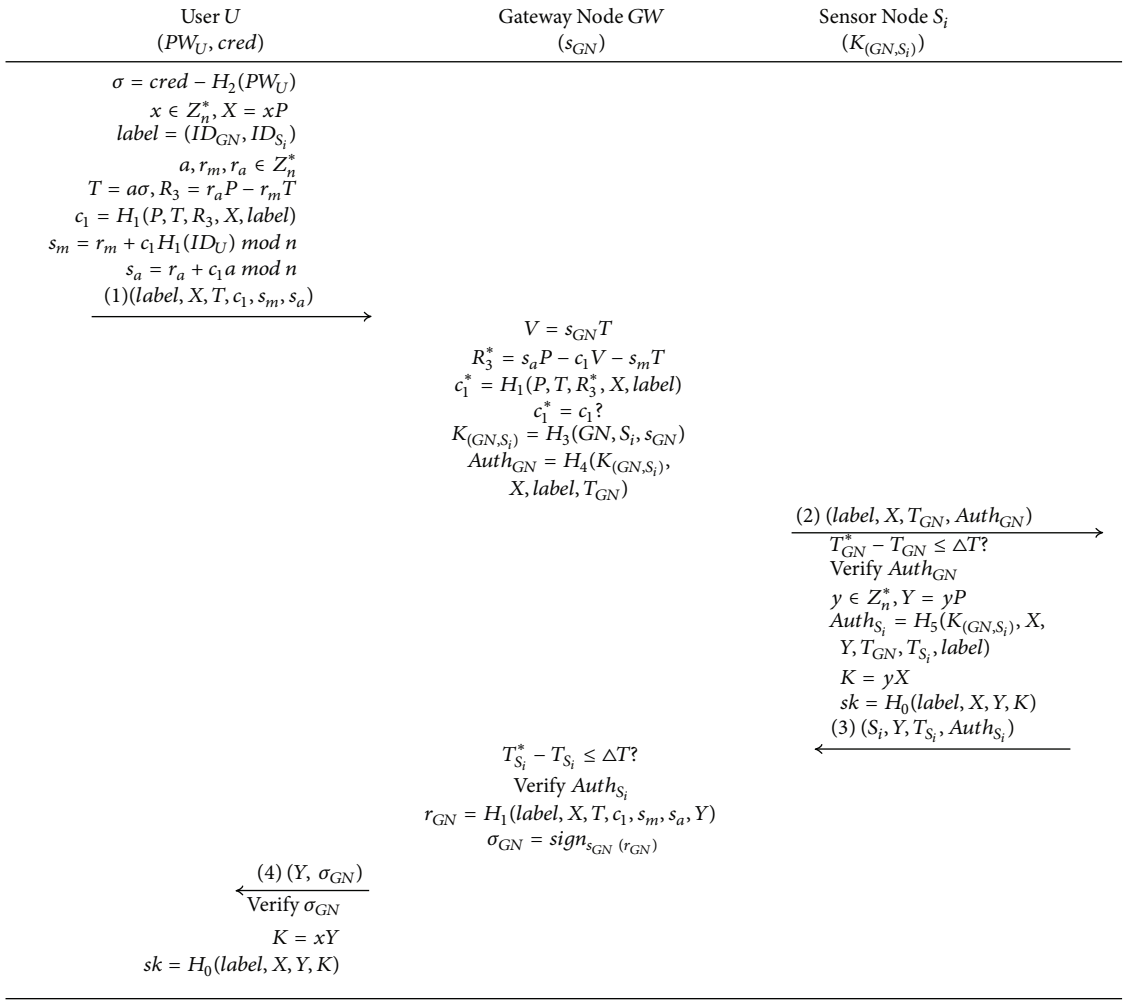


Figure 2: Authentication and key exchange phase.

of the authentication and key exchange phase are described as follows. For a pictorial illustration, please refer to Figure 2.

Step 1. The user U types his password PW_U to his terminal. The terminal will compute $H_2(PW_U)$ and recovers the credential σ from the stored $cred$. U then chooses a random number $x \in Z_n^*$ and computes $X = xP$. U defines the label of this session as $label = (ID_{GN}, ID_{S_i})$. U chooses three random numbers $a, r_m, r_a \in Z_n^*$ and computes $T = a\sigma$, $R_3 = r_aP - r_mT$, $c_1 = H_1(P, T, R_3, X, label)$, $s_m = r_m + c_1 H_1(ID_U) \bmod n$, and $s_a = r_a + c_1 a \bmod n$. Finally, U sends the message $(label, X, T, c_1, s_m, s_a)$ to the gateway node GN.

Step 2. Upon receiving the message $(label, X, T, c_1, s_m, s_a)$ from the user, GN needs to authenticate the user U. GN computes $V = s_{GN}T$, $R_3^* = s_aP - c_1V - s_mT$, and $c_1^* = H_1(P, T, R_3^*, X, label)$. GN checks whether c_1^* is equal to c_1 or not. If the verification is successful, GN authenticates the user U and believes the user U is a valid user. GN then computes the shared key with the sensor node $K_{(GN,S_i)} = H_3(GN, S_i, s_{GN})$ and the authenticator $Auth_{GN} = H_4(K_{(GN,S_i)}, X, lable, T_{GN})$, where T_{GN} is the current timestamp of GN. Finally, GN send the message $(label, X, T_{GN}, Auth_{GN})$ to the sensor node S_i.

Step 3. Upon receiving the message $(label, X, T_{GN}, Auth_{GN})$ from GN at time T_{GN}^*, the sensor node S_i first checks whether $|T_{GN}^* - T_{GN}| \le \triangle T$, where $\triangle T$ is the expected time interval for the transmission delay. If this is true, S_i then verifies the validity of the authenticator $Auth_{GN}$ using its private key $K_{(GN,S_i)}$. If the authenticator is valid, S_i chooses a random number $y \in Z_n^*$ and computes $Y = yP$. S_i then computes the authenticator $Auth_{S_i} = H_5(K_{(GN,S_i)}, X, Y, T_{GN}, T_{S_i}, label)$, where T_{S_i} is the current timestamp of S_i. S_i computes the Diffie-Hellman key $K = yX$ and the session key $sk = H_0(label, X, Y, K)$. Finally, S_i sends the message $(S_i, Y, T_{S_i}, Auth_{S_i})$ to the gateway node GN.

Step 4. Upon receiving the message $(S_i, Y, T_{S_i}, Auth_{S_i})$ from S_i at time $T_{S_i}^*$, GN first checks whether $|T_{S_i}^* - T_{S_i}| \le \triangle T$, where $\triangle T$ is the expected time interval for the transmission delay. If this is true, GN then computes the shared key with the sensor node $K_{(GN,S_i)} = H_3(GN, S_i, s_{GN})$ and verifies the validity of the authenticator $Auth_{S_i}$. If the verification is successful, GN

computes $r_{GN} = H_1(label, X, T, c_1, s_m, s_a, Y)$ and signs the random number r_{GN} using his private key s_{GN}; the signature is denoted by σ_{GN}. Finally, GN sends the message (Y, σ_{GN}) to the user U.

Step 5. Upon receiving the message (Y, σ_{GN}) from GN, U first verifies the validity of the signature σ_{GN}. U computes the random number $r^*_{GN} = H_1(label, X, T, c_1, s_m, s_a, Y)$ and checks if σ_{GN} is a valid signature for r^*_{GN} singed by GN. If the verification is successful, U computes the Diffie-Hellman key $K = xY$ and the session key $sk = H_0(label, X, Y, K)$. U will accept the session and wait for the upcoming communication.

4. Security Proof

In this section, we present the security proof of our protocol. The security proof is conducted in the security model presented in Section 2.

Theorem 1. *Suppose $\mathcal{P}$ is the anonymous authentication and key exchange protocol for WSN described in the previous section and $\mathcal{A}$ is a PPT time adversary against the AKE security of $\mathcal{P}$ who runs in time t and makes at most Q_{send} queries of Send oracle to different instances. If the signature scheme used in our protocol is existential unforgeable against adaptive chosen message attacks, the hash functions $H_i(\cdot)(i = 0, 2, \ldots, 5)$ are all modeled as random oracles; then under the CDH assumption, the advantage of the adversary $\mathcal{A}$ in violating the AKE security of the protocol $\mathcal{P}$ is at most*

$$Adv^{ake}_{\mathcal{P},\mathcal{D}}(\mathcal{A}) \le C' \cdot Q^{s'}_{send} + negl(\kappa). \quad (2)$$

Proof. We use the hybrid experiments technique to prove Theorem 1. These hybrid experiments start with the real attack scenario. We gradually change the simulation rules in each experiment. In the last experiment, the advantage of the adversary in distinguishing the session key is negligible. We also estimate the advantage difference of the adversary between two hybrid experiments and the advantage of the adversary in breaking the AKE security can be calculated. We denote the adversary's advantage in hybrid Exp_i using $Adv_i(\mathcal{A})$. □

Experiment Exp_0. This is the real attack scenario defined in the security model. In this experiment, the adversary has access to all the oracles. According to the definition of $\mathcal{A}$'s advantage, we have the following result:

$$Adv^{ake}_{\mathcal{P},\mathcal{D}}(\mathcal{A}) = Adv_0(\mathcal{A}). \quad (3)$$

Experiment Exp_1. In this experiment, we simulate all the hash function $H_i(\cdot)(i = 0, 1, \ldots, 5)$ by maintaining hash lists $\Lambda_{H_i}(i = 0, 1, \ldots, 5)$ using the following rules:

(i) On a query $H_i(m)$, if a record (i, m, r) exists in Λ_{H_i}, then return r. Otherwise, the output r is chosen according to the following rule: $RuleH_i$.

if $i = 1$, choose a random element r from z^*_n. Then add the record $(1, m, r)$ to Λ_{H_i}.

if $i = 2$, choose a random element r from G. Then add the record $(2, m, r)$ to Λ_{H_2}.

if $i = 0, 3, 4, 5$, choose a random element r from $\{0, 1\}^\kappa$. Then add record (i, m, r) to Λ_{H_i}.

In addition to these lists, we also simulate six private hash oracles $H'_i(i = 0, 1, \ldots, 5)$ by maintaining hash lists $\wedge'_{H_i}(i = 0, 1, \ldots, 5)$. We will use these private hash functions in the following hybrid experiments. It is well known that a hash function can be simulated perfectly in PPT time using the above rules; thus, we have

$$\left|Adv_1(\mathcal{A}) - Adv_0(\mathcal{A})\right| \le negl(\kappa). \quad (4)$$

Experiment Exp_2. In this experiment, we cancel the sessions if some unlikely collisions occur in these sessions. To be more specific, if some collisions occur in the simulation of the hash functions or on the transcripts of $((X, Y, T, c_1, s_m, s_a, \sigma_{GN}))$, we will terminate the session and let the adversary win. Based on the birthday paradox, we have the following result:

$$\left|Adv_2(\mathcal{A}) - Adv_1(\mathcal{A})\right| \le negl(\kappa). \quad (5)$$

Experiment Exp_3. In this experiment, we modify the simulation rules of sessions by *Execute* queries. Whenever we need to compute the session key in a passive session, we use the private hash oracle H'_0 instead of H_0. Moreover, the Diffie-Hellman key K is not used as an input. In other words, the session key of a passive session is computed as $sk = H'_0(label, X, Y)$. The adversary can distinguish the experiment Exp_3 and the previous experiment Exp_2 if and only if the adversary sends a hash query $(label, X, Y, K)$ to the hash oracle H_0 in which X, Y is generated in a passive session and $K = CDH(X, Y)$. However, if the adversary can issue such a query, we can use the ability of the adversary to solve the CDH problem.

Given a CDH instance (U, V), we can embed the instance to all the passive session using the self-reducibility of the CDH problem. In order to do so, we choose four random numbers $a_0, b_0, a_1, b_1 \in Z^*_n$ for each passive session. In simulation the transcripts, we simply set $X = a_0U + b_0P$ and $Y = a_1V + b_1P$. All other transcripts are simulated as usual until the computation of the session key. The session key is computed as $sk = H'_0(label, X, Y)$. If an adversary can distinguish between this experiment and the previous one, then a query $(label, X, Y, K)$ must be issued to the hash oracle H_0. We can compute the Diffie-Hellman value of (U, V) by selecting a random record $(0, (label, X, Y, K), r)$ in $\wedge_{H_0}$ and computing $(K - a_0b_1U - a_1b_0V - b_0b_1P)/a_0a_1$.

Under the intractability assumption of the CDH problem, we have

$$\left|Adv_3(\mathcal{A}) - Adv_2(\mathcal{A})\right| \le negl(\kappa). \quad (6)$$

Experiment Exp_4. In this experiment, we begin to deal with the active sessions. For a $Send(U, (Y, \sigma_{GN}))$ query, if the signature σ_{GN} is a valid signature for this active session, we simply terminate the simulation and let the adversary

win. Since the user U is honest in this session, the message (X, T, c_1, s_m, s_a) is generated by the user U. Besides, we cancel the experiment in which the collision occurs in the output of the hash functions and the transcripts in Exp_2, so the signature σ_{GN} is valid if it is a signature for the random number r_{GN}. The adversary wins the game in this experiment if and only if a new signature is forged. The signature scheme used in our protocol is existential unforgeable against the chosen message attacks, so the advantage of the adversary $\mathcal{A}$ in forging a signature for a new random number is negligible. It is obvious that

$$\left|Adv_4(\mathcal{A}) - Adv_3(\mathcal{A})\right| \leq negl(\kappa). \quad (7)$$

Experiment Exp_5. In this experiment, we continue to deal with the active sessions. For a $Send(GN, (S_i, Y, T_{S_i}, Auth_{S_i}))$ query, if the sensor node S_i is uncorrupted, the timestamp T_{S_i} is within the transmission delay and $Auth_{S_i}$ is a valid authenticator; then we simply terminate the simulation and let the adversary win the attack game. Since the sensor node S_i is uncorrupted, the symmetric key $K_{(GN,S_i)}$ is unknown to the adversary. Moreover, the timestamp T_{S_i} makes the replay attack impossible. The adversary can only produce a valid authenticator $Auth_{S_i}$ by issuing a query $(K_{(GN,S_i)}, X, Y, T_{GN}, T_{S_i}, label)$ to the hash oracle H_5 or the adversary correctly guesses the output of the hash function H_5 without asking the corresponding message. $K_{(GN,S_i)}$ and $Auth_{S_i}$ are two random values chosen from $\{0,1\}^\kappa$; the success probability of the adversary is negligible. Consequently we have the following equation:

$$\left|Adv_5(\mathcal{A}) - Adv_4(\mathcal{A})\right| \leq negl(\kappa). \quad (8)$$

Experiment Exp_6. In this experiment, we deal with the active sessions once again. For a $Send(S_i, (label, X, T_{GN}, Auth_{GN}))$ query, if the timestamp T_{GN} is within the transmission delay and $Auth_{GN}$ is a valid authenticator, then we simply terminate the simulation and let the adversary win the attack game. Since the gateway node is not allowed to be corrupted, the symmetric key $K_{(GN,S_i)}$ is unknown to the adversary and the timestamp T_{GN} ensures the adversary cannot replay an old authenticator. The adversary can only produce a valid authenticator $Auth_{GN}$ by issuing a query $(K_{(GN,S_i)}, X, lable, T_{GN})$ to the hash oracle H_4 or the adversary correctly guesses the output of the hash function H_4 without asking the corresponding message. $K_{(GN,S_i)}$ and $Auth_{S_i}$ are two random values chosen from $\{0,1\}^\kappa$; the success probability of the adversary is negligible. Similarly with the previous experiment, we have

$$\left|Adv_6(\mathcal{A}) - Adv_5(\mathcal{A})\right| \leq negl(\kappa). \quad (9)$$

Experiment Exp_7. In this experiment, we change the simulation rule of *Send* queries for the last time. For a $Send(GN, (label, X, T, c_1, s_m, s_a))$ query, the gateway node will first check the validity of the credential proof. If the credential proof is valid and the message is forged by the adversary, we then terminate the simulation and the adversary is claimed successful. However, the success probability of the adversary in producing a fake proof is bounded by the presentation of an algebraic MAC. With a similar analysis with [23], we get the following result:

$$\left|Adv_7(\mathcal{A}) - Adv_6(\mathcal{A})\right| \leq negl(\kappa). \quad (10)$$

In the last experiment, we can see that all the session keys of passive sessions are chosen randomly from the domain and all the active sessions are terminated without accepting. The only way for the adversary to succeed is to steal the terminal of the user and recover the credential by guessing the password. The adversary has to verify the correctness of the recovered credential by executing the protocol. Consequently, we have

$$\left|Adv_7(\mathcal{A})\right| \leq C' \cdot Q_{send}^{s'}. \quad (11)$$

5. Performance Analysis

In this section, we evaluate the computation and communication costs and the security attributes of our protocol with other related protocols with user anonymity [16–19]. In terms of computation, let "T_M" denote the time of one modular exponentiation computation, "T_{PM}" denote the time cost of one point multiplication computation on elliptic curve, 'T_H" denote the time of one hash function computation, and "T_S" denote the time of one symmetric encryption/decryption operation. According to [24], $T_M \approx 1.169ms$, $T_{PM} \approx 0.508ms$, $T_H \approx 0.069ms$, and $T_S \approx 0.069ms$. Moreover, we only evaluate the computation cost of the authentication and key exchange phase because the registration phase is a one-time job. In terms of communication cost, we assume the length of the identity is 32 bits, the secure parameter κ is 160 bits, the length of the timestamp is 64 bits, an element of cyclic group of ECC can be represented with 320 bits, and an element of cyclic group of RSA can be presented with 1024 bits. We also instantiate the signature scheme using the famous ECDSA signature scheme [25]. The performance of communication and computation is summarized in Table 2. We can see from Table 2 that our protocol is inefficient in terms of computation. However, the communication performance of the compared protocols is more or less the same. The computation cost of our protocols mainly arises from the strong user anonymity; i.e., no one except the user knows his real identity in our protocol, while the gateway node knows the user's real identity in other protocols.

Table 3 summarizes security properties of the proposed protocol with related protocols. It can be seen from Table 3 that our protocol provides all the security features. Moreover, our protocol is the only one which provides strong user anonymity and formal security proof. Considering the computation cost, communication cost, and security attributes as a whole, our protocol outperforms to other protocols. Consequently, the proposed protocol is more suitable for security and privacy critic applications scenarios in WSNs.

Table 2: Comparisons of computation and communication costs.

Protocols	Wu et al.'s [16]	Jiang et al.'s [17]	Wang et al.'s [18]	Li et al.'s [19]	Our protocol
Computation time of user (ms)	$2T_{PM} + T_S + 11T_H \approx 1.04$	$T_{PM} + 8T_H \approx 1.18$	$2T_{PM} + 8T_H \approx 1.04$	$2T_{PM} + 8T_H \approx 1.05$	$4T_{PM} + 4T_H \approx 2.03$
Computation time of gateway (ms)	$2T_S + 11T_H \approx 1.04$	$T_{PM} + 12T_H \approx 1.19$	$2T_{PM} + T_S + 11T_H \approx 1.05$	$T_{PM} + 9T_H \approx 0.52$	$4T_{PM} + 5T_H \approx 2.03$
Computation time of sensor (ms)	$2T_{PM} + T_S + 4T_H \approx 1.05$	$5T_H \approx 0.04$	$2T_{PM} + T_S + 11T_H \approx 1.06$	$4T_H \approx 0.03$	$2T_{PM} + 3T_H \approx 1.02$
Rounds	4	4	4	4	4
Bandwidth	3168bits	2689bits	3968bits	2912bits	2976bits

Table 3: Comparisons of security features.

Protocols	Wu et al.'s [16]	Jiang et al.'s [17]	Wang et al.'s [18]	Li et al.'s [19]	Our protocol
The replay attack	secure	secure	secure	secure	secure
The privileged insider attack	secure	secure	secure	secure	secure
The GW-node impersonation attack	secure	secure	secure	secure	secure
The stolen verifier attack	secure	secure	secure	secure	secure
The off-line dictionary attack	secure	secure	secure	secure	secure
The compromised sensor node attack	secure	secure	secure	secure	secure
Mutual authentication	yes	yes	yes	yes	yes
Session key establishment	yes	yes	yes	yes	yes
Key privacy	yes	no	yes	no	yes
User anonymity	weak	weak	weak	weak	strong
Formal security proof	yes	yes	yes	yes	yes

6. Conclusions

In this paper, we propose an anonymous authentication and key exchange protocol for WSNs. The most attractive property of our protocol is its strong user anonymity such that no one except the user knows the real identity of himself. Besides this, our protocol also enjoys formal security proof in the random oracle model and efficient communication complexity. The only disadvantage is that it consumes more computation resources. In wireless communication networks, establishing a channel usually consumes more energy than computation does. As a result, the heavy computation cost is not a serious problem. Due to its high security and strong anonymity, our protocol is very suitable for security and privacy critic application scenarios in WSNs.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work is supported by the Funding of Science and Technology on Information Assurance Laboratory (no. KJ-17-001) and Key Scientific and Technological Project of Henan Province (no. 122102210126).

References

[1] Y. Liu, W. Guo, C. Fan, L. Chang, and C. Cheng, "A practical privacy-preserving data aggregation (3PDA) scheme for smart grid," *IEEE Transactions on Industrial Informatics*, pp. 1-1, 2018.

[2] D. He, N. Kumar, H. Wang, L. Wang, K. R. Choo, and A. Vinel, "A Provably-Secure Cross-Domain Handshake Scheme with Symptoms-Matching for Mobile Healthcare Social Network," *IEEE Transactions on Dependable and Secure Computing*, pp. 1-1, 2016.

[3] J. Shen, T. Zhou, D. He, Y. Zhang, X. Sun, and Y. Xiang, "Block design-based key agreement for group data sharing in cloud computing," *IEEE Transactions on Dependable and Secure Computing*, vol. PP, no. 99, 2017.

[4] J. Shen, J. Shen, X. Chen, X. Huang, and W. Susilo, "An efficient public auditing protocol with novel dynamic structure for cloud data," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 10, pp. 2402–2415, 2017.

[5] D. He, S. Zeadally, and L. Wu, "Certificateless public auditing scheme for cloud-assisted wireless body area networks," *IEEE Systems Journal*, vol. 12, no. 1, pp. 64–73, 2018.

[6] Q. Jiang, Z. Chen, B. Li et al., "Security analysis and improvement of bio-hashing based three-factor authentication scheme for telecare medical information systems," *Journal of Ambient Intelligence and Humanized Computing*, 2017.

[7] Q. Jiang, J. Ma, C. Yang, X. Ma, J. Shen, and S. A. Chaudhry, "Efficient end-to-end authentication protocol for wearable health monitoring systems," *Computers and Electrical Engineering*, 2017.

[8] M. L. Das, "Two-factor user authentication in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1086–1090, 2009.

[9] M. K. Khan and K. Alghathbar, "Cryptanalysis and security improvements of 'two-factor user authentication in wireless sensor networks'," *Sensors*, vol. 10, no. 3, pp. 2450–2459, 2010.

[10] H.-L. Yeh, T.-H. Chen, P.-C. Liu, T.-H. Kim, and H.-W. Wei, "A secured authentication protocol for wireless sensor networks using Elliptic Curves Cryptography," *Sensors*, vol. 11, no. 5, pp. 4767–4779, 2011.

[11] K. Xue, C. Ma, P. Hong, and R. Ding, "A temporal-credential-based mutual authentication and key agreement scheme for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 316–323, 2013.

[12] D. He, N. Kumar, H. Shen, and J.-H. Lee, "One-to-many authentication for access control in mobile pay-TV systems," *Science China Information Sciences*, vol. 59, no. 5, pp. 1–14, 2016.

[13] J.-J. Yuan, "An enhanced two-factor user authentication in wireless sensor networks," *Telecommunication Systems*, vol. 55, no. 1, pp. 105–113, 2014.

[14] D. Wang and P. Wang, "Understanding security failures of two-factor authentication schemes for real-time applications in hierarchical wireless sensor networks," *Ad Hoc Networks*, vol. 20, pp. 1–15, 2014.

[15] J. Shen, S. Chang, J. Shen, Q. Liu, and X. Sun, "A lightweight multi-layer authentication protocol for wireless body area networks," *Future Generation Computer Systems*, vol. 78, no. 3, pp. 956–963, 2018.

[16] F. Wu, L. Xu, S. Kumari, and X. Li, "A new and secure authentication scheme for wireless sensor networks with formal proof," *Peer-to-Peer Networking and Applications*, vol. 10, no. 1, pp. 16–30, 2017.

[17] Q. Jiang, S. Zeadally, J. Ma, and D. He, "Lightweight three-factor authentication and key agreement protocol for internet-integrated wireless sensor networks," *IEEE Access*, vol. 5, pp. 3376–3392, 2017.

[18] C. Wang, G. Xu, and J. Sun, "An enhanced three-factor user authentication scheme using elliptic curve cryptosystem for wireless sensor networks," *Sensors*, vol. 17, no. 12, article no. 2946, 2017.

[19] X. Li, J. Niu, S. Kumari, F. Wu, A. K. Sangaiah, and K. R. Choo, "A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments," *Journal of Network and Computer Applications*, vol. 103, pp. 194–204, 2018.

[20] C. Wang, D. Wang, G. Xu, and Y. Guo, "A lightweight password-based authentication protocol using smart card," *International Journal of Communication Systems*, vol. 30, no. 16, pp. 1–11, 2017.

[21] D. Wang, H. Cheng, P. Wang et al., "Zipfs law in passwords," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

[22] F. Wei, P. Vijayakumar, J. Shen, R. Zhang, and L. Li, "A provably secure password-based anonymous authentication scheme for wireless body area networks," *Computers and Electrical Engineering*, 2017.

[23] Z. Zhang, K. Yang, X. Hu, and Y. Wang, "Practical anonymous password authentication and TLS with anonymous client authentication," in *Proceedings of the 23rd ACM Conference on Computer and Communications Security, CCS 2016*, pp. 1179–1191, October 2016.

[24] 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*, 2016.

[25] C. Schnorr, "Efficient signature generation by smart cards," *Journal of cryptology*, vol. 4, no. 3, pp. 161–174, 1991.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 9436120, 14 pages
https://doi.org/10.1155/2018/9436120

Research Article

LIP-PA: A Logistics Information Privacy Protection Scheme with Position and Attribute-Based Access Control on Mobile Devices

Qi Gao, Junwei Zhang, Jianfeng Ma, Chao Yang, Jingjing Guo, and Yinbin Miao

School of Cyber Engineering, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Junwei Zhang; jwzhang@xidian.edu.cn

Received 13 April 2018; Accepted 27 June 2018; Published 15 July 2018

Academic Editor: Ding Wang

With the fast development of Logistics Internet of Things and smart devices, the security of express information processed by mobile devices in Logistics Internet of Things has attracted much attention. However, the existing secure express schemes only focus on privacy protection of personal information but do not consider the security of the logistics information against couriers with malicious mobile devices. For example, a privacy-preserving delivery path should be required in order to prevent the privacy leakage in the express delivery procedure. Therefore, besides the security of personal information, the privacy protection of logistics information and authentication of mobile devices used in express company are important to security in Logistics Internet of Things. In this paper, we propose a secure logistics information scheme LIP-PA to provide privacy protection of both personal information and logistics information. First, we define the basic requirements of Logistics Internet of Things. Then, using attribute-based encryption and position-based key exchange, we propose a logistics information privacy protection scheme with position and attribute-based access control for mobile devices. The analysis results show that our scheme satisfies the defined requirements. Finally, the performance of our scheme is evaluated and the experiment results show that our scheme is efficient and feasible for mobile devices in real parcel delivery scenario.

1. Introduction

With the rapid development of Internet of Things technology [1] and continuous optimization of the logistics operation process, the Logistics Internet of Things (LIoT) has become an indispensable pattern for modern logistics industry. Based on many network communication technologies, such as RFID [2], QR code [3], NFC [4], and D2D [5] technology, LIoT has adopted lots of mobile devices to deal with the express business.

At the same time, the fast development of modern logistics industry has also brought a lot of security issues, especially the security of mobile devices in LIoT [6]. Without security protection, the customer's private information is expressly visible on mobile devices, so that any adversary can see the private information. Besides, the adversary can easily track the parcel according to the logistics information in mobile devices and even analyze the customer's information such as personal hobbies, family members, and economic conditions [7, 8]. In addition, express delivery between different express stations can lead to data leakage and the malicious courier can analyze the flow of express station [9]. The privacy protection of express information has attracted widespread attention.

However, traditional privacy protection schemes of express information all focus on customer's personal information [10, 11], and they are unable to achieve a secure management of internal logistics staff. It is obvious that the couriers, who actually move parcels between sender and receiver, can obtain lots of express information including personal information and logistics information. Therefore, how to guarantee the express information privacy and securely manage the large-scale couriers based on mobile devices in LIoT is a focus.

As far as we know, the most popular authentication methods for couriers on mobile devices can be classified as public key authentication [12–14], password authentication [15, 16], and multifactors authentication [17–19]. However, these methods are not suitable for management of large-scale couriers on mobile devices. Group-based key exchange can

realize efficient group-based management [20, 21], but group-based methods could not provide fine-grained access control on large-scale mobile devices. As a result, attribute-based encryption with fine-grained access control can be applied to the security solutions in LIoT.

Related Work. For privacy protection in LIoT, Wei et al. [22] proposed a K-anonymous model to take the anonymous process of logistics information. This method randomly breaks the relationship between attribute values in the record to anonymous data. However, the order still retains the receiver's name and phone number, and K-anonymity will cause some loss of information. Zhang et al. [23] proposed a logistics information system privacy protection system, which can solve the contradiction between privacy protection and logistics business process by the segment encryption design. However, the disadvantage of this solution is that the two-dimensional code needs to be constantly updated at each logistics station; also the processes of encryption and decryption are repeated. Qi et al. [24] proposed a new express management system based on encrypted QR code. The real-time logistics information of goods is automatically updated through GPRS or Wi-Fi. The APP provides an optimal delivery route for couriers by employing the improved genetic algorithm. Obviously, the above works lack access control and authentication for internal logistics staff. They cannot ensure the privacy protection of logistics information.

Li et al. [25] designed a privacy-preserving express delivery system, i.e., PriExpress. This system introduces the ciphertext-policy attribute-based encryption (CP-ABE) method into the privacy protection in logistics information system (LIS). With CP-ABE [26], the parcel sender specifies an attribute-based access policy for enforcing fine-grained access control to his delivery order which contains sensitive personal information. However, the PriExpress does not separate personal information from logistics information. Customers in this scheme need abundant computation capability and remain online when they deal with the express delivery procedure.

In summary, all above schemes lack a position and attribute-based access control for both logistics information and personal information.

Privacy Protection in LIoT. Different from IoT, the security requirements of LIoT have the following characteristics.

First, privacy-preserving of logistics information should be guaranteed in LIoT. For complete logistics information, it is necessary to keep confidentiality for untrusted express staff. In traditional privacy protection schemes, logistics information is not separated from personal information. Some malicious couriers may sell the logistics information to criminals who can analyze the flow of express station according to the delivery path. Thus, privacy-preserving of logistics information is one of the required properties in LIoT.

Second, couriers who actually deliver parcels are very important to the guarantee of the security of express delivery process. But there are so large-scale couriers that it is difficult to manage them. On one hand, for some malicious couriers, they may sell privacy information to criminals. On the other hand, once the hackers get the courier's master secret, they can impersonate a valid courier to obtain the express information illegally. Therefore, it is necessary to guarantee the attribute-based access control for privacy information and ensure that the courier with specific attributes can obtain privacy information.

Due to the high turnover of couriers, it is significant to make sure that only the couriers that work online at delivery station can obtain the order information. So position-based access control is also one of the required properties of LIoT. Besides, there may be some couriers that are curious about the information which they cannot possess, so they maybe collude together and share their own attributes to obtain much more information. Therefore, it is necessary to implement the authentication of couriers. Specifically, the security requirements on couriers should guarantee anticollusion attack to attribute-based access control and anticollusion attack to position-based access control.

Third, customers always are remote and offline. Therefore, it is hard to achieve authentication of customers online [27]. A dishonest sender may deny that he (or she) has sent some harmful parcel to someone. Besides, a dishonest receiver may personate a legal receiver to take away the parcel which does not belong to himself (or herself). Therefore, it is necessary to achieve the verifiability of receiver and verifiability of parcel [28]. Last but not least, it is necessary to keep customer's unlinkability for administrator and others.

Fourth, in the real delivery scene, customers always have irregular operations such as leaving wrong addresses and phone numbers, which leads to the fact that it is difficult to manage or recover parcel. Therefore, the security requirements of LIoT should guarantee the undeniability of sender and receiver. Specifically, a sender cannot deny a parcel sent by himself (or herself), while a receiver cannot deny a parcel received by himself (or herself).

Our Contributions. Although the most popular authentication methods on mobile devices are not suitable for management of large-scale couriers [29, 30], attribute-based encryption can achieve fine-grained access control, so we need to use the ABE technology to realize position and attribute-based access control of couriers based on mobile devices in LIoT. However, this study is very challenging with the following reasons. First, it is hard to realize both attribute-based encryption and location-based access control. Second, how to securely verify the validity of a courier's claimed position is also a problem [31].

This paper proposes a logistics information privacy protection scheme with position and attribute-based access control on mobile devices. First, in order to realize fine-grained access control of encrypted logistics information, we adopt ciphertext-policy attribute-based encryption (CP-ABE) [26], which encrypts segmented logistics information in different access policies. Different couriers can only decrypt different segments of the express order in accordance with their respective attributes. Second, we apply position-based key exchange [32], which uses the courier's physical position information as credential, to realize position-based access control on couriers. Third, we utilize public key encryption

to achieve the confidentiality of personal information. At the same time, we use digital signature to ensure the verifiability of parcel and the undeniability of customers [33].

Our contributions in this paper are fourfold.

(1) We classify the required properties of our scheme including attribute-based access control, position-based access control, privacy-preserving of logistics information, confidentiality of personal information, verifiability of receiver, verifiability of parcel, anticollusion attack to attribute-based access control, anticollusion attack to position-based access control, undeniability, and unlinkability.

(2) We propose a logistics information privacy protection scheme LIP-PA. In the scheme, logistics information is divided into segments and, respectively, encrypted; administrator can prebuild access tree which contains the position attribute.

(3) We theoretically analyze the security of LIP-PA. We show that LIP-PA satisfies the above required properties.

(4) We report experimental evaluations of our scheme. Our results show that LIP-PA is efficient and feasible for mobile devices in real parcel delivery scenario.

2. Preliminaries

In this paper, some basic cryptographic algorithms are necessary. First, we will use public key encryption to protect the personal information [34] in our scheme. Second, we also need digital signature [35, 36] to realize the undeniability of customers. Third, hash functions [37] are used in generation of order information. Besides the above algorithms, ciphertext-policy attribute-based encryption and position-based key exchange are core algorithms in our scheme.

2.1. Ciphertext-Policy Attribute-Based Encryption. A ciphertext-policy attribute-based encryption scheme [26] consists of four fundamental algorithms: Setup, Encrypt, KeyGen, and Decrypt.

Setup: The setup algorithm takes no input other than the implicit security parameter. It outputs the public parameters PK and a master key MK.

Encrypt(PK, M, T): The encryption algorithm takes as input the public parameters PK, a message M, and an access tree T. The algorithm will encrypt M and produce a ciphertext CT such that only a user with a set of attributes that satisfies the access tree T will be able to decrypt CT.

Key Generation(MK, S): The key generation algorithm takes as input the master key MK and a set of attributes S. It outputs the private key SK, which is used by users to decrypt ciphertext.

Decrypt(PK, CT, SK): The decryption algorithm takes as input the public parameters PK, a ciphertext CT, and a private key SK. If the set S satisfies the access tree T then the algorithm will return a message M.

2.2. Position-Based Key Exchange. Based on the bounded storage model (BSM) and BSM pseudorandom generators (PRG), Chandran et al. construct the provable secure position-based key exchange (PBKE) protocol against colluding adversaries [32]. BSM assumes that any party including adversary can only store a part of information with high min-entropy. BSM PRG: $\{0,1\}^n \times \{0,1\}^d \longrightarrow \{0,1\}^m$ is an ε-secure BSM PRG for storage rate β and min-entropy rate α if and only if, for every αn-source X on $\{0,1\}^n$ and for every function A: $\{0,1\}^n \longrightarrow \{0,1\}^{\beta n}$, the random variable $(PRG(X,K), A(X), K)$ is ε-close to $(U_m, A(X), K)$, where $K \xleftarrow{R} \{0,1\}^d$. Different from traditional key exchange, PBKE applies the user's physical position as the unique credential to negotiate a shared key K between verifiers and a prover at a legal position P. At the end of PBKE, the shared key K and a random number are indistinguishable from the view of the colluding adversaries.

However, the previous position-based key exchange protocol is not suitable for realizing the position-based access control in ABE. Thus, an improved position-based key exchange protocol is proposed as shown in Algorithm 4.

3. Problem Formulations

3.1. System Model. As shown in Figure 1, the system model consists of following entities: customers including a sender and a receiver, couriers, and administrator with attribute authority and landmarks.

When a logistic transit process begins, as shown in Figure 1, a sender first generates an order and submits the order information to administrator. The order information consists of address information, i.e., customers' address information, and customers' personal information including the names and telephone numbers. If the administrator accepts this order, the administrator will generate a logistics information which is a delivery plan including some independent delivery steps for couriers. During the parcel delivery, every courier in one delivery step only distributes the express from one station to another station. The parcel finally arrives at the last delivery station and the receiver takes away the parcel after authenticated. More specifically, we have the following.

Customers. Customers can be divided into a sender and a receiver of parcel. They aim to absolutely protect personal information against administrator and couriers and obtain a privacy-preserving protection service for logistics information from administrator; i.e., any courier managed by administrator can only know a part of logistics information but not all the logistics information.

Administrator. Administrator, as a general management institution of logistics company, provides trusted express service for customers and completes the privacy protection of logistics information against couriers. Specifically, the administrator first employs an attribute authority (AA) to realize attribute-based management for couriers. Then, the administrator configures all the access trees for all couriers involved by the target order according to the delivery plan and encrypts the segmented logistics information with different attribute policies. If a courier delivers a parcel in some one delivery step, the courier can only decrypt the required

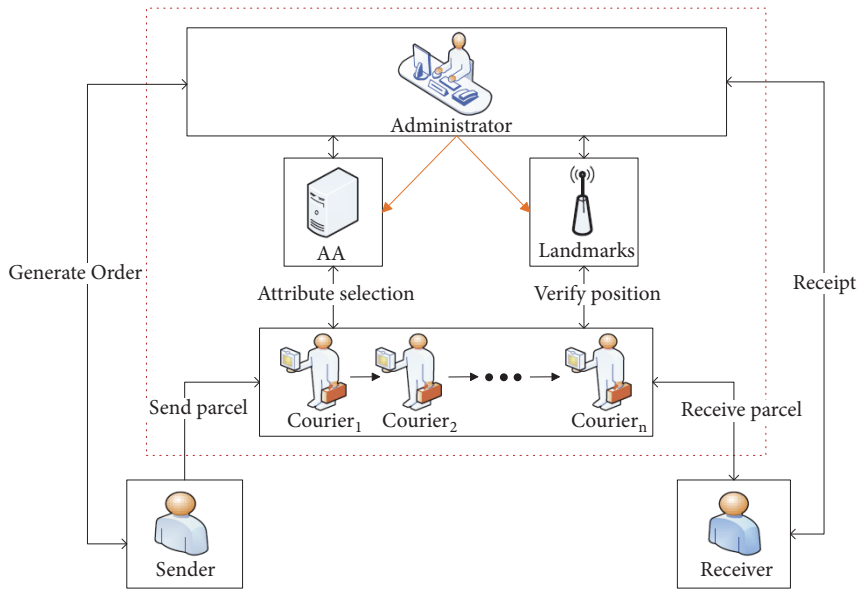


Figure 1: The system model.

information segment (i.e., the next station information of the current delivery step). At the same time, the administrator employs some landmarks to verify the location of the courier in order to guarantee that the courier from a legal station in a valid time slot can get the required information segment to perform the delivery task.

Couriers. Couriers are employed by express company. They are the entities that actually deliver parcels between a sender and a receiver. A courier is only responsible for delivering the parcels from one station to another station in a delivery step. The courier just needs to obtain the required information segment (e.g., the next station) with position and attribute-based access control. In this way, a parcel delivery process consists of multiple delivery steps performed by multiple couriers.

3.2. Threat Model

Administrator. Administrator, as a manager of an express company, is honest but curious. On one hand, the administrator is responsible for privacy protection on customers' address information and logistics information against couriers. For such a service, the administrator is in charge of securely encrypting the logistics information and building suitable access policy to couriers. AA and landmarks are employed by administrator to guarantee position and attribute-based access control. On the other hand, in order to obtain more potential benefits, the administrator is also very interested in customers' personal information, such as name and telephone number.

Customers. A customer of the parcel (i.e., sender or receiver) may have the following dishonest behaviors. First, a dishonest sender may deny that he (or she) has sent some harmful parcels to someone. Second, a dishonest receiver may personate a legal receiver to take away the parcel which does not belong to himself (or herself).

Couriers. Couriers, as the entities that actually deliver parcels, are very important for the guarantee of the security of the express delivery process. However, in order to obtain illegal individual benefit, some couriers may have four dishonest behaviors: First, couriers are curious about the customers' address and personal information; they may sell them to obtain economic benefit. Second, some colluding couriers may attempt to steal the logistics information even if they are not located at the valid delivery station at work time. Third, some couriers with different attributes may collude with others in order to decrypt the extra information which they could not know originally. Fourth, couriers may modify the order information in order to disturb the express delivery process, such as changing the valid receiver information into an illegal receiver.

3.3. Design Goals. According to the requirements and the adversary model, the proposed scheme should satisfy the following properties:

(1) Attribute-based access control (ABAC): Our scheme should achieve a fine-grained access control of encrypted logistics information based on the attributes of couriers. The required part of logistics information could only be decrypted by a valid courier whose attributes satisfy the access policy. Apart from this, there is no way to obtain the other parts of logistic information for valid couriers.

(2) Position-based access control (PBAC): Our scheme should ensure that a courier that is going to obtain the

Table 1: Comparison with related work.

	LIPPS	NEMS	PriExpress	Our scheme
ABAC	×	×	√	√
PBAC	×	×	×	√
PPLI	√	×	√	√
CPI	×	√	×	√
VR	√	√	√	√
VP	×	×	×	√
ACA-A	-	-	√	√
ACA-P	-	-	-	√
UD	×	×	×	√
UL	×	×	√	√

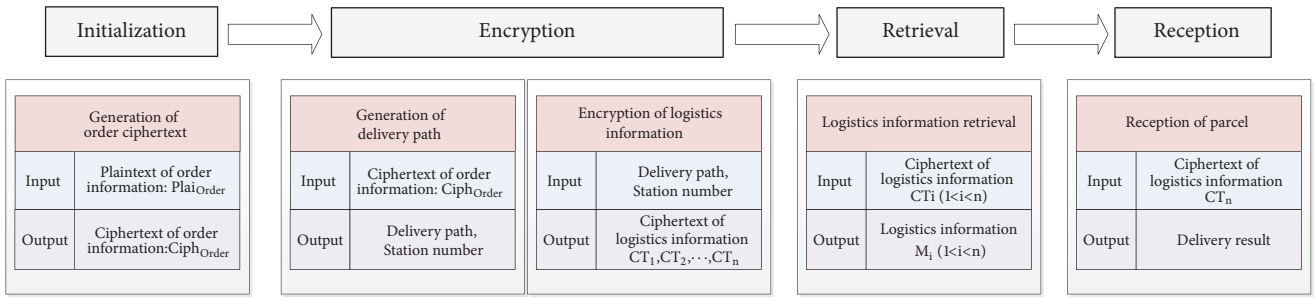


Figure 2: The LIP-PA framework.

required part of logistics information must be at a valid delivery station at the expected work time.

(3) Privacy-preserving of logistics information (PPLI): First and foremost, our scheme should guarantee confidentiality of logistics information. For logistics company, the whole delivery path should only be known by administrator, so a valid courier can only decrypt the partial logistics information which is necessary for the delivery process. The complete logistics information keeps privacy protection to all couriers.

(4) Confidentiality of personal information (CPI): Our scheme should ensure the confidentiality of personal information even to administrator. Customers' personal information contains sender's and receiver's names, phone numbers, etc. It is only visible among the sender and the receiver.

(5) Verifiability of receiver (VR): Our scheme should provide the verifiability of receiver. Only when the receiver of parcel is the expected one according to the order information, can receiver take away the parcel from the final courier.

(6) Verifiability of parcel (VP): A receiver in our scheme should ensure that the order information of the parcel is the correct and unforged during the express delivery process, and then he (or she) will receive this parcel.

(7) Anticollusion attack to attribute-based access control (ACA-A): Our scheme should ensure that the colluding couriers with different attributes cannot obtain the additional logistics information under the colluding attribute sets.

(8) Anticollusion attack to position-based access control (ACA-P). Our scheme should achieve that the colluding couriers that are not located at the valid station cannot obtain the logistics information according to the position-based access control policy.

(9) Undeniability (UD): A sender cannot deny a parcel sent by himself (or herself), while a receiver cannot deny a parcel received by himself (or herself).

(10) Unlinkability (UL): Although the same sender sends lots of parcels to a receiver, the administrator and others cannot distinguish whether the encrypted order information in the many delivery processes originates from the same sender.

3.4. Comparison with Related Work. In this section, we compare our scheme with the related schemes including LIPPS [23], NEMS [24], and PriExpress [25]. Table 1 shows the comparison results, where "√" means satisfied, "×" means dissatisfied, and "–" means uninvolved.

From the Table 1, we can see that both LIPPS and NEMS cannot provide ABAC. It is obvious that all the related schemes including LIPPS, NEMS, and PriExpress cannot guarantee PBAC, VP, and UD. Meanwhile PriExpress cannot provide totally CPI. Our scheme can satisfy all the properties in Table 1.

4. The Proposed Scheme

4.1. Overview of Scheme LIP-PA. The scheme LIP-PA consists of four phase: initialization, encryption, retrieval, and reception as shown in Figure 2.

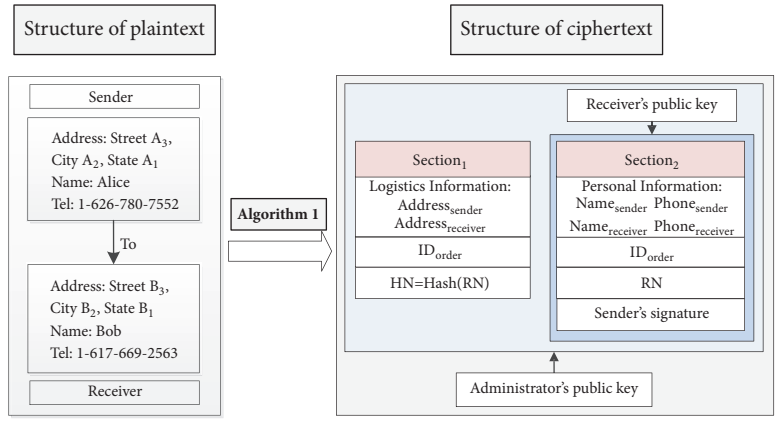


Figure 3: Initialization of order information.

Input: plaintext of order information $Plai_{Order}$
Output: ciphertext of order information $Ciph_{Order}$
1: generate random number ID_{Order}, RN
2: $HN \longleftarrow Hash(RN)$
3: $signature_{Sen} \longleftarrow E(PRI_{Sen}, [ID_{Order}||Plai_{Order}])$
4: $Section_1 \longleftarrow Add_{Sen}||Add_{Rec}||ID_{Order}||HN$
5: $Section_2 \longleftarrow Nam_{Sen}||Nam_{Rec}||Pho_{Sen}||Pho_{Rec}||ID_{Order}||RN||signature_{Sen}$
6: $Ciph_{Section_2} \longleftarrow E(PUB_{Rec}, [Section_2])$
7: $Ciph_{Order} \longleftarrow E(PUB_{Adm}, [Ciph_{Section_2}||Section_1])$
8: **return** $Ciph_{Order}$

Algorithm 1: Generation of order ciphertext.

In the initialization phase, a sender will generate an encrypted order and launch an order request to administrator. The encrypted order information consists of customers' address information and personal information like names, telephone numbers, etc.

Upon receiving the order request from sender, the administrator constructs logistics information based on the planned delivery path in the encryption phase. Then, the administrator formulates the position and attribute-based access control and encrypts different segments of logistics information with different access control policies.

In the retrieval phase, a courier located at a valid station at work time can run a position-based key exchange protocol to obtain a secret key about position attribute. Based on the position-based key, the courier who satisfies the desired attribute policy can retrieve required logistics information in order to transmit the parcel from one station to next station.

When the parcel arrives at last station, the courier will transmit the parcel to target receiver in the reception phase. In this phase, the courier should verify the authenticity of receiver. The receiver would also verify the correctness of order information on parcel. If both authentication of receiver and validity of parcel are verified, the parcel will be delivered to receiver successfully.

4.2. Initialization. We assume that sender, receiver, administrator, and courier must be registered and have a certified public/private key pair, respectively, before initialization.

In our scheme, order information can be divided into two types: logistics information and personal information. Logistics information including sender's address (Add_{Sen}) and receiver's address (Add_{Rec}) should be encrypted using public key of administrator to provide the property of secrecy against adversarial couriers. Personal information, such as sender's name (Nam_{Sen}), receiver's name (Nam_{Rec}), sender's phone number (Pho_{Sen}), and receiver's phone number (Pho_{Rec}), should be always secret during the express delivery process considering the confidentiality. Figure 3 illustrates the structure of encrypted order information sent from sender to administrator. Algorithm 1 shows the detailed generation process of order ciphertext, where $E()$ is the public key encryption algorithm, PRI_{user} means the user's private key, and PUB_{user} means the user's public key.

As we can see, the ciphertext mainly contains two sections: $Section_1$; $Section_2$. $Section_1$ contains logistics information and $Section_2$ contains personal information. Note that, in line 2 of Algorithm 1, a random number ID_{Order} is unique identity of parcel. In addition, the sender generates a random number RN in $Section_2$. At the same time, the

Input: ciphertext of order information $Ciph_{Order}$
Output: delivery path $Optimalpath$, station number $Num_{station}$
1: $D(PRI_{Adm}, [Ciph_{Order}]) \longrightarrow Ciph_{Section_2}||Section_1$
2: **if** $A_2 = B_2$ **then**
3: $Num_{station} = 3$
4: $A_3A_2B_3 \longrightarrow Optimalpath$
5: **else**
6: **if** $A_1 = B_1$ **then**
7: $Num_{station} = 5$
8: $A_3A_2A_1B_2B_3 \longrightarrow Optimalpath$
9: **else**
10: $Num_{station} = 6$
11: $A_3A_2A_1B_1B_2B_3 \longrightarrow Optimalpath$
12: **end if**
13: **end if**
14: **return** $Optimalpath$, $Num_{station}$

Algorithm 2: Generation of delivery path.

Input: $Optimalpath$, $Num_{station}$
Output: ciphertext of logistics information $(CT_1, CT_2, \ldots, CT_n)$
1: $n = Num_{station}$
2: **for** $i = 1; i \leq n; i++$ **do**
3: $Position_i \longleftarrow$ *delivery station on* $Optimalpath[i]$
4: picks $K_{i6} \overset{R}{\leftarrow} \{0,1\}^m$
5: $\mathcal{S}_{iSpatial} \longleftarrow \{Position_i, T_i\}$
6: $attr(\mathcal{S}_{iSpatial}) \longleftarrow K_{i6}$
7: **end for**
8: **for** $i = 1; i \leq n-1; i++$ **do**
9: $M_i \longleftarrow Optimalpath[i]$
10: **end for**
11: $M_n \longleftarrow Ciph_{Section_2}||HN$
12: Setup $\longrightarrow (PK, MK)$
13: Key Generation$(PK, MK, \mathcal{S}) \longrightarrow (SK)$
14: **for** $j = 1; j \leq n; j++$ **do**
15: Encrypt $(PK, M_j, T_j) \longrightarrow (CT_j)$
16: **end for**
17: **return** $(CT_1, CT_2, \ldots, CT_n)$

Algorithm 3: Encryption of logistics information.

sender computes $HN = Hash(RN)$ in $Section_1$. In line 5 of Algorithm 1, the sender's digital signature is added to $Section_2$ for the verifiability of parcel. Finally, Algorithm 1 outputs the ciphertext of order information, which will be submitted to administrator.

4.3. Encryption. The encryption phase contains two steps which administrator needs to complete. First, the administrator runs the delivery path generation algorithm (i.e., Algorithm 2) and generates the logistics information according to sender's and receiver's addresses in $Ciph_{Order}$. Second, based on the delivery path, the administrator formulates the position and attribute policies $(T_1, T_2, \ldots, T_n)$ and encrypts segments of logistics information using CP-ABE with related policies. Algorithm 3 shows the detailed encryption process of logistics information.

In line 1 of Algorithm 2, the administrator conducts $D(PRI_{Adm}, [Ciph_{Order}])$, where $D()$ is a public key decryption algorithm, and obtains Add_{Sen}, Add_{Rec}, ID_{Order}, HN, and $Ciph_{Section_2}$. Because address can be divided into three sections, street, city, and state, all address structures can be combined to form multiple tree structures. According to sender's and receiver's addresses, the administrator will generate an optimal path, as shown in Figure 4.

Obviously, the sender's address is different from receiver's address. In other words, the street $A_3 \neq B_3$. There are three types of paths according to addresses of sender and receiver. The parcel will be distributed according to the optimal path.

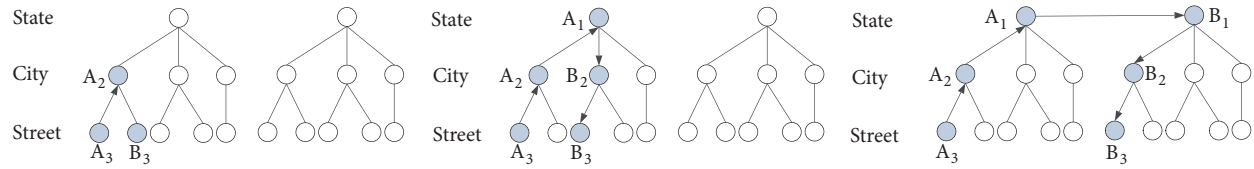


Figure 4: Delivery path.

So the number of delivery stations can be divided into three types:

(1) When $A_2 = B_2$, as shown in Figure 4(a), the optimal path is $A_3A_2B_3$, so there are three delivery stations on the optimal path.

(2) When $A_1 = B_1$ *and* $A_2 \neq B_2$, as shown in Figure 4(b), the optimal path is $A_3A_2A_1B_2B_3$, so there are five delivery stations on the optimal path.

(3) When $A_1 \neq B_1$, as shown in Figure 4(c), the optimal path is $A_3A_2A_1B_1B_2B_3$, so there are six delivery stations on the optimal path.

It is worth noting that, in real parcel delivery scenario, the numbers of delivery stations may be changed, such as wrong delivery or route change. The alternative solutions are as follows.

When the administrator has encrypted multiple paths and made the ciphertext into a QR code, the courier can decrypt multiple optional station addresses. He (or she) needs to choose a suitable station according to actual situation. When the existing optional paths are all unavailable, the administrator will encrypt a new path and send the new ciphertext to courier. If the parcel is wrongly transmitted, the solution is different. For example, the original path is $A_3A_2B_3$, as shown in Figure 4(a); the real path becomes $A_3A_2CB_3$. It means the courier at A_2 transmits the parcel to the wrong station C; at the same time, the courier at C cannot decrypt the logistics information. So he (or she) will return the parcel to station A_2. The courier at A_2 needs to tell the event to administrator. Then administrator will adjust the access policy of logistics information.

In lines 2-7 of Algorithm 3, the administrator predetermines the position attribute according to delivery stations on delivery path for couriers, where $attr(\mathcal{S})$ means the attribute $\mathcal{S}$'s value. Our scheme guarantees that the courier needs to arrive at the correct delivery position, so that he (or she) can obtain a secret key about the policy of position-based access control.

In lines 8-11 of Algorithm 3, the administrator divides the path into segments according to delivery station nodes and obtains segments $M_1, M_2, \ldots, M_n$. In lines 12-15 of Algorithm 3, the administrator, respectively, encrypts the segments $M_1, M_2, \ldots, M_n$ under CP-ABE. This method guarantees a fine-grained access control of encrypted logistics information. Besides, the position attribute $\mathcal{S}_{Spatial}$ is added to leaf nodes of access tree. So our scheme achieves a position-based access control.

After completing Algorithm 3, the administrator makes the ciphertext $CT_1, CT_2, \ldots, CT_n$ combined with ID_{Order}. At the same time, he (or she) sends encrypted PK and SK to couriers using his (or her) public key. When the sender drops off the parcel at local delivery station to courier, the courier can inquire about the ciphertext according to ID_{Order}. Then he (or she) makes the ciphertext into a QR code, which will be pasted on the parcel.

4.4. Retrieval. When the parcel arrives at the delivery station, the courier at this station scans QR code to get ciphertext CT_i and decrypts logistics information M_i; then he (or she) transmits the parcel to next station.

Before logistics information retrieval, the courier can decrypt and obtain PK and SK sent by administrator. In the retrieval phase, the courier who is located at a valid station at work time needs to run an improved position-based key exchange protocol to obtain the secret key about policy of position-based access control. In Algorithm 4, the courier performs the improved position-based key exchange (I-PBKE) with landmarks to obtain $\mathcal{S}_{Spatial}$, where $F(X_i, K_j)$ means the BSM PRG function.

In the previous position-based key exchange protocol (PBKE) [32], it assumes that landmarks must store $\{X_i\}$ in order to compute the expected response from courier. However, this position-based key exchange protocol is not suitable for realizing the position-based access control with ABE in our scheme. There are two reasons. First, X_i is a long string which is drawn from the landmark's reverse block entropy source. Thus the landmark's storage capacity needs to be large enough. What is more, the landmark generates X_i randomly along with the protocol execution. It means the final exchange key K_6 is determined by all landmarks after protocol execution, so that the administrator cannot prebuild access tree for couriers.

The improved position-based key exchange protocol is shown in Algorithm 4. The landmarks predetermine the keys $K_1, K_2, K_3, K_4, K_5, K_6$ that are to be used at every iteration of the application of the PRG. Now, the expected exchange key K_6 is known before protocol execution to all landmarks.

Obviously, there are two advantages of the improved position-based key exchange (I-PBKE) compared with the previous PBKE [32]:

(1) The expected exchange key K_6 is known by all landmarks before protocol execution. In other words, the position attribute's value is already determined, so the administrator can prebuild the access tree T which contains the position

1: Landmarks$\{L_1, L_2, L_3, L_4\}$ executes:
2: pick keys K_1,K_2,K_3,K_4,$K_5 \overset{R}{\leftarrow} \{0,1\}^m$
3: broadcast K_1,K_2,K_3,K_4,K_5 over their private channel.
4: At time $T - T_1$, L_1 picks large string X_4, computes $K_5' = F(X_4, K_4) \oplus K_5$, and broadcasts (X_4, K_1, K_5')
5: At time $T - T_2$, L_2 picks large string X_1,X_5, computes $K_2' = F(X_1, K_1) \oplus K_2$, $K_6' = F(X_5, K_5) \oplus K_6$, and broadcasts (X_1, X_5, K_2', K_6')
6: At time $T - T_3$, L_3 picks large string X_2, computes $K_3' = F(X_2, K_2) \oplus K_3$, and broadcasts (X_2, K_3')
7: At time $T - T_4$, L_4 picks large string X_3, computes $K_4' = F(X_3, K_3) \oplus K_4$, and broadcasts (X_3, K_4')
8: Courier executes:
9: At time T, the courier receives all of the strings
10: **for** $i = 1$; $i < 6$; $i++$ **do**
11: compute $K_{i+1} = F(X_i, K_i) \oplus K_{i+1}'$
12: end for
13: **return** exchange key K_6

Algorithm 4: Improved position-based key exchange.

Input: ciphertext CT_i
Output: logistics information M_i
1: Courier executes:
2: $K_{i6} \longleftarrow$ I-PBKE
3: $attr(\mathcal{S}_{Spatial}) \longleftarrow K_{i6}$
4: $CT_i = (T_i, \widetilde{C}_i, C_i)$
5: **if** $\mathcal{S}$ satisfies T_i **then**
6: $r \longleftarrow$ node of access tree T_i
7: $A \longleftarrow DecryptNode(CT_i, SK, r)$
8: $M_i \longleftarrow \widetilde{C}_i / (e(C_i, SK)/A)$
9: end if
10: **return** logistics information M_i

Algorithm 5: Logistics information retrieval.

attribute, instead of waiting for the courier's response when the parcel arrives at delivery station.

(2) The landmarks need not store long strings $\{X_i\}$.

In line 2 of Algorithm 5, the courier obtains the secret key K_6 according to Algorithm 4 (I-PBKE); thus the courier possesses suitable attribute set $\mathcal{S}$. Since the courier has got PK and SK sent by administrator and he (or she) satisfies the desired access tree T, the courier can retrieve required logistics information M_i, as shown in lines 5-9 of Algorithm 5. After decrypting the logistics information M_i, the courier will transmit parcel to next station. Multiple couriers collaborate to complete the express delivery process.

4.5. Reception. When the parcel arrives at last station, the final courier gets ciphertext CT_n; then he (or she) decrypts logistics information and transmits the parcel to target receiver.

Note that, in line 2 of Algorithm 6, the courier decrypts M_n according to Algorithm 5 (LIR). Specifically, the courier obtains the secret key about the position attribute, decrypts the ciphertext CT_n, and gets $Ciph_{Section_2}$ and HN.

When the receiver comes to pick up parcel, he (or she) should verify the correctness of order information on parcel and the courier should verify the authenticity of receiver. After mutual verification between courier and receiver, the receiver will get the parcel from courier successfully.

The details of reception of parcel are illustrated in Algorithm 6. Firstly, the receiver can get $E(PUB_{Rec}, [Section_2])$ from courier and decrypt it using his (or her) private key. Then the receiver obtains the following information: ID_{Order}, RN, $signature_{Sen}$, and personal information: Nam_{Sen}, Nam_{Rec}, Pho_{Sen}, and Pho_{Rec}. By verifying $signature_{Sen}$, he (or she) can confirm the integrity of parcel information. In addition, the receiver can compute $HN_1 = Hash(RN)$. By comparing the string HN_1 shown by receiver with the information HN which courier possesses, the courier can verify the validity of receiver. Then the receiver conducts $E(PUB_{Adm}, [ID_{Order} || RN])$ and sends it to administrator. The administrator can compute $HN_2 = Hash(RN)$ and check $HN_2 = HN$, in order to verify that the receiver has received the parcel.

5. Analysis of Scheme

In this section, we demonstrate that our scheme satisfies all the required properties.

5.1. Attribute-Based Access Control (ABAC). In our scheme, the logistics information is encrypted under CP-ABE. With

Input: ciphertext CT_n
Output: delivery result
1: courier executes:
2: $M_n \longleftarrow LIR(CT_n)$
3: $Ciph_{Section_2}||HN \longleftarrow M_n$
4: send $Ciph_{Section_2}$ to receiver
5: send $E(PUB_{Adm}, [HN])$ to administrator
6: receiver executes:
7: $Section_2 \longleftarrow D(PRI_{Rec}, [Ciph_{Section_2}])$
8: $HN_1 \longleftarrow Hash(RN)$
9: **if** $D(PUB_{Sen}, [signature_{Sen}]) = ID_{Order}||Plai_{Order}$ **then**
10: $Ciph_{FSC} \longleftarrow E(PUB_{Cou}, [ID_{Order}||HN_1])$
11: send $Ciph_{FSC}$ to courier
12: courier executes:
13: $ID_{Order}||HN_1 \longleftarrow D(PRI_{Cou}, [Ciph_{FSC}])$
14: **if** $HN_1 = HN$ **then**
15: receiver executes:
16: $Ciph_{FSA} \longleftarrow E(PUB_{Adm}, [ID_{Order}||RN])$
17: send $Ciph_{FSA}$ to administrator
18: administrator executes:
19: $ID_{Order}||RN \longleftarrow D(PRI_{Adm}, [Ciph_{FSA}])$
20: $HN \longleftarrow D(PRI_{Adm}, [E(PUB_{Adm}, [HN])])$
21: $HN_2 \longleftarrow Hash(RN)$
22: **if** $HN_2 = HN$ **then**
23: delivery result ⟵ Success
24: end if
25: **else**
26: delivery result ⟵ Receiver Wrong
27: end if
28: **else**
29: delivery result ⟵ Parcel Wrong
30: end if
31: **return** delivery result

Algorithm 6: Reception of parcel.

CP-ABE, the courier is specified with an attribute-based access policy for fine-grained access control of logistics address.

Specifically, the administrator selects attributes $\mathcal{S}$ and builds access tree T for couriers. The required logistics information M_i can only be decrypted by the valid courier whose attributes satisfy the access policy T_i. Different courier can only decrypt different segments of logistics information in accordance with their respective private keys, which correspond to different attributes sets $\{\mathcal{S}_i\}$ satisfying the access policy.

5.2. Position-Based Access Control (PBAC). Different from traditional cryptography, position-based cryptography uses the user's spatial position information as the only credential for user. In our scheme, the courier's position of valid delivery station at the expected work time is considered as one of the indispensable attributes. Specifically, the courier needs to obtain a secret key K_6 about position attribute $\mathcal{S}_{Spatial}$, so that he can continue the decryption process.

In addition, landmarks predetermine the keys K_6. It means that the expected value of position attribute is already known by all landmarks before protocol execution. Then the position attribute $\mathcal{S}_{Spatial}$ will be added to leaf nodes of access tree, so that the administrator can prebuild the access tree T.

5.3. Privacy-Preserving of Logistics Information (PPLI). In our scheme, besides sender and receiver, the whole delivery path is only known by administrator. The sender submits encrypted order information $Ciph_{Order}$ to administrator; then the administrator, respectively, encrypts the segmented logistics information $\{M_i\}$ under CP-ABE, as shown in Algorithm 3.

On the one hand, for general people who do not have the correct attributes, the only information they can get is ID_{Order}, which is used to uniquely identify the parcel. Apart from this, they cannot obtain anything about the plaintext of logistics information, i.e., Add_{Sen} and Add_{Rec}. On the other hand, the courier who actually moves parcel can only decrypt partial address information M_i according to attributes. M_i is the next delivery station address, which is necessary for the courier's delivery process.

5.4. Confidentiality of Personal Information(CPI). In initialization phase, customer's personal information which contains Nam_{Sen}, Nam_{Rec}, Pho_{Sen}, and Pho_{Rec} is encrypted by

receiver's public key PUB_{Rec}. In last phase, the parcel arrives at the final delivery station and the receiver obtains $Ciph_{Section_2}$ from the final courier and executes $D(PRI_{Rec}, [Ciph_{Section_2}])$, so that the receiver gets the personal information which is included in $Section_2$. In the whole parcel transit process, only the target receiver can decrypt the personal information. For other people including administrator and couriers, the probability of obtaining $Section_2$ is negligible even if they have $Ciph_{Section_2}$.

5.5. Verifiability of Receiver (VR). The courier can verify the correctness of receiver. In the phase of reception, the final courier can obtain HN using $LIR(CT_n)$. At the same time, the target receiver can get $Section_2$ using $D(PRI_{Rec}, [Ciph_{Section_2}])$, where RN is included in $Section_2$. Then the receiver computes $HN_1 = Hash(RN)$. By comparing the hash value HN_1 which receiver shows with HN, the courier can verify the correctness of receiver. For the adversary who wants to simulate the receiver, he (or she) must obtain RN. Because hash function is noninvertible and collision resistant and $Section_2$ cannot be decrypted without PRI_{Rec} which is kept secretly by receiver, the adversary has a negligible probability to obtain RN and HN. Consequently, our scheme can prevent forging identity by adversary to take away the parcel which does not belong to himself (or herself).

5.6. Verifiability of Parcel (VP). The digital signature guarantees the origin and integrity of parcel. The sender generates digital signature $signature_{Sen}$ by $E(PRI_{Sen}, [ID_{Order}||Plai_{Order}])$ and adds it into $Section_2$. Then $Section_2$ is encrypted using receiver's public key Pub_{Rec}. In the reception phase, the receiver can decrypt $Section_2$ and obtain $signature_{Sen}$. He (or she) can decrypt $E(PRI_{Sen}, [ID_{Order}||Plai_{Order}])$ using sender's public key and get $Plai_{Order}$, ID_{Order}. $Plai_{Order}$ contains receiver's personal information and address information, so that the receiver can confirm that the parcel is not forged. As a result, our scheme can prevent parcel forgery by malicious couriers during the express delivery process.

5.7. Anticollusion Attack to Attribute-Based Access Control (ACA-A). In the retrieval phase, the courier with appropriate attributes can only decrypt a segment M_i. However, couriers are so curious about other information that they may collude with others in order to enlarge their privileges. Our scheme ensures the courier only can decrypt the specific information according to his (or her) attributes.

For example, assuming that couriers C_1, C_2 have the attribute set $\mathcal{S}_1, \mathcal{S}_2$. Couriers C_1 and C_2 want to collude together. There is a courier C_3 who has the attribute set $\mathcal{S}_3$; let $\mathcal{S}_3 = S_1 \cup \mathcal{S}_2$. So C_1, C_2 want to obtain C_3's secret key and decrypt C_3's information. In our scheme, C_1 and C_2 must recover $e(g, g)^{\alpha s}$ in order to obtain C_3's secret key. In the phase of encryption, the string s from different couriers is randomized, so C_1 and C_2 cannot recover $e(g, g)^{\alpha s}$. It means C_3's ciphertext cannot be decrypted even if C_1 and C_2 collude. In other words, the collusion of multiple couriers is useless for decryption of addition logistics information.

5.8. Anticollusion Attack to Position-Based Access Control (ACA-P). The improved position-based key exchange (I-PBKE) which we propose is shown in Algorithm 4. It achieves that the colluding couriers who are not located at the valid station cannot obtain the logistics information. In other words, the position attribute's value K_6 cannot be distinguished with other random strings for colluding couriers.

Suppose there exists a set of malicious couriers. Let C_j be the malicious couriers between the honest courier at position P and landmarks $\{L_j \mid (1 \leq j \leq 4)\}$. At time T, C_1 can store$(K_1, K_5', A(X_4, K_1, K_5'))$, C_2 can store $(K_2', K_6', A(X_1, X_5, K_2', K_6'))$, C_3 can store $(K_3', A(X_2, K_3'))$, and C_4 can store $(K_4', A(X_3, K_4'))$, where $A(X_i, K_j)$ is any arbitrary adversarial algorithm.

After time T, the sequence of string reaching at adversary is different. Particularly, as for C_1, the string (X_1, X_5, K_2', K_6') first arrives and C_1 computes $K_2 = F(X_1, K_1) \oplus K_2'$. Later, the string (X_2, K_3') arrives and C_1 computes $K_3 = F(X_2, K_2) \oplus K_3'$. Finally, the string (X_3, K_4') arrives and C_1 computes $K_4 = F(X_3, K_3) \oplus K_4'$. Even if the malicious couriers collude together, according to properties of the ε-secure BSM PRG, the probability of the malicious couriers correctly guessing $K_5 = F(A(X_4), K_4) = F(X_4, K_4) \oplus K_5'$ is $\varepsilon + 2^{-\varphi}$, which is negligible in security parameter by choice of κ and r; thus $r \geq (2/\sigma)\kappa lb(n)$. So K_5 is still a random string to adversaries; thus they cannot find anything about K_6. Similarly, as for other possible reaching sequences, even if the malicious couriers collude together, they cannot find anything about final key K_6.

5.9. Undeniability (UD). Specifically, the sender cannot deny a parcel sent by himself (or herself), while the receiver cannot deny a parcel received by himself (or herself). In the step of initialization, the sender generates digital signature $signature_{Sen}$ which uses his (or her) private key PRI_{Sen} and then adds the signature $signature_{Sen}$ to $Section_2$. So, after decrypting $Section_2$, the receiver can confirm that the parcel was sent by sender. In the step of reception, when the receiver wants to take away the parcel, he (or she) should conduct $E(PUB_{Adm}, [ID_{Order}||RN])]$ and send it to administrator. Then the administrator can decrypt $E(PUB_{Adm}, [ID_{Order}||RN])]$ using Pri_{Adm}. At the same time, the administrator can receive HN sent by courier. By computing $HN_2 = Hash(RN)$ and checking $HN_2 = HN$, the administrator can verify that the receiver has received the parcel.

5.10. Unlinkability (UL). In the initialization phase, the sender's encrypted order information $Ciph_{Order}$ which contains Nam_{Sen}, Nam_{Rec}, Pho_{Sen}, Pho_{Rec}, Add_{Sen}, and Add_{Rec} and is transmitted to receiver. It is impossible for administrator and couriers to reveal the identity of customers from the encrypted order information.

As shown in Algorithm 1, a different random number RN is used in each generation of order ciphertext. $Section_2$ contains RN and $Section_1$ contains $HN(HN = Hash(RN))$, so each order information is different even if it is sent by the same sender. Since RN is random, the administrator and

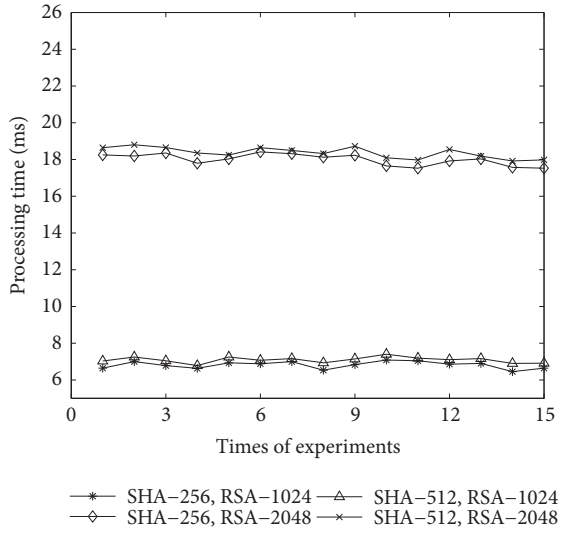


Figure 5: Computation overhead in initialization phase.

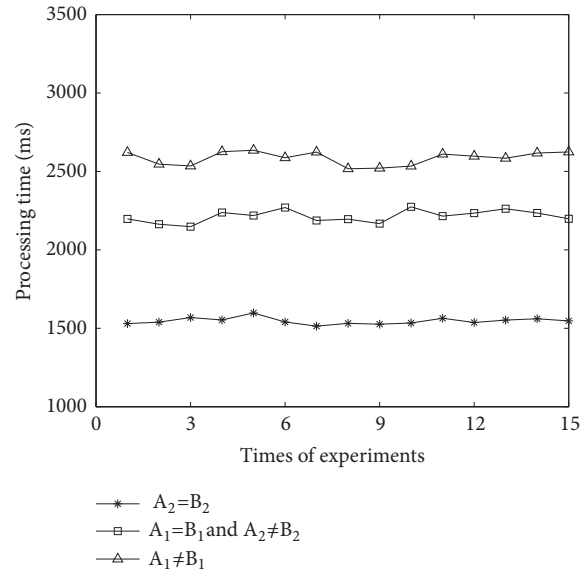


Figure 6: Computation overhead in encryption phase.

others are unable to tell whether these encrypted orders have the same logistics information (Add_{Sen}, Add_{Rec}) and personal information (Nam_{Sen}, Nam_{Rec}, Pho_{Sen}, and Pho_{Rec}). In other words, except customers, other people cannot distinguish whether the encrypted order information $Ciph_{Order}$ in many delivery processes originates from the same sender.

6. Performance Evaluation

In this section, we mainly focus on evaluation of computation overhead of our proposed scheme. The performance evaluation consists of four parts according to LIPPA, i.e., initialization, encryption, retrieval, and reception. The experiments are implemented on an Android phone (Band: Samsung Galaxy S7 Edge, CPU: Quad Core 2.15GHz, Operating System: Android 6.0, ROM:32G, RAM:4G). Our implementation is based on Java Pairing-Based Cryptography Library (JPBC).

Initialization. In the initialization phase, a sender wants to submit an encrypted order to administrator. Specifically, the sender needs to complete $Hash(RN)$, $E(PUB_{Rec}, [Section_2])$, $E(PUB_{Adm}, [Ciph_{Section_2}||Section_1])$, and $signature_{Sen}$. In Figure 5, we adopt hash function and RSA algorithm with different parameters to evaluate the sender's computation overhead. From Figure 5, we notice that the hash functions with different parameters lead to slightly different computation overhead. RSA algorithm with different parameters has greater impact on computational cost. When the hash function is SHA-512 and RSA algorithm is RSA-2048, the computation cost is still lower than 20ms. In general, the computation overhead of initialization is low for sender.

Encryption. In this phase, the computational cost of administrator mainly reflects on encryption of logistics information. In Figure 6, we compare the administrator's computation overhead under three types of delivery paths. The administrator, respectively, encrypts the segmented logistics information $M_1, M_2, \ldots, M_n$ (where $n = 3$ if $A_2 = B_2$; $n = 5$ if $A_2 \neq B_2$ and $A_1 = B_1$; $n = 6$ if $A_1 \neq B_1$) under CP-ABE. We set that the number of attributes in private key is fixed to 10, the number of leaf nodes in policy is fixed to 5, and the size of logistics information is 4kB. As shown in Figure 6, the processing time almost is between 1.5s and 2.6s. The computation overhead of encryption is efficient for administrator in practice.

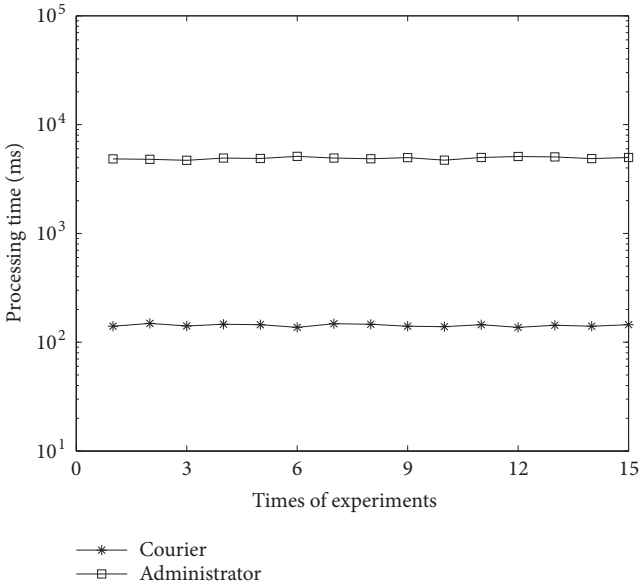


Figure 7: Computation overhead in retrieval phase.

Retrieval. In this phase, the courier and administrator collaborate to complete the retrieval. As shown in Figure 7, we evaluate the calculation costs of the administrator and the

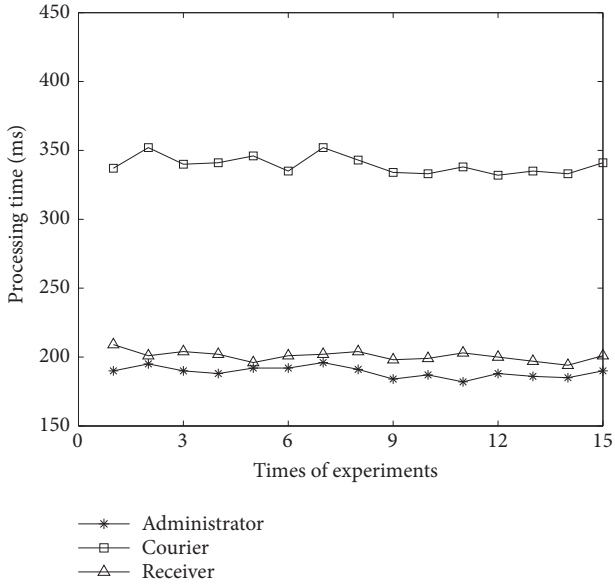


Figure 8: Computation overhead in reception phase.

courier, respectively. For courier, he (or she) needs to run I-PBKE to obtain K_6. Besides, the courier would decrypt the logistics information M_i. As a result, the computation overhead of courier consists of six $F(X_i, K_j)$ operations and one CP-ABE decryption operation. For administrator, he (or she) can employ landmarks to complete I-PBKE. At this protocol, landmarks need to generate five random large strings $\{X_i\}$. From Figure 7, we can see that the courier's computation overhead is far less than administrator's computation overhead.

Reception. In the reception phase, the courier delivers the parcel to receiver with the help of administrator. They use public key encryption and hash function specifically. We adopt SHA-256 and RSA-1024 to evaluate the computation overhead of administrator, courier, and receiver, respectively. As illustrated in Figure 8, the computation costs of administrator and receiver are all about 200ms. As for courier, he (or she) needs to complete extra decryption of logistics information. As a result, the courier's computation overhead is higher. In general, their computation overhead is all acceptable for real parcel delivery process.

In general, our scheme is efficient and feasible in practice. What is more, our scheme satisfies all the security requirements of LIoT. So the LIP-PA is available for mobile devices in real parcel delivery scenario.

7. Conclusions

In this paper, we propose LIP-PA, a logistics information privacy protection scheme with position and attribute-based access control on mobile devices. Different from existing schemes, our scheme provides privacy protection for both personal information and logistics information. In our scheme, customers could achieve verifiability and undeniability. The administrator could encrypt the logistics information based on the policy of position and attribute-based access control. In order to transmit the parcel to next station, couriers could only decrypt the required segment of logistics information but not all the logistics information. As a further contribution, we prove that our scheme can satisfy all the security requirements and show that it is available for mobile devices in practice based on the experiment results.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was funded by National Natural Science Foundation of China (61472310, U1536202, U1405255, 61672413, 61672415, 61671360, 61602360, and 61702404) and China 111 Project (Grant B16037).

References

[1] L. D. Xu, W. He, and S. Li, "Internet of things in industries: a survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.

[2] C. Sun, "Application of RFID technology for logistics on internet of things," *AASRI Procedia*, vol. 1, pp. 106–111, 2012.

[3] L. Tarjan, I. Šenk, S. Tegeltija, S. Stankovski, and G. Ostojic, "A readability analysis for QR code application in a traceability system," *Computers and Electronics in Agriculture*, vol. 109, pp. 1–11, 2014.

[4] L. Ye, Y. Wang, and J. Chen, "Research on the intelligent warehouse management system based on near field communication (NFC) technology," *International Journal of Advanced Pervasive and Ubiquitous Computing*, vol. 8, no. 2, pp. 38–55, 2016.

[5] A. Orsino, A. Ometov, G. Fodor et al., "Effects of heterogeneous mobility on D2D- and drone-assisted mission-critical MTC in 5G," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 79–87, 2017.

[6] D. Wang, H. Cheng, D. He, and P. Wang, "On the challenges in designing identity-based privacy-preserving authentication schemes for mobile devices," *IEEE Systems Journal*, vol. 12, no. 1, pp. 916–925, 2018.

[7] Y. Zhang, M. Yang, G. Gu, and H. Chen, "Rethinking permission enforcement mechanism on mobile systems," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2227–2240, 2016.

[8] J. Yu, B. Zhang, Z. Kuang, D. Lin, and J. Fan, "IPrivacy: image privacy protection by identifying sensitive objects via deep multi-task learning," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 5, pp. 1005–1016, 2017.

[9] J. Shen, T. Zhou, F. Wei, X. Sun, and Y. Xiang, "Privacy-preserving and lightweight key agreement protocol for V2G in the social internet of things," *IEEE Internet of Things Journal*, pp. 1-1, 2017.

[10] Q. Wei, C. Wang, and X. Li, "Express information privacy protection application based on RSA," *Application of Electronic Technique*, vol. 40, no. 7, pp. 58–60, 2014.

[11] W. Hu, Q. Wu, and C. Gu, "Scheme design of logistic personal information privacy protection based on QR code," *Communications Technology*, vol. 50, no. 9, pp. 2074–2079, 2017.

[12] D. He, M. Ma, S. Zeadall, N. Kumar, and K. Liang, "Certificateless public key authenticated encryption with keyword search for industrial internet of things," *IEEE Transactions on Industrial Informatics*, 2017.

[13] J. Shen, J. Shen, X. Chen, X. Huang, and W. Susilo, "An efficient public auditing protocol with novel dynamic structure for cloud data," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 10, pp. 2402–2415, 2017.

[14] D. He, S. Zeadally, N. Kumar, and W. Wu, "Efficient and anonymous mobile user authentication protocol using self-certified public key cryptography for multi-server architectures," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2052–2064, 2016.

[15] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, "Zipfs law in passwords," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

[16] D. Wang, Z. Zhang, P. Wang, J. Yan, and X. Huang, "Targeted online password guessing: an underestimated threat," in *Proceedings of the 23rd ACM Conference on Computer and Communications Security (CCS '16)*, pp. 1242–1254, October 2016.

[17] D. He, N. Kumar, M. K. Khan, and J.-H. Lee, "Anonymous two-factor authentication for consumer roaming service in global mobility networks," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 4, pp. 811–817, 2013.

[18] D. Wang, D. He, P. Wang, and C.-H. Chu, "Anonymous two-factor authentication in distributed systems: certain goals are beyond attainment," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 428–442, 2015.

[19] 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. 1-1, 2016.

[20] J. Shen, T. Zhou, D. He, Y. Zhang, X. Sun, and Y. Xiang, "Block design-based key agreement for group data sharing in cloud computing," *IEEE Transactions on Dependable and Secure Computing*, vol. PP, no. 99, 2017.

[21] J. Shen, T. Zhou, X. Chen, J. Li, and W. Susilo, "Anonymous and traceable group data sharing in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 912–925, 2018.

[22] Q. Wei and L. I. Xing-Yi, "Express information protection application based on K-anonymity," *Application Research of Computers*, vol. 31, no. 2, pp. 555–567, 2014.

[23] X. Zhang, H. Li, Y. Yang, G. Sun, and G. Chen, "LIPPS: logistics information privacy protection system based on encrypted QR code," in *Proceedings of the 2016 IEEE Trustcom/BigDataSE/ISPA*, pp. 996–1000, Tianjin, China, August 2016.

[24] H. Qi, D. Chenjie, Y. Yingbiao, and L. Lei, "A new express management system based on encrypted QR code," in *Proceedings of the 2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pp. 53–56, Nanchang, China, June 2015.

[25] T. Li, R. Zhang, and Y. Zhang, "PriExpress: privacy-preserving express delivery with fine-grained attribute-based access control," in *Proceedings of the 2016 IEEE Conference on Communications and Network Security, CNS 2016*, pp. 333–341, USA, October 2016.

[26] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *Proceedings of the IEEE Symposium on Security and Privacy (SP '07)*, pp. 321–334, May 2007.

[27] Q. Jiang, J. Ma, C. Yang, X. Ma, J. Shen, and S. A. Chaudhry, "Efficient end-to-end authentication protocol for wearable health monitoring systems," *Computers and Electrical Engineering*, vol. 63, pp. 182–195, 2017.

[28] J. Shen, D. Liu, M. Z. Bhuiyan, J. Shen, X. Sun, and A. Castiglione, "Secure verifiable database supporting efficient dynamic operations in cloud computing," *IEEE Transactions on Emerging Topics in Computing*, pp. 1-1, 2017.

[29] D. He, S. Zeadally, N. Kumar, and J.-H. Lee, "Anonymous authentication for wireless body area networks with provable security," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2590–2601, 2016.

[30] D. He and D. Wang, "Robust biometrics-based authentication scheme for multiserver environment," *IEEE Systems Journal*, vol. 9, no. 3, pp. 816–823, 2015.

[31] J. Zhang, J. Ma, C. Yang, and L. Yang, "Universally composable secure positioning in the bounded retrieval model," *Science China Information Sciences*, vol. 58, no. 11, pp. 1–15, 2015.

[32] N. Chandran, V. Goyal, R. Moriarty, and R. Ostrovsky, "Position based cryptography," in *Advances in cryptology—CRYPTO 2009*, vol. 5677 of *Lecture Notes in Comput. Sci.*, pp. 391–407, Springer, Berlin, 2009.

[33] Q. Jiang, Z. Chen, B. Li, J. Shen, L. Yang, and J. Ma, "Security analysis and improvement of bio-hashing based three-factor authentication scheme for telecare medical information systems," *Journal of Ambient Intelligence and Humanized Computing*, no. 5, pp. 1–13, 2017.

[34] K. Jia, X. Chen, and G. Xu, "The improved public key encryption algorithm of kerberos protocol based on braid groups," in *Proceedings of the 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, pp. 1–4, Dalian, China, October 2008.

[35] J. Zhang, J. Ma, and S. Moon, "Universally composable one-time signature and broadcast authentication," *Science China Information Sciences*, vol. 53, no. 3, pp. 567–580, 2010.

[36] X. Dong, H. Qian, and Z. Cao, "Provably secure RSA-type signature based on conic curve," *Wireless Communications and Mobile Computing*, vol. 9, no. 2, pp. 217–225, 2009.

[37] W.-B. Hsieh and J.-S. Leu, "A dynamic identity user authentication scheme in wireless sensor networks," in *Proceedings of the 2013 9th International Wireless Communications and Mobile Computing Conference, IWCMC 2013*, pp. 1132–1137, Italy, July 2013.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 8539674, 11 pages
https://doi.org/10.1155/2018/8539674

Research Article

Cryptanalysis and Security Enhancement of Three Authentication Schemes in Wireless Sensor Networks

Wenting Li[1,2] **Bin Li**[1] **Yiming Zhao**[1] **Ping Wang**[1,3] **and Fushan Wei**[2]

[1]*School of Software and Microelectronics, Peking University, Beijing 100871, China*
[2]*State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China*
[3]*National Engineering Research Center for Software Engineering, Peking University, Beijing 100871, China*

Correspondence should be addressed to Ping Wang; pwang@pku.edu.cn

Received 4 April 2018; Accepted 28 May 2018; Published 5 July 2018

Academic Editor: Joseph Liu

Nowadays wireless sensor networks (WSNs) have drawn great attention from both industrial world and academic community. To facilitate real-time data access for external users from the sensor nodes directly, password-based authentication has become the prevalent authentication mechanism in the past decades. In this work, we investigate three foremost protocols in the area of password-based user authentication scheme for WSNs. Firstly, we analyze an efficient and anonymous protocol and demonstrate that though this protocol is equipped with a formal proof, it actually has several security loopholes been overlooked, such that it cannot resist against smart card loss attack and violate forward secrecy. Secondly, we scrutinize a lightweight protocol and point out that it cannot achieve the claimed security goal of forward secrecy, as well as suffering from user anonymity violation attack and offline password guessing attack. Thirdly, we find that an anonymous scheme fails to preserve two critical properties of forward secrecy and user friendliness. In addition, by adopting the "perfect forward secrecy (PFS)" principle, we provide several effective countermeasures to remedy the identified weaknesses. To test the necessity and effectiveness of our suggestions, we conduct a comparison of 10 representative schemes in terms of the underlying cryptographic primitives used for realizing forward secrecy.

1. Introduction

Currently, wireless sensor networks (WSNs) have become one of the most standard services employed in commercial and industrial applications and proved to be a leading area of research [1–3]. Like many advanced technologies, the original appliance of WSNs can be found in military and heavy industrial applications. In the 1950s, the first modern WSN—the Sound Surveillance System (SOSUS)—is developed by the United States Military and used for detecting Soviet submarines [4]. Nowadays, WSNs thrive in industrial and consumer applications, including machine health monitoring, environmental sensing, natural disaster prevention, and health care monitoring [5–7].

A wireless sensor network generally includes a central gateway node (*GWN*, so-called base station), a large number of circulating, self-directed and low powered devices named sensor nodes, and a set of end users. The *GWN* acts as a bridge between WSNs and the other networks and also a powerful data managing and processing center. Sensor nodes are multifunctional, energy efficient devices and are spatially distributed over the networks for caringly collecting, processing, and transferring data.

In many critical applications, remote users are usually keen on real-time accessing with sensor nodes [8, 9], yet if data queries are carried out by the gateway node, efficiency and accuracy might not be guaranteed over the long transmission path between *GWN* and the sensors. Accordingly, password-based user authentication proves to be a proper solution for this issue as its security, simplicity, and portability [10–12]. That is, users are first authenticated by remote sensor nodes before being permitted to access data.

In 2006, Wong et al. [13] proposed the first password-based authentication scheme for wireless sensor networks that allows legitimate users to query sensitive information at every sensor of the network. However, shortly after this protocol was presented, Tseng et al. [14] and Das et al. [15] pointed out that Wong et al.'s scheme [13] is vulnerable

to replay attack, forgery attack, and node capture attack separately, then an enhanced one based on smart card was firstly proposed by Das et al. [15]. Unfortunately, Khan and Alghathbar [16], Chen et al. [17], and Yeh et al. [18] pointed out some weaknesses in Das et al.'s scheme, such as suffering from impersonation attack, insider attack, and the violation of user anonymity and key agreement. Then some improvements are made in these works.

However, in 2013, Shi and Gong [19] found that Yeh et al.'s scheme [18] cannot achieve mutual authentication and user anonymity, then they proposed an efficient ECC-based authentication scheme for WSNs. At the same time, Khan and Alghathbar's protocol [16] was also proven insecure against insider attack, smart card loss attack and forgery attack by Vaidya et al. [20] and Chen et al.'s scheme [17] was shown as vulnerable to impersonation attack, replay attack and *GWN* by passing attack in [21]. Later, Choi et al. [22] demonstrated that Shi and Gong's scheme [19] is vulnerable to smart card loss attack and an enhanced scheme was given in [22]. Meanwhile, Xue et al. [23] presented a temporal credential-based two-factor (i.e., smart card and password) authentication scheme for WSNs. Although their scheme retains many admirable properties, there are some weaknesses being found by researchers [8, 11, 24], such as offline password guessing, insider, impersonation, and tracking attacks.

Quite recently, Li et al. [25] analyzed the security of Jiang et al.'s scheme [11] (an improvement based on [23]) and showed that their scheme suffers from user friendliness issue, desynchronization problem, and is inapplicable for WSN environments. Then a new scheme was proposed in [25]; however, in this paper, we reveal that Li et al.'s scheme [25] still fails to eliminate the security pitfalls of smart card loss attack and the violation of forward secrecy. At the same time, we find that the newly proposed schemes by Amin et al. [8] and Wu et al. [9] are prone to the same security defects with Li et al.'s scheme [25].

From the above analysis, it can be seen that many of the previous protocols are not much satisfactory. On the one hand, this is because the lack of necessary principles. Some principles that have been proven are still ignored in the design of the protocol, such as user anonymity principle [7] and perfect forward secrecy principle [26]. On the other hand, the protocol designers usually do not follow unified evaluation criteria, and they tend to emphasize the advantages of their new designed protocol, but ignore its inadequacies. Besides reporting the security flaws in [8, 9, 25], we also provide effective countermeasures and refinements to overcome these pitfalls, accordingly, examine the necessary of our suggestions.

Contributions. In this work, we mainly review and analyze three state-of-the-art authentication protocols proposed by Li et al. [25], Amin et al. [8], and Wu et al. [9]. And reveal that all these three schemes suffer from smart card loss issue and cannot achieve forward secrecy. Then we suggest several possible countermeasures to overcome these pitfalls. We also provide a comparison of 10 representative schemes for wireless sensor networks which emphatically considered how and with what technology did they realize forward secrecy. This illustrates the necessity and effectiveness of our suggestions and provides a better understanding of the exiting schemes.

Table 1: Notations and abbreviations.

Symbol	Description
U_i	i^{th} user
S_j	j^{th} remote sensor node
GWN	gateway node
$\mathscr{A}$	malicious attacker
ID_i, SID_j	identity of U_i and S_j
PW_i	password of U_i
x, K	the secret key of gateway node GWN
SC	a smart card
$\oplus$	the XOR operation
$\|$	the concatenation operation
$h(\cdot)$	one-way hash function
$\longrightarrow$	the public channel
$\Longrightarrow$	the secure channel

Organization. The remainder of this paper is organized as follows. Section 2 reviews and demonstrates the pitfalls of Li et al.'s scheme. Section 3 cryptanalyzes Amin et al.'s protocol with proper countermeasures over discovered flaws. Section 4 describes the weaknesses of Wu et al.'s protocol and compares 10 representative schemes. The conclusion is made in Section 5.

2. Cryptanalysis of Li et al.'s Scheme

Earlier in 2018, Li et al. [25] presented a three-factor anonymous and efficient authentication scheme for wireless sensor networks. Although their scheme has many attractive properties, such as the provision of user anonymity and local password change, it still fails to attain many of the claimed goals. In this section, we will demonstrate that though Li et al. try to settle the user friendliness issue of Jiang et al.'s scheme [11], their solution leads to offline dictionary attack. And we also observe that Li et al.'s scheme cannot preserve forward secrecy, which is the most crucial goal for WSNs.

2.1. Review of Li et al.'s Scheme. In this subsection, we briefly revisit Li et al.'s scheme [25]. For ease of description, some intuitive notations and abbreviates are listed in Table 1 and will appear throughout this paper. Their scheme includes three main phases: registration, login and authentication, and password change. We will follow their presentations as close as possible.

2.1.1. Registration Phase. Before the registration phase of Li et al.'s [25] scheme, the gateway node defines a finite cyclic group $\mathbb{G} = < P >$ of order a large prime number n. This group could be an elliptic curve group, or it could be a prime order subgroup of $\mathbb{Z}_p^*$. Then GWN chooses two random numbers x,

K as its master secret key and computes $X = xP$ as its public key. Ultimately, *GWN* publishes $\{E(F_p), G, P, X\}$ and stores x, K securely.

Sensor Registration Phase. *GWN* chooses an identity SID_j and computes the secret key $K_S = h(SID_j \parallel K)$ for each sensor node. Then, *GWN* embeds $\{SID_j, K_S\}$ in the memory of S_j and deploys it in the particular area.

User Registration Phase. When a user U_i aims to acquire the sensitive information of remote sensor nodes, the following procedure is carried out by U_i firstly.

(1) U_i chooses an identity ID_i, a password PW_i, and a nonce a_i and calculates $RPW_i = h(PW_i \parallel a_i)$. Then U_i imprints his/her biometric b_i on a specific device.

(2) $U_i \Longrightarrow GWN$: $\{ID_i, RPW_i, b_i\}$.

(3) Once obtaining U_i's registration request, *GWN* generates a random codeword $c_i \in C$ and computes $F(c_i, b_i) = (\alpha, \delta)$, where $\alpha = h(c_i)$ and $\delta = c_i \oplus b_i$. *GWN* further computes $A_i = h(ID_i \parallel RPW_i \parallel c_i)$ and $B_i = h(ID_i \parallel K) \oplus h(RPW_i \parallel c_i)$ and keeps $\{\alpha, \delta, A_i, B_i, X, f(\cdot)\}$ into a new smart card *SC*. At the same time, *GWN* stores ID_i in its database.

(4) $GWN \Longrightarrow U_i : \{\alpha, \delta, A_i, B_i, X, f(\cdot)\}$.

(5) When receiving the smart card, U_i stores a_i into it.

2.1.2. Login and Authentication Phase. In this phase, the following steps are performed by U_i, S_j, and *GWN* as well as negotiating a session key.

(1) U_i inserts *SC* into a card reader and inputs b_i^* on a specific device. Then *SC* computes $c_i^* = f(\delta \oplus b_i^*) = f(c_i \oplus (b_i \oplus b_i^*))$ and checks whether $h(c_i^*) \stackrel{?}{=} \alpha = h(c_i)$. If not, *SC* terminates the session. Otherwise, *SC* asks U_i to input ID_i and PW_i and computes $A_i^* = h(ID_i \parallel h(PW_i \parallel a_i) \parallel c_i^*)$. Then *SC* checks whether $A_i^* \stackrel{?}{=} A_i$. If it does not hold, *SC* rejects the session. Otherwise, *SC* chooses two random numbers r_i and s and then calculates $M_1 = B_i \oplus h(h(PW_i \parallel a_i) \parallel c_i^*)$, $M_2 = sP$, $M_3 = sX = sxP$, $M_4 = ID_i \oplus M_3$, $M_5 = M_1 \oplus r_i$, $M_6 = h(ID_i \parallel r_i) \oplus SID_j$, and $M_7 = h(M_1 \parallel SID_j \parallel M_3 \parallel r_i)$.

(2) $U_i \longrightarrow GWN$: $\{M_2, M_4, M_5, M_6, M_7\}$.

(3) Upon receiving the login request, *GWN* computes $M_3 = xM_2 = xsP$ and $ID_i = M_4 \oplus M_3$ and verifies if ID_i is in the database. If not, the request is aborted. Otherwise, *GWN* computes $M_1 = h(ID_i \parallel K)$, $r_i = M_5 \oplus M_1$, $SID_j = M_6 \oplus h(ID_i \parallel r_i)$, and $M_7^* = h(M_1 \parallel SID_j \parallel M_3 \parallel r_i)$ and checks whether $M_7^* \stackrel{?}{=} M_7$. If it does not hold, *GWN* terminates the session. Otherwise, *GWN* selects a random number r_g and computes $K_S = h(SID_j \parallel K)$, $M_8 = ID_i \oplus K_S$, $M_9 = r_g \oplus h(ID_i \parallel K_S)$, $M_{10} = r_g \oplus r_i$, and $M_{11} = h(ID_i \parallel SID_j \parallel K_S \parallel r_i \parallel r_g)$.

(4) $GWN \longrightarrow S_j : \{M_8, M_9, M_{10}, M_{11}\}$.

(5) When receiving the message, S_j computes $ID_i = M_8 \oplus K_S$, $r_g = h(ID_i \parallel K_S) \oplus M_9$, $r_i = r_g \oplus M_{10}$, and $M_{11}^* = h(ID_i \parallel SID_j \parallel K_S \parallel r_i \parallel r_g)$ and checks whether $M_{11}^* \stackrel{?}{=} M_{11}$. If not, the session is rejected. Otherwise, S_j selects a random number r_j and computes $M_{12} = r_j \oplus K_S$, $SK_j = h(ID_i \parallel SID_j \parallel r_i \parallel r_g \parallel r_j)$, and $M_{13} = h(K_S \parallel SK_j \parallel r_j)$.

(6) $S_j \longrightarrow GWN$: $\{M_{12}, M_{13}\}$.

(7) After getting the response message, *GWN* computes $r_j = M_{12} \oplus K_S$, $SK = h(ID_i \parallel SID_j \parallel r_i \parallel r_g \parallel r_j)$, and $M_{13}^* = h(K_S \parallel SK \parallel r_j)$ and checks whether $M_{13}^* \stackrel{?}{=} M_{13}$. If not, *GWN* aborts the session. Otherwise, *GWN* calculates $M_{14} = M_1 \oplus r_g$, $M_{15} = r_i \oplus r_j$, and $M_{16} = h(ID_i \parallel SK \parallel r_g \parallel r_j)$.

(8) $GWN \longrightarrow U_i : \{M_{14}, M_{15}, M_{16}\}$.

(9) When U_i receiving the response message, *SC* computes $r_g = M_{14} \oplus M_1$, $r_j = M_{15} \oplus r_i$, $SK_i = h(ID_i \parallel SID_j \parallel r_i \parallel r_g \parallel r_j)$, and $M_{16}^* = h(ID_i \parallel SK_i \parallel r_g \parallel r_j)$ and checks whether $M_{16}^* \stackrel{?}{=} M_{16}$. If it does not hold, U_i terminates the connection. Otherwise, U_i and S_j establish a connection with a session key.

2.2. Cryptanalysis of Li et al.'s Scheme. A concrete and concise adversarial model is essential for a good design of user authentication scheme in wireless sensor networks. Though lacking of specification in Li et al.'s scheme [25], the following assumptions about the adversary's capabilities are implicitly made in [25]:

(1) Two communication channels exist: one is a secure, or a private channel which is mainly used for registration; another is a public channel which acts on login and authentication phases. As in the conventional authentication protocols, the adversary $\mathscr{A}$ is modeled to have full control of the public channel; i.e., $\mathscr{A}$ can eavesdrop, intercept, and modify and redirect any transmitted messages between the communication parties [3, 6].

(2) The user-memorable identities and passwords are of low entropy and can be offline enumerated by $\mathscr{A}$ at the same time within polynomial time.

(3) When considering truly multifactor authentication (i.e., the scheme is secure even if one or more factors are cracked [10]), it is rational to assume that $\mathscr{A}$ may (i) learn a victim's password such as phishing or shoulder surfing attacks, (ii) extract the secret parameters in the lost smart card by side-channel attack, or (iii) obtain a victim's biometric information via malicious device, but cannot achieve all. Otherwise, it is a trivial case.

(4) To delineate the critical feature of forward secrecy, $\mathscr{A}$ is allowed to corrupt any valid entities to obtain its longterm secret key(s). In addition, previous session key(s) may be revealed by $\mathscr{A}$ as a possible reason of improper erasure [10, 27].

It is worth noting that the above adversarial model, following the existing works in [3, 6, 7, 10, 28], is one of the few ones that apply to multifactor authentication in WSNs. For the sake of user friendliness, many protocols allow their users to select his/her identity *ID* and password *PW*. However, the user usually chooses easy-to-remember identity (e.g., email, phone number) and password, which are of low entropy ($|\mathscr{D}_{id}| \leq |\mathscr{D}_{pw}| \leq 10^6$ [29, 30]) and can be offline enumerated by $\mathscr{A}$ within polynomial time. Besides, assumption (3) specifies truly three-factor security and assumption (4) is used to capture the crucial notion of forward secrecy when *GWN* or any sensor node S_j is corrupted. In the following sections, our analysis will take account of these four assumptions.

2.2.1. Smart Card Loss Attack. In [25], Li et al. pointed out that Jiang et al.'s scheme [11] lacks timely detection mechanism, which means once a user inputs wrong identity or password unintentionally, the system will remain executing the following login and authentication phases. Undoubtedly, this interaction process will bring extra cost. In reality, it is a common accident as users usually involve in countless applications and manage various pairs of identity and password [7]. To solve this problem, Li et al.'s scheme [25] inserts a verification item $A_i = h(ID_i \parallel h(PW_i \parallel a_i) \parallel c_i)$ in the smart card for the purpose of providing timely detection and performing password change without any interaction with the *GWN*. However, their modification goes back to the "security-usability" balance problem proposed by Huang et al. [12]; that is, it realizes local password change but brings offline dictionary attack. We illustrate this attack as below.

Step 1. $\mathscr{A}$ chooses a pair (ID_i^*, PW_i^*) from $\mathscr{D}_{id} \times \mathscr{D}_{pw}$, where $\mathscr{D}_{id}$ represents the identity space and $\mathscr{D}_{pw}$ represents the password space.

Step 2. $\mathscr{A}$ computes $A_i^* = h(ID_i^* \parallel h(PW_i^* \parallel a_i) \parallel c_i)$, where a_i is extracted from the victim's smart card and c_i can be obtained by computing $c_i = f(\delta \oplus b_i)$ with the help of malicious device.

Step 3. $\mathscr{A}$ verifies the correctness of (ID_i^*, PW_i^*) pair by checking whether the computed A_i^* equals the extracted A_i.

Step 4. $\mathscr{A}$ repeats the above Steps 1 ~ 3 until the right values are found.

Besides the previous reasonable assumption (3), it should be pointed out that, in the registration phase of Li et al.'s scheme [25], U_i imprints his/her biometric information b_i on a specific device and simply submits the plain-text b_i to *GWN*. Then, *GWN* employs the fuzzy commitment technology [31] and the generated α to compute δ. In such situation, if a privileged insider, e.g., the administrator, has learned the user's biometric information, she is able to complete the above offline guessing attack. Of course, she is able to impersonate the victim to login other applications as biometric characteristics cannot be easily changed.

For another, in order to realize user friendliness, most password-based authentication schemes (e.g., [8, 9, 11]) allow users to choose his/her own *ID* and *PW*, and Li et al.'s scheme is no exception. However, users usually tend to choose easy-to-remember and thus of low entropy identities and passwords, so that it is reasonable to make the assumption (2) that $\mathscr{A}$ can offline enumerate all the (ID, PW) pairs within polynomial time. The running time of the above attack procedure is $\mathcal{O}(2T_H \times |\mathscr{D}_{id}| \times |\mathscr{D}_{pw}|)$, where $|\mathscr{D}_{id}|$ denotes the number of identities, $|\mathscr{D}_{pw}|$ denotes the number of passwords, and T_H is the running time for Hash operation. Since $|\mathscr{D}_{id}|$ and $|\mathscr{D}_{pw}|$ are very limited in practice (e.g., $|\mathscr{D}_{id}| \leq |\mathscr{D}_{pw}| \leq 10^6$ [29, 30]), our above attack is meaningful and poses a real challenge to user authentication protocols for wireless sensor networks.

2.2.2. The Violation of Forward Secrecy. WSNs are generally deployed in security-critical applications, such as battlefield surveillance and health care monitoring [7, 27, 32, 33]. The sensor nodes at risk had been driven: on one hand, due to the unattended environments and low-cost considerations, it is easier for an adversary $\mathscr{A}$ to focus on sensors access to breakthrough success; on the other hand, sensors often perform extremely sensitive tasks and thus, they preserve sensitive information and exhibit greater attack surface. Consequently, sensor nodes are more vulnerable to serious attacks, so that an admired authentication scheme for WSNs ought to be guaranteed against node capture attack.

Unfortunately, Li et al.'s scheme [25] cannot resist against this severe node capture attack. Let us consider the following scenarios. In case a sensor node S_k has been compromised by an adversary $\mathscr{A}$ and the stored secret key K_S can be extracted. This assumption is sound as made in assumption (4) and it is also implicitly described in Li et al.'s scheme [25]. With the extracted K_S, $\mathscr{A}$ can successfully obtain the previous session key between S_k and any user U_i, as follows.

Step 1. Eavesdrop and intercept the message $\{M_8, M_9, M_{10}, M_{11}\}$ sending from *GWN* to S_k.

Step 2. Compute $ID_i = M_8 \oplus K_S$, $r_g = h(ID_i \parallel K_S) \oplus M_9$, and $r_i = r_g \oplus M_{10}$.

Step 3. Intercept the message $\{M_{12}, M_{13}\}$ sending from S_k to *GWN*.

Step 4. Compute $r_k = M_{12} \oplus K_S$.

Step 5. Intercept the login message $\{M_2, M_4, M_5, M_6, M_7\}$ sending from U_i to *GWN*.

Step 6. Compute $SID_k = M_6 \oplus h(ID_i \parallel r_i)$.

Step 7. Compute the previous session key as $SK = h(ID_i \parallel SID_k \parallel r_i \parallel r_g \parallel r_k)$.

There are some points to be noted regarding the aforementioned attack. Firstly, the reason why we add Steps 5 and 6 is that these two steps are conducive to check the parameters though $\mathscr{A}$ has already known SID_k. Then, it is not hard to see that $\mathscr{A}$ only needs to eavesdrop over the public channel with simple computations to complete the aforementioned attack procedure. Consequently, the desirable security goal of perfect forward secrecy (PFS) cannot be attained by Li et al.'s scheme.

Despite considerable attention has been paid to forward secrecy issue, many prior works still explicitly or implicitly use an incorrect computation for the session key(s) (e.g., [8, 9, 21, 34]). This is mainly due to the violation of the "PFS principle" suggested in [26]: (i) public-key techniques are indispensable; (ii) at least two exponentiation operations are conducted on server side. Though Ma et al. [26] emphasize this principle on client-server architecture, after careful analysis, we find this "PFS principle" is suitable for WSN environments (i.e., three-party environment). In this cases,

we will take *GWN* and sensors as server side, while keeping users as client.

Accordingly, elliptic curves cryptosystem (ECC) is a reasonable choice for overcoming this pitfall, whereas in their original scheme [25] Li et al. employ this mechanism to greatly attain user anonymity. To make a precisely modification, we assume r_i to be r_iP and r_j to be r_jP, where point P is a generator mentioned before and r_i, r_j are two random numbers chosen by U_i and S_j separately. Note that *GWN* has no need to be involved in negotiating the session key. Then in this way, the session key can be recalculated as $SK = h(ID_i \parallel SID_j \parallel r_iP \parallel r_jP \parallel r_ir_jP)$. As it is generated by session-variant random numbers r_i and r_j and computationally infeasible to guess r_ir_jP from transmitted message due to discrete logarithm problem, Li et al.'s scheme [25] will be secure against node capture attack and provide forward secrecy perfectly after slight modifications.

2.2.3. Mistakes in the Proof. The emergence of BAN logic opens up a new chapter in the proof of user authentication protocol [35, 36]; it can not only be used to prove whether the protocol achieves some desired goals, but also be employed to find some defects in the protocol. However, there still are some problems in the application of BAN logic. On the one hand, BAN logic cannot prove whether the protocol achieves all security goals and desirable properties. For example, it cannot prove that the protocol resists against parallel session attack, denial-of-service attack, node capture attack, etc. On the other hand, the analysis of BAN logic depends on some basic assumptions and the initial hypotheses. If the initial hypotheses was not sound, the formal analysis will lead to erroneous conclusions.

In the formal proof of Li et al.'s scheme [25] with BAN logic, there are several minor problems. Firstly, Li et al. add a new logic rule, *session keys rule*:

$$\frac{P \mid\equiv \#(X), P \mid\equiv Q \mid\equiv X}{P \mid\equiv P \mid \overset{K}{\longleftrightarrow} Q} \quad (1)$$

However, it is better to explain the calculation method of K and the key role of X in K. Otherwise, we cannot derive that P believes P and Q share K from the upper part of the equation.

Secondly, we suggest that the initiative premises p13 and p14, i.e., $GWN| \equiv S_j \Longrightarrow r_j \oplus K$ and $S_j| \equiv GWN \Longrightarrow r_g \oplus h(ID_i \parallel K)$, respectively, should be derived from the translation messages, but not in the premises. Finally, they may ignore some details in the formal proof, such as in the D5, it is better to add $GWN| \equiv \#r_i$, which we cannot find in the assumption or derive from the front. It also can be seen that the correctness of the protocol cannot be guaranteed only by using the formal proof.

3. Cryptanalysis of Amin et al.'s Scheme

Recently, Amin et al. [8] proposed a lightweight protocol for IoT-enabled devices for cloud computing environments. The private information is usually stored in distributed cloud servers (e.g., sensors), so that distributed nodes are confronted with the same security threats of sensors in wireless sensor networks. After careful analysis, we find that though equipped with a formal proof and exhibiting great application prospects, Amin et al.'s scheme still cannot resist against smart card loss attack and also fail to provide user anonymity and forward secrecy.

3.1. Review of Amin et al.'s Scheme. Here we briefly review the scheme proposed by Amin et al. [8], an enhancement over Xue et al.'s scheme [37] and Chuang et al.'s scheme [38].

3.1.1. Registration Phase. The registration phase of Amin et al.'s scheme can be divided into cloud server registration and user registration.

Cloud Server Registration Phase. In this phase, any cloud server S_j sends a self-chosen identity and random number pair $\{SID_j, d\}$ to control server (CS). Then CS chooses a random number y, computes $P_j = h(SID_j \parallel d)$, $BS_j = h(P_j \parallel y)$, and responds $\{BS_j\}$ to S_j securely. Finally, S_j stores $\{BS_j, d\}$ in the memory.

User Registration Phase. Firstly, a user U_i chooses his/her identity ID_i, password PW_i, and two random numbers $< b_1, b_2 >$. Then U_i computes $A_i = h(PW_i \parallel b_1)$, $PID_i = h(ID_i \parallel b_2)$, and $bb_i = b_2 \oplus A_i$ and sends $\{A_i, PID_i\}$ to CS via secure channel. Upon receiving the registration request, the CS computes $C_i = h(A_i \parallel PID_i)$, $D_i = h(PID_i \parallel x)$, and $E_i = D_i \oplus A_i$ with its secret key. Finally, CS replies U_i a smart card with $\{C_i, E_i, h(\cdot)\}$. After getting the smart card SC, U_i computes $DP = h(ID_i \parallel PW_i) \oplus b_1$ and records DP, bb_i into it.

3.1.2. Login and Authentication Phase. In order to access remote server resources, a legal user U_i inserts his/her smart card into a card reader and inputs ID_i, PW_i. Then the following steps are performed:

(1) SC computes $b_1 = DP \oplus h(ID_i \parallel PW_i)$, $A_i = h(PW_i \parallel b_1)$, $b_2 = bb_i \oplus A_i$, $PID_i = h(ID_i \parallel b_2)$, and $C_i^* = h(A_i \parallel PID_i)$ and verifies whether $C_i^* \overset{?}{=} C_i$. If so, SC selects a random number N_i and computes $D_i = E_i \oplus A_i$, $G_i = h(PID_i \parallel SID_j \parallel N_i \parallel T_i \parallel D_i)$, $F_i = D_i \oplus N_i$, and $Z_i = SID_j \oplus h(D_i \parallel N_i)$, where SID_j is S_j's identity chosen by U_i and T_i is the current timestamp. Otherwise, SC terminates the session.

(2) $U_i \longrightarrow S_j : \{G_i, F_i, Z_i, PID_i, T_i\}$.

(3) Upon receiving the login request, S_j checks whether $|T_j - T_i| < \Delta T$ holds, where T_j is S_j's current timestamp and ΔT is the expected valid time interval. If it does not hold, S_j rejects the connection. Otherwise, S_j produces a random number N_j and computes $J_i = BS_j \oplus N_j$, $K_i = h(N_j \parallel BS_j \parallel G_i \parallel T_j)$.

(4) $S_j \longrightarrow CS : \{J_i, K_i, P_j, G_i, F_i, Z_i, PID_i, T_i, T_j\}$.

(5) Once receiving the message from S_j, CS first checks the validity of time interval $|T_{cs} - T_j| < \Delta T$. If the verification holds, CS continues to compute $D_i = h(PID_i \parallel x)$, $N_i = F_i \oplus D_i$, $SID_j = Z_i \oplus h(D_i \parallel N_i)$, and $G_i^* = h(PID_i \parallel SID_j \parallel N_i \parallel T_i \parallel D_i)$ and checks whether $G_i^* \overset{?}{=} G_i$. If either of the above

verification fails, CS terminates the procedure. Otherwise, CS keeps on calculating $BS_j = h(P_j \parallel y)$, $N_j = BS_j \oplus J_i$, and $K_i^* = h(N_j \parallel BS_j \parallel G_i \parallel T_j)$ and verifies whether the computed K_i^* equals the received one. If not, CS aborts the session. Otherwise, CS chooses a random number N_{cs} and computes $P_{cs} = N_j \oplus N_{cs} \oplus h(N_i \parallel D_i)$, $R_{cs} = N_i \oplus N_{cs} \oplus h(BS_j \parallel N_j)$, $SK_{cs} = h(N_i \oplus N_j \oplus N_{cs})$, $Q_{cs} = h((N_j \oplus N_{cs}) \parallel SK_{cs})$, and $V_{cs} = h((N_i \parallel N_{cs}) \parallel SK_{cs})$.

(6) $CS \longrightarrow S_j : \{P_{cs}, R_{cs}, Q_{cs}, V_{cs}\}$.

(7) While receiving the message from CS, S_j computes $W_j = h(BS_j \parallel N_j)$, $N_i \oplus N_{cs} = R_{cs} \oplus W_j$, $SK_j = h(N_i \oplus N_j \oplus N_{cs})$, and $V_{cs}^* = h((N_i \parallel N_{cs}) \parallel SK_j)$ and checks the condition $V_{cs}^* \stackrel{?}{=} V_{cs}$ holds or not. If it does not hold, S_j terminates the connection. Otherwise, S_j sends $\{P_{cs}, Q_{cs}\}$ to U_i via public channel.

(8) After receiving the response message from S_j, U_i computes $L_i = h(N_i \parallel D_i)$, $N_j \oplus N_{cs} = P_{cs} \oplus L_i$, $SK_i = h(N_i \oplus N_j \oplus N_{cs})$, and $Q_{cs}^* = h((N_j \oplus N_{cs}) \parallel SK_{cs})$ and verifies whether $Q_{cs}^* \stackrel{?}{=} Q_{cs}$. If so, U_i successfully authenticates S_j and CS and establish a session key $SK_i = SK_j = SK_{cs}$.

3.2. Cryptanalysis of Amin et al.'s Scheme. The four assumptions made in Section 2.2 are also explicitly employed in Amin et al.'s work [8] when they analyze the security of Xue et al.'s scheme [37] and Chuang et al.'s scheme [38] and proof the safety of their scheme. Consequently, our following discussions will base on these four assumptions.

3.2.1. No Provision of User Anonymity. Nowadays, privacy concerns are attracting more and more attention among governments, organizations, and individuals, and anonymous privacy-preserving authentication protocols are of particular interest. This is because the violation of user anonymity, say the leakage of some user-specific (static) information, may facilitate a malicious adversary to track the victim's current activities and login history [7, 39]. Generally, there are two kinds of user anonymity attributes, basic and advanced [7]: (i) user ID protection, which means $\mathscr{A}$ cannot obtain the real ID of the user; (ii) user untraceability, which means $\mathscr{A}$ is unable to tell who the user is and distinguish whether two communications are coming from the same user. In wireless sensor networks, the latter notion has been widely adopted (e.g., [40–42]), so does Amin et al.'s scheme.

In 2014, Das et al. [43] firstly introduced a "dynamic ID technique" to achieve user anonymity: a user's real ID is concealed in the session-variant pseudonym identities. Subsequently, many schemes (e.g., [25, 44, 45]) follow this technique, which are so-called "dynamic ID" schemes, and Amin et al.'s scheme [8] falls into this category. However, after careful analysis, we find that Amin et al.'s scheme cannot achieve user anonymity in practice. To be specific, in the login phase of their scheme, Amin et al. try to compute a pseudonym identity $PID_i = h(ID_i \parallel b_2)$ as a dynamic identity. On one hand, PID_i is specific to the legitimate user U_i; on the other hand, PID_i is kept static and transmitted in plain of all the U_i's login messages $\{G_i, F_i, Z_i, PID_i, T_i\}$.

Accordingly, this specific value PID_i can be seen as U_i's "identification", and thus $\mathscr{A}$ can exploit it to identify and track U_i in the whole system. To conduct the aforementioned attack, an adversary $\mathscr{A}$ only needs to eavesdrop the transmission channel without other contact operations and computations. This well serves to show the violation of user anonymity on Amin et al.'s scheme [8], thereby contradicting their claim.

3.2.2. Smart Card Loss Attack. Amin et al. [8] showed that, in Xue et al.'s protocol [37], users' passwords can be offline guessed once $\mathscr{A}$ has somehow obtained (lost or stolen) the victim's smart card and extracted the stored secret information. Then Amin et al. attempt to overcome this pitfall in their new proposed scheme. However, precisely the same deficiency still exits in Amin et al.'s enhanced version. Let us consider the following scenario, suppose that $\mathscr{A}$ has obtained the secret parameters $\{C_i, E_i, DP, bb_i, h(\cdot)\}$ stored in U_i's smart card (e.g., by side-channel attack [46–48] and reverse engineering technique [49]), which is reasonable under assumption (3). Then $\mathscr{A}$ can conduct the following procedure to guess U_i's password.

Step 1. Choose a pair of (ID_i^*, PW_i^*) from the identity space $\mathscr{D}_{id}$ and password space $\mathscr{D}_{pw}$.

Step 2. Compute $b_1 = DP \oplus h(ID_i^* \parallel PW_i^*)$, $A_i = h(PW_i^* \parallel b_1)$, $b_2 = bb_i \oplus A_i$, $PID_i = h(ID_i^* \parallel b_2)$, and $C_i^* = h(A_i \parallel PID_i)$.

Step 3. Verify whether the computed C_i^* equals the extracted C_i.

Step 4. Repeat Steps 1, 2, and 3 until finding the correct values.

Let $|\mathscr{D}_{id}|$ and $|\mathscr{D}_{pw}|$ denote the size of $\mathscr{D}_{id}$ and $\mathscr{D}_{pw}$, and the time complexity of the aforementioned attack is $\mathcal{O}(4T_H \times |\mathscr{D}_{id}| \times |\mathscr{D}_{pw}|)$, which is linearly associated with the running time of Hash operation and can be finished in a few days as the limited size of $|\mathscr{D}_{id}| \le |\mathscr{D}_{pw}| \le 10^6$ [29, 30].

Further, according to assumption (1), $\mathscr{A}$ is capable of eavesdropping and intercepting the normal (previous successful) login message $\{G_i, F_i, Z_i, PID_i, T_i\}$ between U_i and S_j over the public channel. It is fair to assume that $\mathscr{A}$ has already obtained the correct value of PID_i, then Step 2 might be changed to compute b_1, A_i, b_2, and PID_i^* and compared the computed PID_i^* with the intercepted PID_i in Step 3. In this way, the time complexity of the above procedure reduces to $\mathcal{O}(3T_H \times |\mathscr{D}_{id}| \times |\mathscr{D}_{pw}|)$, where the exclusive and concatenation operations are too small to overlook.

Note that both of the above two attacks are carried out *offline* without any interaction with the control server. Hence, there is no way for CS to find abnormality and the adversary $\mathscr{A}$ can impersonate U_i at anytime until CS revokes the victim's smart card. All in all, our analysis demonstrates the feasibility of smart card loss attack on Amin et al.'s scheme [8].

3.2.3. The Violation of Forward Secrecy. As mentioned in Section 2.2.2, Amin et al.'s scheme [8] also subjects to node

capture attack. In such cases, the captured nodes may enable an adversary to compromise communications between other noncaptured nodes or obtain previous session keys. We will show this pitfall in this subsection. Assume that a malicious adversary $\mathcal{A}$ has compromised a cloud server S_k and extracted the secret parameters $\{BS_j, d\}$ stored in its memory, $\mathcal{A}$ can recover the previous session key as follows.

Step 1. Intercept the message $\{J_i, K_i, P_k, G_i, F_i, Z_i, PID_i, T_i, T_k\}$ sending from S_k to CS.

Step 2. Compute $N_k = BS_k \oplus J_i$, where BS_k is extracted from the compromised node S_k.

Step 3. Intercept the message $\{P_{cs}, R_{cs}, Q_{cs}, V_{cs}\}$ sending from CS to S_k.

Step 4. Compute $N_i \oplus N_{cs} = R_{cs} \oplus h(BS_k \parallel N_k)$.

Step 5. Compute the session key $SK = h(N_k \oplus N_i \oplus N_{cs})$.

In light of N_k and $N_i \oplus N_{cs}$ which are all correct values, $\mathcal{A}$ manages to find the previous session key. Hence, the desirable property of forward secrecy can not be attained by Amin et al.'s scheme [8]. Similar to Li et al.'s scheme [25], this also due to the violation of "PFS principle". Except the ECC technique mentioned before, we suggest this issue to be well addressed by introducing another high-efficiency technique, i.e., Chebyshev polynomials semigroup property (so-called chaotic maps).

For this property, given x, $T_u(x)$, and $T_v(x)$, it is intractable to find $T_u T_v(x)$, where x is a variable and u, v denote the integer degree [45]. Assume the control server chooses and writes a variable value x in each user's smart card in the registration phase. Then we slightly modify the random numbers N_i to be $T_{N_i}(x)$ and N_k to be $T_{N_k}(x)$, and thus the session key can be calculated as $SK = h(T_{N_i}(x) \parallel T_{N_k}(x) \parallel T_{N_i} T_{N_k}(x))$. For higher security, it is better to involve other secret parameters such as ID_i, SID_j. In this way, the improvement of Amin et al.'s scheme [8] can achieve perfect forward secrecy based on computational Diffie-Hellman problem.

3.2.4. Mistakes in the Proof. Similarly, the security proof in Amin et al.'s scheme [8] does not capture realistic security threat. There are three main reasons: (1) The error of initial hypothesis. In the formal proof of Amin et al.'s scheme [8], they make an assumption A11: $S_j| \equiv U_i \overset{SK}{\longleftrightarrow} S_j$, which is the same as **Goal 3**. This demonstrates that the proof of **Goal 3** is not necessary. (2) The wrong usage of logic rules. We take Step S2 as an example. This step is based on the message meaning rule and derives that S_j believes U_i said N_i from A11 and S1. However, according to the message meaning rule, we cannot obtain this conclusion from A11. Hence, A11 should be changed to $S_j| \equiv U_i \overset{D_i}{\longleftrightarrow} S_j$. (3) Using undefined new rules. Amin et al. [8] also employ a new *session keys rule*, but they did not give a definition of the new rule.

4. Cryptanalysis of Wu et al.'s Scheme

In this section, we will review and analyze Wu et al.'s scheme [9], which is a lightweight and relatively robust two-factor authentication scheme for wireless medical sensor networks. In [9], Wu et al. have found some security pitfalls in historical schemes and attempted to overcome all these flaws in the new proposed one. Besides, Wu et al. [9] use NS-3, a simulation tool to prove the security of their proposed protocol. Note that, the simulation process can only prove the validity of their protocol, including the viable communication between the sensor node and the user, the probable communication time, system size, etc. However, it can not prove whether their protocol resists against various known attacks. In the following section, we find Wu et al.'s improved scheme still fails to attain the most important goal of forward secrecy and is prone to user friendliness issue.

4.1. Review of Wu et al.'s Scheme. This subsection briefly reviews Wu et al.'s [9] scheme, which involves four critical phases: registration, login, authentication and password change, and a previous initialization. We simplify initialization phase in the registration phase.

4.1.1. Registration Phase. Initially, GWN is equipped with an identity GID and its own secret key G. The registration phase is further divided into sensor node registration and user registration.

Sensor Node Registration Phase. Each sensor node S_j chooses an identity SID_j and sends to GWN via a secure channel. Then GWN decides to deploy it in a sensor set numbered N_l and computes the secret key $SG = h(SID_j \parallel G \parallel N_l)$. Finally, $\{SID_j, SG, GID\}$ is injected to the memory of S_j and (SID_j, N_l) is stored in the database of GWN.

User Registration Phase. In this phase, U_i first selects an identity ID_i, a password PW_i, and a nonce r_0, and then

(1) U_i computes $HPW_i = h(PW_i \parallel r_0)$;

(2) $U_i \Longrightarrow GWN$: $\{ID_i, HPW_i\}$;

(3) GWN checks if ID_i has already existed in the database. If so, it denies the registration request. Otherwise, GWN chooses a pseudoidentity CID_i and computes $A_1 = h(CID_i \parallel GID \parallel G) \oplus HPW_i$ and $A_2 = h(ID_i \parallel G) \oplus h(ID_i \parallel HPW_i)$ and then stores ID_i in database;

(4) $GWN \Longrightarrow U_i$: a smart card contains sensitive parameters $\{A_1, A_2, CID_i, GID\}$;

(5) after receiving the message, U_i computes $A_3 = h(ID_i \parallel PW_i) \oplus r_0$ and inserts it into SC.

4.1.2. Login and Authentication Phase. U_i conducts the following procedures to access sensitive information of the target sensor S_j:

(1) U_i inputs ID_i and PW_i to the smart card. Then SC computes $r_0 = A_3 \oplus h(ID_i \parallel PW_i)$ and $HPW_i = h(PW_i \parallel r_0)$. SC chooses a random number r_u and the required sensor node SID_j and further computes $B_1 = A_1 \oplus HPW_i = h(CID_i \parallel GID \parallel G)$, $B_2 = B_1 \oplus r_u$, $B_3 = ID_i \oplus h(r_u \parallel B_1)$, and $B_4 = h(CID_i \parallel GID \parallel SID_j \parallel B_1 \parallel ID_i \parallel r_u)$.

(2) $U_i \longrightarrow GWN$: $\{CID_i, GID, SID_j, B_2, B_3, B_4\}$.

(3) When receiving the message from U_i, GWN first checks if GID is correct. If so, GWN computes $B_1 = h(CID_i \parallel GID \parallel G)$, $r_u = B_1 \oplus B_2$, and $ID_i = B_3 \oplus h(r_u \parallel B_1)$ and verifies whether ID_i is in the database and $B_4 \stackrel{?}{=} h(CID_i \parallel GID \parallel SID_j \parallel B_1 \parallel ID_i \parallel r_u)$. If either of the two verifications does not hold, GWN will terminate the session. Otherwise, GWN searches (SID_j, N_l) from the database, generates a random number r_g, and computes $SG = h(SID_j \parallel G \parallel N_l)$, $B_5 = h(SG \parallel GID) \oplus r_u$, $B_6 = h(r_u) \oplus r_g$, and $B_7 = h(SG \parallel r_u \parallel r_g)$.

(4) $GWN \longrightarrow S_j : \{SID_j, B_5, B_6, B_7\}$.

(5) Once receiving the message, the corresponding node S_j checks if SID_j is correct and computes $r_u = B_5 \oplus h(SG \parallel GID)$ and $r_g = B_6 \oplus h(r_u)$. Then S_j verifies whether $B_7 \stackrel{?}{=} h(SG \parallel r_u \parallel r_g)$. If either is incorrect, S_j rejects the session. Otherwise, S_j generates r_s and computes $sk_s = h(r_u \parallel r_g \parallel r_s)$, $B_8 = h(SG \parallel r_g) \oplus r_s$, and $B_9 = h(sk_s \parallel SID_j \parallel GID \parallel r_s)$.

(6) $S_j \longrightarrow GWN$: $\{B_8, B_9\}$.

(7) Once received the response message, GWN computes $r_s = B_8 \oplus h(SG \parallel r_g)$, $sk_g = h(r_u \parallel r_g \parallel r_s)$ and checks whether $B_9 \stackrel{?}{=} h(sk_g \parallel SID_j \parallel GID \parallel r_s)$. If so, GWN chooses a new pseudoidentity CID_i^{new} and calculates $B_{10} = h(CID_i^{new} \parallel GID \parallel G) \oplus h(r_u \parallel CID_i)$, $B_{11} = h(r_u \parallel ID_i) \oplus r_g$, $B_{12} = h(r_u \parallel r_g) \oplus r_s$, $B_{13} = h(h(ID_i \parallel G) \parallel r_s) \oplus CID_i^{new}$, and $B_{14} = h(sk_g \parallel ID_i \parallel B_{10} \parallel CID_i^{new})$.

(8) $GWN \longrightarrow U_i : \{B_{10}, B_{11}, B_{12}, B_{13}, B_{14}\}$.

(9) When receiving the response message, SC computes $r_g = B_{11} \oplus h(r_u \parallel ID_i)$, $r_s = B_{12} \oplus h(r_u \parallel r_g)$, $sk_u = h(r_u \parallel r_g \parallel r_s)$, and $CID_i^{new} = B_{13} \oplus h((A_2 \oplus h(ID_i \parallel HPW_i)) \parallel r_s)$ and verifies whether $B_{14} \stackrel{?}{=} h(sk_u \parallel ID_i \parallel B_{10} \parallel CID_i^{new})$. If it is equal, SC computes $A_1^{new} = B_{10} \oplus h(r_u \parallel CID_i) \oplus HPW_i$ and replaces (A_1, CID_i) with (A_1^{new}, CID_i^{new}).

4.2. Cryptanalysis of Wu et al.'s Scheme. Due to its simplicity and admirable provision of user anonymity, Wu et al.'s scheme [9] exhibits great application prospects, and yet there are still some security pitfalls being overlooked by Wu et al. In the following, we will demonstrate that Wu et al.'s scheme [9] has some user friendliness issue and fails to achieve the critical property of forward security.

4.2.1. No Provision of User Friendliness. According to the collected data from Dashlane [56], "we are online hoarders" that the average user maintains over 107 accounts registered to one email address and this figure will rise to 207 by 2020. This statistical shows that users are creating and virtually stashing more online account information than ever, which leads to an insanely high number of accounts to manage. In that case, freely password change is a recommended practice, for users have to reset a forgotten password (an average of 37 accounts [56]) and the fixed password is definitely vulnerable. Moreover, users may make a slip in writing passwords or identities; the rapid response and decisive action are quite necessary for a user friendly authentication protocol.

Early in 1968, Robert Miller [57] published a classic paper about response time in man-computer conversational transactions, which pointed out that "response times exceed 10 seconds will completely lose the user's attention". In this way, locally secure password change, i.e., providing an explicit and secure process to verify the correctness of user-keyed password in smart card, is essential. That is, the smart card has no need to interact with remote server in user input and password changing phases. However, as stated above, both Li et al.'s scheme [25] and Amin et al.'s scheme [8] provide local password change, but their strategies introduce new vulnerabilities-offline dictionary attack.

Back to Wu et al.'s scheme [9], there is no verifier in the smart card, which means their scheme even cannot provide timely detection mechanism and reasonable password change. Fortunately, Wang et al. [10] introduced a "fuzzy verifier" technique to effectively solve this security-usability issue. In the following, we will take Wu et al.'s scheme [9] as an example to show this strategy. Firstly, U_i submits $\{ID_i, HPW_i\}$ to GWN in the registration phase. Then GWN computes $A_i = h((h(ID_i) \oplus HPW_i)) \bmod n)$ and stores it in U_i's smart card, where n denotes the size of (ID, PW) pool and $2^6 \le n \le 2^8$. Assume $|\mathscr{D}_{id}| \approx |\mathscr{D}_{pw}| \approx 10^6$ and $n = 2^8$ [29, 30], we can be assured that there have the possibilities of $(|\mathscr{D}_{id}| \times |\mathscr{D}_{id}|)/n \approx 2^{32}$ identity and password pairs to thwart the adversary from guessing out the correct password.

The same considerations can also be applied to Li et al.'s scheme [25] and Amin et al.'s scheme [8]. The large-scale candidates will effectively frustrate $\mathscr{A}$ from random guessing the password by a brute force method as well as providing a timely detection of the mistyped identity or password.

4.2.2. The Violation of Forward Secrecy. Forward secrecy is an important property, for the unattended environment and security-critical applications in wireless sensor networks [7, 11]. In [9], Wu et al. explicitly stated that "the sensor nodes may be captured by the intruder", which accords with assumption (4) made in Section 2.2. Under this statement, we find that Wu et al.'s scheme cannot achieve the forward secrecy. Once a sensor node S_k has been compromised, the stored information SG might be obtained by $\mathscr{A}$ and the following attacks can be launched.

Step 1. Intercept the message $\{CID_i, GID, SID_k, B_2, B_3, B_4\}$ sending from U_i to GWN and the message $\{SID_k, B_5, B_6, B_7\}$ sending from GWN to S_k.

Step 2. Compute $r_u = B_5 \oplus h(SG \parallel GID)$, $r_g = B_6 \oplus h(r_u)$, where SG is extracted from the compromised node S_k.

Step 3. Intercept the message $\{B_8, B_9\}$ sending from GWN to S_k.

Step 4. Compute $r_s = B_8 \oplus h(SG \parallel r_g)$.

Step 5. Compute the session key $= h(r_u \parallel r_g \parallel r_s)$.

The above attack demonstrates that once a sensor node S_k has been captured, the previous sessions might be decoded. This is the same failure with Li et al.'s scheme [25] and Amin et al.'s scheme [8]. Besides the above two techniques

Table 2: Performance and security comparison.

Related schemes	Computation cost on log-auth phases	PFS	Key technology			
			Hash	ECC	CM*	RSA
Li et al. (2018) [25]	$21T_H + 3T_E$	×	√	√	×	×
Amin et al. (2018) [8]	$22T_H$	×	√	×	×	×
Wu et al. (2017) [9]	$34T_H$	×	√	×	×	×
Roy et al. (2017) [45]	$15T_H + 3T_C$	×	√	×	√	×
Moon et al. (2017) [50]	$11T_H + 2T_C$	×	√	×	√	×
Srinivas et al. (2017) [51]	$29T_H$	×	√	×	×	×
Das et al. (2016) [52]	$31T_H + 4T_E$	√	√	√	×	×
Chang et al. (2016) [53]	$20T_H$	×	√	×	×	×
Vaidya et al. (2016) [54]	$29T_H$	×	√	×	×	×
Odelu et al. (2015) [55]	$15T_H + 6T_M$	×	√	×	×	√

Note[1]. T_H: one-time hash operation time; T_E: elliptic curve point multiplication computation time; T_C: running time of chaotic maps; T_M: time for modular multiplication/division.
Note[2]. × means the corresponding scheme fails to achieve this property; CM denotes chaotic maps; PFS denotes perfect forward secrecy.

(ECC cryptosystem and chaotic maps), we also suggest employing some other public-key cryptography techniques, such as Pairing [58] and RSA cryptosystem. Note that when using RSA cryptosystem to achieve forward secrecy, a new temporary RSA key must be generated by user side for each session [59].

To demonstrate the necessity and effectiveness of our suggestions, we provide a comparison of 10 recently proposed schemes by assessing whether they achieve forward secrecy and what main technology do they use. The result are shown in Table 2. One can see that only Das et al.'s scheme [52] successfully provides forward secrecy. This failure is mainly due to the fact that half of them (i.e., [8, 9, 51, 54]) only use Hash operation that are virtually impossible to provide forward secrecy ("PFS principle" [26]), yet the other 4 schemes (i.e., [25, 45, 50, 55]) that make use of public-key techniques (e.g., ECC, Chaotic maps, RSA) violate the principle that the random numbers must be generated by U_i and S_j separately and cannot be transmitted over the public channel.

5. Conclusion

In this paper, we first analyze three state-of-the-art authentication schemes presented by Li et al., Amin et al., and Wu et al., which are mainly applied to realize real-time data access for security-critical wireless sensor networks. We demonstrate that although their schemes are equipped with formal proof, they still suffer from smart card loss attack and fail to achieve some important properties of forward secrecy, user anonymity, and user friendliness. Our cryptanalysis results discourage the practical application of these three schemes and reveal some challenges in designing a robust scheme for WSNs. We then suggest several possible countermeasures on account of their weaknesses and provide a comparison of 10 representative schemes in terms of forward secrecy and key technology to demonstrate the necessity of our suggestions. For the future work, a natural direction is to employ our recommended technologies and countermeasures to design robust and efficient schemes for WSNs.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors have declared that no conflicts of interest exist.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (NSFC) under Grants no. 61472016 and no. 61772548, the National Key R&D Program of China under Grants no. 2016YFB0800603 and no. 2017YFB1200700, and the Foundation of Science and Technology on Information Assurance Laboratory No. KJ-17-001.

References

[1] K.-A. Shim, "BASIS: a practical multi-user broadcast authentication scheme in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1545–1554, 2017.

[2] J.-H. Bang, Y.-J. Cho, and K. Kang, "Anomaly detection of network-initiated LTE signaling traffic in wireless sensor and actuator networks based on a hidden semi-Markov model," *Computers & Security*, vol. 65, pp. 108–120, 2017.

[3] 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.

[4] E. C. Whitman, "Sosus the 'secret weapon' of undersea surveillance," *Undersea Warfare*, vol. 7, no. 2, 2005.

[5] C. M. Talbot, M. A. Temple, T. J. Carbino, and J. A. Betances, "Detecting rogue attacks on commercial wireless Insteon home automation systems," *Computers & Security*, vol. 74, pp. 296–307, 2017.

[6] X. Li, M. H. Ibrahim, S. Kumari, A. K. Sangaiah, V. Gupta, and 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.

[7] D. Wang and P. Wang, "On the anonymity of two-factor authentication schemes for wireless sensor networks: attacks, principle and solutions," *Computer Networks*, vol. 73, pp. 41–57, 2014.

[8] 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.

[9] 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*, 2017.

[10] 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*, 2016.

[11] 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.

[12] X. Huang, X. Chen, J. Li, Y. Xiang, and L. Xu, "Further observations on smart-card-based password-authenticated key agreement in distributed systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 7, pp. 1767–1775, 2014.

[13] K. H. M. Wong, Z. Yuan, C. Jiannong, and W. Shengwei, "A dynamic user authentication scheme for wireless sensor networks," in *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, vol. 1, pp. 244–251, Taichung, Taiwan, June 2006.

[14] H. R. Tseng, R. H. Jan, and W. Yang, "An improved dynamic user authentication scheme for wireless sensor networks," in *Proceedings of the 50th Annual IEEE Global Telecommunications Conference (GLOBECOM '07)*, pp. 986–990, November 2007.

[15] M. L. Das, "Two-factor user authentication in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1086–1090, 2009.

[16] M. K. Khan, S. K. Kim, and K. Alghathbar, "Cryptanalysis and security enhancement of a more efficient secure dynamic id-based remote user authentication scheme," *Computer Communications*, vol. 34, no. 3, pp. 305–309, 2011.

[17] T.-H. Chen and W.-K. Shih, "A robust mutual authentication protocol for wireless sensor networks," *ETRI Journal*, vol. 32, no. 5, pp. 704–712, 2010.

[18] H.-L. Yeh, T.-H. Chen, P.-C. Liu, T.-H. Kim, and H.-W. Wei, "A secured authentication protocol for wireless sensor networks using elliptic curves cryptography," *Sensors*, vol. 11, no. 5, pp. 4767–4779, 2011.

[19] W. B. Shi and P. Gong, "A new user authentication protocol for wireless sensor networks using elliptic curves cryptography," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 730831, 7 pages, 2013.

[20] B. Vaidya, D. Makrakis, and H. T. Mouftah, "Improved two-factor user authentication in wireless sensor networks," in *Proceedings of the 6th Annual IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob '10)*, pp. 600–606, October 2010.

[21] M. S. Farash, M. Turkanović, S. Kumari, and M. Hölbl, "An efficient user authentication and key agreement scheme for heterogeneous wireless sensor network tailored for the Internet of Things environment," *Ad Hoc Networks*, vol. 36, pp. 152–176, 2016.

[22] Y. Choi, D. Lee, J. Kim, J. Jung, J. Nam, and D. Won, "Security enhanced user authentication protocol for wireless sensor networks using elliptic curves cryptography," *Sensors*, vol. 14, no. 6, pp. 10081–10106, 2014.

[23] K. Xue, C. Ma, P. Hong, and R. Ding, "A temporal-credential-based mutual authentication and key agreement scheme for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 316–323, 2013.

[24] D. He, N. Kumar, and N. Chilamkurti, "A secure temporal-credential-based mutual authentication and key agreement scheme with pseudo identity for wireless sensor networks," *Information Sciences*, vol. 321, Article ID 11403, pp. 263–277, 2015.

[25] X. Li, J. Niu, S. Kumari, F. Wu, A. K. Sangaiah, and K. R. Choo, "A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments," *Journal of Network and Computer Applications*, vol. 103, pp. 194–204, 2018.

[26] C.-G. Ma, D. Wang, and S.-D. Zhao, "Security flaws in two improved remote user authentication schemes using smart cards," *International Journal of Communication Systems*, vol. 27, no. 10, pp. 2215–2227, 2014.

[27] J. Zhao, "On resilience and connectivity of secure wireless sensor networks under node capture attacks," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 557–571, 2017.

[28] X. Huang, Y. Xiang, E. Bertino, J. Zhou, and L. Xu, "Robust multi-factor authentication for fragile communications," *IEEE Transactions on Dependable and Secure Computing*, vol. 11, no. 6, pp. 568–581, 2014.

[29] D. Wang, Z. Zhang, P. Wang, J. Yan, and X. Huang, "Targeted online password guessing: an underestimated threat," in *Proceedings of the 23rd ACM Conference on Computer and Communications Security (CCS '16)*, pp. 1242–1254, October 2016.

[30] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, "Zipfs law in passwords," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

[31] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proceedings of the 6th ACM Conference on Computer and Communications Security (ACM CCS '99)*, pp. 28–36, November 1999.

[32] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS '02)*, pp. 41–47, Washington, DC, USA, November 2002.

[33] T. M. Vu, R. Safavi-Naini, and C. Williamson, "Securing wireless sensor networks against large-scale node capture attacks," in *Proceedings of the 5th ACM Symposium on Information, Computer and Communication Security (ASIACCS '10)*, pp. 112–123, April 2010.

[34] M. Turkanović, B. Brumen, and M. Hölbl, "A novel user authentication and key agreement scheme for heterogeneous ad hoc wireless sensor networks, based on the Internet of Things notion," *Ad Hoc Networks*, vol. 20, pp. 96–112, 2014.

[35] M. Burrows, M. Abad, and M. Needham, "A logic of authentication," *Proceedings of the Royal Society A Mathematical, Physical and Engineering Sciences*, vol. 426, no. 1871, pp. 233–271, 1989.

[36] V. Odelu, A. K. Das, and A. Goswami, "A Secure Biometrics-Based Multi-Server Authentication Protocol Using Smart

Cards," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1953–1966, 2015.

[37] K. Xue, P. Hong, and C. Ma, "A lightweight dynamic pseudonym identity based authentication and key agreement protocol without verification tables for multi-server architecture," *Journal of Computer and System Sciences*, vol. 80, no. 1, pp. 195–206, 2014.

[38] M.-C. Chuang and M. C. Chen, "An anonymous multi-server authenticated key agreement scheme based on trust computing using smart cards and biometrics," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1411–1418, 2014.

[39] C. Tang and D. O. Wu, "Mobile privacy in wireless networks-revisited," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 1035–1042, 2008.

[40] R. Ali, A. K. Pal, S. Kumari, M. Karuppiah, and M. Conti, "A secure user authentication and key-agreement scheme using wireless sensor networks for agriculture monitoring," *Future Generation Computer Systems*, vol. 84, pp. 200–215, 2018.

[41] J. Jung, J. Moon, D. Lee, and D. Won, "Efficient and security enhanced anonymous authentication with key agreement scheme in wireless sensor networks," *Sensors*, vol. 17, no. 3, p. 644, 2017.

[42] M. Wazid, A. K. Das, V. Odelu, N. Kumar, M. Conti, and M. Jo, "Design of secure user authenticated key management protocol for generic iot networks," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 269–282, 2018.

[43] M. L. Das, A. Saxena, and V. P. Gulati, "A dynamic ID-based remote user authentication scheme," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 2, pp. 629–631, 2004.

[44] P. K. Dhillon and S. Kalra, "Secure multi-factor remote user authentication scheme for internet of things environments," *International Journal of Communication Systems*, vol. 30, no. 16, pp. 1–20, 2017.

[45] S. Roy, S. Chatterjee, A. K. Das, S. Chattopadhyay, S. Kumari, and M. Jo, "Chaotic map-based anonymous user authentication scheme with user biometrics and fuzzy extractor for crowdsourcing internet of things," *IEEE Internet of Things Journal*, 2017.

[46] Y. Xiao, M. Li, S. Chen, and Y. Zhang, "Stacco: differentially analyzing side-channel traces for detecting ssl/tls vulnerabilities in secure enclaves," in *Proceedings of the ACM SIGSAC Conference*, pp. 859–874, Dallas, TX, USA, October 2017.

[47] N. Veyrat-Charvillon and F. Standaert, "Generic side-channel distinguishers: improvements and limitations," in *Advances in Cryptology—CRYPTO 2011*, vol. 6841 of *Lecture Notes in Computer Science*, pp. 354–372, Springer Berlin Heidelberg, 2011.

[48] Y. Zhou, Y. Yu, F.-X. Standaert, and J.-J. Quisquater, "On the need of physical security for small embedded devices: a case study with COMP128-1 implementations in SIM cards," in *Proceedings of the FC*, vol. 7859, pp. 230–238, 2013.

[49] G. Chalupar, S. Peherstorfer, E. Poll, and J. De Ruiter, "Automated reverse engineering using lego," in *Proceedings of the 8th USENIX WOOT*, vol. 14, pp. 1–10, 2014.

[50] J. Moon, Y. Lee, J. Kim, and D. Won, "Improving an anonymous and provably secure authentication protocol for a mobile user," *Security and Communication Networks*, vol. 2017, Article ID 1378128, 13 pages, 2017.

[51] J. Srinivas, S. Mukhopadhyay, and D. Mishra, "Secure and efficient user authentication scheme for multi-gateway wireless sensor networks," *Ad Hoc Networks*, vol. 54, pp. 147–169, 2017.

[52] A. K. Das, S. Kumari, V. Odelu, X. Li, F. Wu, and X. Huang, "Provably secure user authentication and key agreement scheme for wireless sensor networks," *Security and Communication Networks*, vol. 9, no. 16, pp. 3670–3687, 2016.

[53] C.-C. Chang and H.-D. Le, "A provably secure, efficient, and flexible authentication scheme for ad hoc wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 357–366, 2016.

[54] B. Vaidya, D. Makrakis, and H. Mouftah, "Two-factor mutual authentication with key agreement in wireless sensor networks," *Security and Communication Networks*, vol. 9, no. 2, pp. 171–183, 2016.

[55] V. Odelu, A. K. Das, and A. Goswami, "An effective and robust secure remote user authenticated key agreement scheme using smart cards in wireless communication systems," *Wireless Personal Communications*, vol. 84, no. 4, pp. 2571–2598, 2015.

[56] T. L. Bras, "Online overload its worse than you thought," July 2015, https://bit.ly/2IjgkGL.

[57] R. B. Miller, "Response time in man-computer conversational transactions," in *Proceedings of the Fall Joint Computer Conference (AFIPS '68)*, pp. 267–277, San Francisco, Calif, USA, December 1968.

[58] D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing," in *Advances in Cryptology—CRYPTO 2001*, vol. 2139 of *Lecture Notes in Computer Science*, pp. 213–229, Springer, 2001.

[59] D. Park, C. Boyd, and S. Moon, "Forward secrecy and its application to future mobile communications security," in *Public Key Cryptography*, vol. 1751 of *Lecture Notes in Computer Science*, pp. 433–445, Springer Berlin Heidelberg, 2000.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 1640167, 15 pages
https://doi.org/10.1155/2018/1640167

Review Article

Lightweight Cryptographic Techniques for Automotive Cybersecurity

Ahmer Khan Jadoon[1], **Licheng Wang**[1], **Tong Li**[2] and **Muhammad Azam Zia**[3]

[1]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*School of Computer Science, Guangzhou University, China*
[3]*Department of Computer Science, University of Agriculture Faisalabad, Pakistan*

Correspondence should be addressed to Tong Li; litongziyi@mail.nankai.edu.cn

Received 10 March 2018; Accepted 24 May 2018; Published 26 June 2018

Academic Editor: Joseph Liu

A new integration of wireless communication technologies into the automobile industry has instigated a momentous research interest in the field of Vehicular Ad Hoc Network (VANET) security. Intelligent Transportation Systems (ITS) are set up, aiming to offer promising applications for efficient and safe communication for future automotive technology. Vehicular networks are unique in terms of characteristics, challenges, architecture, and applications. Consequently, security requirements related to vehicular networks are more complex as compared to mobile networks and conventional wireless networks. This article presents a survey about developments in vehicular networks from the perspective of lightweight cryptographic protocols and privacy preserving algorithms. Unique characteristics of vehicular networks are presented which make the embedded security applications computationally hard as well as memory constrained. The current study also deals with the fundamental security requirements, essential for vehicular communication. Furthermore, awareness of security threats and their cryptographic solutions in terms of future automotive industry are discussed. In addition, asymmetric, symmetric, and lightweight cryptographic solutions are summarized. These strategies can be enhanced or incorporated all in all to meet the security perquisites of future cars security.

1. Introduction

There has been a tremendous increase in the number of vehicles compared to the number of roads. This situation leads to many challenges like heavy traffic jams, economy, pollution, and many other issues related to efficiency and safety of transportation systems. Many initiatives have already been taken in response to these challenges in order to overcome the situation. For this scenario, utilization of wireless technology in vehicular networks makes a huge difference to overcome the traffic issues and reduce the chances of accidents or injuries. Intelligent transportation systems (ITS) [1] are developed, aiming to improve the efficiency and safety of transportation systems. This technology mainly relies on the information sharing and authentication of vehicles. Moreover, it makes them traceable to law enforcement authorities in case of overspeeding, crash or collection of tolls, etc. The authentication of vehicles can be performed through radio links, instead of conventional methods such as reading license plates. Vehicles also need to be authenticated by other vehicles and infrastructure for secure communication. Many service providing companies exchange information with vehicles to facilitate the use in terms of location services or other helpful applications. All these authentications are carried out by cryptographic algorithms to ensure the identity of sender and receiver.

A general vehicular network consists of three types of communications links, i.e., Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), and Infrastructure to Infrastructure (I2I) communication. All these links require being protected in order to insure the security of network. Vehicles are equipped with On-Board Units (OBUs) to communicate with each other and Road Side Units (RSUs). Validation and authentication of information exchange between the vehicles

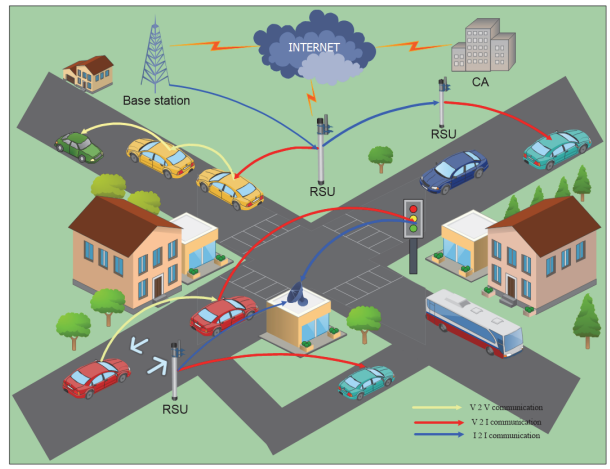


Figure 1: Vehicular network architecture.

are a key concern for the traffic safety. Furthermore, driver's privacy also needs to be considered (their details must be confidential from unauthorized entities) and their confidential information can only be accessed by a legitimate authority. The main goal is to achieve both anonymity and traceability at the same time [2]. Privacy in a vehicular network is more considerable as compared to mobile network, since a mobile phone can be switched off at any time but a license plate needs to be accessed by the law enforcement authorities all the time.

However, the protection against many malicious attacks like message suppression, denial of services, dropping down of packets from network, broadcast of false information, getting control over the network, and several other attacks is still unknown to manufacturers and suppliers. Conventional cryptographic algorithms such as public-key infrastructure (PKI), elliptic curve cryptography (ECC), HASH functions, and symmetric key cryptography may not be applied directly in vehicular networks due to their high mobility and dynamic network topology [3–5]. Vehicular networks require real-time response and cannot tolerate delay in communication. Therefore, the conventional protocols that are developed for traditional networks fail to provide high throughput performance, low latency, and reliability for vehicular networks. So, there is a need to implement secure lightweight cryptographic algorithms (as well as lightweight PKI [6]) on small embedded devices at acceptable execution time. Recently, scholars focus on developing the lightweight cryptographic algorithms and key generations schemes which may provide security for vehicular networks with high performance efficiently.

Primary target of this article is to give an overview on advancements of vehicular systems insight into the lightweight cryptographic conventions and security protecting calculations. The public acceptance for new technology in vehicular networks can only be ensured by optimizing the security and privacy of users. Moreover, awareness of security threats regarding the malicious attacks in future automotive industry should be known by the users and manufacturers. Security concerns for the future automotive industry are hurdle in the way of extensive deployment of vehicular networks commercially. Furthermore, we provide knowledge about resource constraints and challenges during the implementation of cryptographic algorithms for vehicular networks. We also present suggestions and lightweight cryptographic solutions to overcome the problems in future automotive industry.

The paper is organized as follows: In Section 2, we present the architecture of vehicular networks in sight of characteristics and security requirements. In Section 3, we discuss the security attacks on vehicular systems. Lightweight cryptographic protocols for vehicular networks are characterized in Section 4. Finally, Section 5 presents concluding remarks.

2. Architecture of Vehicular Ad Hoc Network

VANETs consist of the two basic wireless terminals, namely, On-Board Unit (OBU) and Road Side Unit (RSU) [1, 7–9]. The OBUs are embedded wireless devices installed in vehicles to communicate with RSUs and other OBUs. While RSUs are located at important points alongside the road or infrastructure and represent the wireless access points for communication. Each terminal acts like a node that can receive and relay messages within a wireless network. These nodes function as a router to other nodes in the network as shown in Figure 1. There is also interroadside communication of these access points with each other or with other devices. For instance, traffic lights may communicate with each other or RSU may communicate with cellular base stations etc. IEEE 802.11p Dedicated Short Range Communication (DSRC) has been selected as a standard for V2V and V2I communication in order to provide high data transfer with low latency [9]. It works in 5.9 GHz frequency band with 75 MHz bandwidth and has 300-1000 m range with several vehicle velocities in different environments [9, 10].

The existing Wireless Access in Vehicular Environments (WAVE) uses DSRC protocol to broadcast the services provided at RSUs [3]. Basically short messages contain vital

information about location, speed, and direction as well as emergency information with respect to airbag deployment, accident report, emergency brakes, etc. However, current approach of broadcast mechanism can result in network traffic congestion due to the insignificant usage of network resources. This issue can be resolved by tracking the addresses of OBUs and their connections with respective RSUs to perform efficient mobility management. There have been many suggestions related to mobility management in WAVE.

Chun et al. proposed two types of mobility management schemes, i.e., location estimation-based mobility management (LEMM) and basic mobility management (BMM) for WAVE services [11]. In LEMM scheme, positioning systems (e.g., GPS) are utilized to determine the location of OBUs in a fast moving vehicle, whereas, in BMM scheme, all RSUs are divided into different location areas which can determine the locations of OBUs by their MAC addresses. Torrent-Moreno et al. present a scheme for congestion mitigation based on distributed and fair transmit power-control [12]. Tielert et al. proposed a message-rate controller which uses disseminating congestion information over multiple hops to achieve global fairness [13]. Most of the schemes proposed a common feature for congestion control in the literary work and their goal was to attain unweighted fair sharing of the scarce channel resource to all vehicles [12–14]. Recently, Xu et al. proposed a lightweight scheme known as Dynamic Fully Homomorphic encryption-based Merkle Tree (FHMT) for lightweight streaming authenticated data structures which can be adopted in vehicular networks as a congestion mitigation technique. By leveraging the computing capability of fully homomorphic encryption, FHMT shifts almost all of the computation tasks to the server, reaching nearly no overhead for the client [15]. These schemes are significant to vehicular security requirements; however DSRC requires cryptographic protocols for authentication and authorization purposes which can result in network congestion. Therefore, lightweight cryptographic algorithms should be the first choice to ensure efficient security in automotive technology.

2.1. VANET Characteristics. VANETs use ad hoc approach to execute the wireless communication. The combination of properties of both wireless medium and ad hoc approach is generally defined as characteristics of VANET which makes it unique. We list some of the unique characteristics of VANETs as follows:

(1) **High mobility:** Mobility in VANETs is relatively higher as compared to MANETs. Generally, each node moves at higher speed in VANET. Therefore, network's communication time is reduced due to high mobility of the nodes [16, 17].

(2) **Time critical data exchange:** In VANET, the transfer of information to legitimate nodes should be reached within a specific time limit in order to execute rapid actions based on decisions made by the node.

(3) **Dynamic network topology:** The high mobility of vehicles makes the VANET topology irregular. Rapid changes in topology make the vehicular network vulnerable to attacks. Under such conditions, malicious vehicles are quite hard to detect.

(4) **Unbounded network density:** In VANETs, the density of network mainly relies on number of vehicles that may be high in traffic jams and low in suburban and rural areas. There is no bound to the number of vehicles joining the network.

(5) **Frequent disconnections:** Vehicles mostly use wireless medium to communicate in VANET, so frequent disconnection may occur due to high density of the vehicles or worse weather conditions.

(6) **Wireless medium:** Since only transmission medium that can be used in VANETs is wireless medium, therefore, the transmission of data should be anonymous. If the medium of transmission is not properly protected, then security of whole network can be jeopardized by using the same operating frequency [18].

(7) **Power constraints:** As compared to MANETs, the vehicular nodes do not experience power issues because of an uninterrupted power supply which can be arranged for OBUs by using long life battery.

(8) **Limited power transmission:** The architecture of wireless access of vehicles (WAVE) supports maximum range of 0 to 28.8 dBm for transmission power and associated coverage of distance range from 10 m to 1 km. So, coverage area distance is limited due to limitation of transmission power [9].

(9) **Wireless transmission limitations:** The factors like reflection, scattering diffraction, and refraction present in the urban areas makes the performance of DSRC wireless communication limited [19].

(10) **Computing capacity and energy storage:** The energy or storage breakdown problems are not present in VANETs. However, processing of very large amount of information is required due to huge scaling environment which becomes certainly a big challenge.

2.2. VANET Security Requirements. The main objective of VANET is to provide the comfort and safety to the driver as well as passengers. Communication between OBUs and RSUs can be employed to realize the active safety services like collision warnings, active navigation systems, real-time traffic information or weather information, etc. Facilities like multimedia or Internet connectivity are provided in the wireless coverage of a car. VANETs also include automatic parking payments and electronic toll collections. To ensure the efficient working of all these applications and services, a network needs to authenticate every message sent or received by the nodes. A small error or attack may result in a big damage for the safety and security of public. Certain security requirements for the V2V and V2I communication links in a basic vehicular network are listed as follows:

(1) **Message authentication and integrity:** Message authentication is the fundamental part of vehicular

Table 1: Attacks on VANET and their impact on security requirements.

Security Requirements	Attacks	Reference (Security Requirements/Attacks)
Message authentication and integrity	Sybil/Impersonation/Replay attacks	[20–23]/[46, 47, 53]
Availability	DoS/Sybil/Bogus information/Routing attacks	[24–26]/[46, 49, 50]
Confidentiality	Sybil/Impersonation attacks	[20]/[46–48]
Non-Repudiation	Impersonation attacks	[24, 25]/[48]
Privacy	Impersonation/Location Trailing/Eavesdropping attacks	[28, 29]/[41–45]

security. It ensures that each received message arrives in the same condition as it was sent out by the sender. Moreover, ID, location, and property of a sender must be authenticated and it is made sure that legitimate sender transmits reliable information [20]. Integrity check allows the receiver to verify if there has been any kind of fabrication or modification within the duration in which message was sent and received. We can find related work on message authentication and data integrity in literature [21–23].

(2) **Availability:** Availability of information is directly related to the efficiency of vehicular network. It ensures that network resources such as session key and applications must be available to legitimate nodes in a certain period of time without affecting operation of the network even in the presence of faults or malicious nodes [24, 25]. A number of multipath algorithms have been proposed to transfer information via multiple disjointed paths in order to reduce the chances of transmission breaks as an effect of a path failure. Ad hoc On-demand Distance Vector Multipath (AODVM) [26] and Ad hoc On-demand Multipath Distance Vector (AOMDV) [27] are extensions to the general Ad hoc On-demand Distance Vector (AODV) routing protocol.

(3) **Confidentiality:** All drivers private information has to be confined. This security prerequisite is to ensure that confidential information will only be read by permitted users. Requirement of confidentiality is needed in group communications, where only authorized group members are allowed to read such data. Confidentiality is considered a security issue when some message contains sensitive information like session key or toll payment data, etc. [20].

(4) **Access Control:** The security mechanism must guarantee that only authorized users can access the ad hoc network resources and information provided by the certificate authority. Access control provides protection against malicious vehicle to access unauthorized services and sensitive information of certificate authority. These messages must be encrypted using cryptographic encryption techniques.

(5) **Nonrepudiation:** Nonrepudiation is a service that requires a vehicle sending a safety message to other vehicles which cannot deny having sent message [24, 25]. This requirement is important as in case of any dispute a user of the vehicle shall not deny its fault.

(6) **Privacy:** Unauthorized node should not be able to access personal information of a driver. While the information in a vehicular network is broadcast publicly, there is a big threat to privacy. An adversary can collect and analyze this information to harm the users. An eavesdropper should not have the ability to distinguish two distinct information messages which came from same node [5, 20]. The fundamental concept of privacy preservation schemes in VANET is to periodically change the pseudonyms. There exists many schemes that have been proposed by researchers in which the concept of changing pseudonym is used to preserve the privacy of user [28–36]. Data owners often suppress their data for an untrusted trainer to train a classifier due to privacy concerns. Li et al. proposed a privacy preserving solution for learning algorithms based on differentially private naive Bayes learning, allowing a trainer to build a classifier over the data from a single owner [37]. Data privacy also becomes a central consideration in vehicular networks, where outsourcing data to the cloud server is done [38]. L-EncDB is a lightweight framework for privacy preserving scheme for efficient data outsourcing [39]. Recently HybridORAM [40] scheme is proposed which provides a better solution to securely outsource data to the cloud.

3. Cyberattacks in Automotive Technology

In this section, we present a summary of various attacks on vehicular networks, which can be found in the literature [41–53]. Some of these attacks are performed by the member nodes already registered with the network and called insider attacks. When a nonregistered node carries out an attack, it is known as outsider attack. These attacks can also be categorized as active and passive attacks. In an active attack, the attacker may generate new packets to damage the network or falsify the legitimate information, while the passive attacker can only eavesdrop the channel and acquires sensitive information. We categorize these attacks according to the violation of security services provided by vehicular networks. However, some of the attacks may violate more than one security service as shown in Table 1. The following are the common types of attacks which can be harmful to the security of vehicular network.

3.1. Attacks Related to Authentication. Attacks related to authentication are performed by unauthorized nodes entering

into the network and gaining access to the network privileges or claiming illegal authority. The most frequent attacks related to authentication of vehicles are summarized as follows:

(1) **Sybil attack:** In Sybil attack a node asserts itself as several nodes by simulating multiple identities [46, 47]. An attacker sends several messages with multiple identities and announces its various positions at the same time. Multiple copies of a node create confusion in the network and hence claim all the fake and illegal authority. Sybil attacks are harmful to the network topology and cause bandwidth consumption [10].

(2) **Impersonation attack:** In this type of attack, an attacker characterizes itself as an authorized node [48]. Objective of these attacks is to either gain access to the network privileges or to disturb the network. These attacks are potentially possible through possession of false attributes or identity theft.

(3) **Bogus information:** Attackers may send fake or bogus information to the system for their own advantage. For instance, an attacker sends bogus information of a heavy traffic jam due to an accident on a certain road to make its route clear. These attacks compromise the authentication requirement of vehicular network [20].

(4) **Session hijacking:** The attacker targets unique Session Identifier (SID) allocated for each new session and may get control over that session. An attacker gets edge of the fact that authentication at the network layer is done only once. No authentication is done after generation and allocation of the SID; therefore attackers get advantage of this feature [17].

(5) **Replay attacks:** The attacker impersonates itself as a legitimate vehicle or RSU to capture information packets and then sends out the replica of the captured signal to another node for its own benefits [53]. Replay attacks are considered threat to confidentiality and authenticity of the system.

(6) **GPS spoofing:** The Global Positioning System (GPS) Satellite stores geographical locations of vehicles and their identities in form of a location table. The attacker may alter these location table readings to mislead the vehicle. Signal simulators can be used by the attacker to generate signals stronger than the actual signals generated by satellite.

3.2. Attacks Related to Network Efficiency. Attacker may try to jam the network or produces delays in communication of vehicles which severely affects the performance and efficiency of vehicular network. Time is very critical issue in a vehicular network as a small delay can result into accidents or severe traffic issues. There is a need to apply antijamming techniques for better network efficiency [54]. Some of the common attacks related to efficiency and performance of vehicular network are described as follows:

(1) **Denial of service attacks:** DoS attacks can have severe effect on the efficiency and performance of vehicular network. The attack is performed by sending dummy messages to the network and making a victim node unavailable to other legitimate users by SYN flooding, jamming, or distributed DoS attack [49].

(2) **Routing attacks:** Routing attacks generally exploit the loopholes and vulnerability in routing protocols of a network. These attacks can be categorized as follows:

(i) **Black hole attack:** In this attack, malicious node first sends false route with lower hop count to attract the source node to send packet through itself. After source node sends data packet to the route, attacking node silently drops these packets [50].

(ii) **Gray hole attacks:** Similar to black hole attack, the compromised node drops packet but this dropping is performed only on selective packets. Selection is done according to requirement and intensions of attacker [50].

(iii) **Wormhole attack:** Two or more nodes work together to make tunnels within a network. The malicious node receives the packets and routes it to the other end of the tunnel. Through this tunneling process, hop count of the route decreases and the compromised nodes attract packets. In this way attacker node gets strong position than other deserving nodes in the network and thus it can carry out DoS attacks, replay attacks, etc.

(3) **Timing attacks:** In timing attack, the attacker node creates a delay in communication by altering time slot of the received packet. Due to this alteration, the neighbors of malicious node might not receive sensitive messages on time. In vehicular network, information is time critical with respect to its sensitivity and hence a small delay can result in accidents or severe traffic issues.

(4) **Intruder attack:** An unregistered node or application tries to enter the network in order to disturb the efficiency of network or gain false attributes. Intrusion detection systems (IDSs) are widely deployed in various networks in order to identify cyberthreats and possible incidents [55]. Li et al. proposed a malware detection system based on permission usage analysis by significant permission identification technique. 3 levels of pruning by mining the permission data are developed to identify the most significant permissions [56]. These recently proposed techniques can be incorporated with vehicular networks to mitigate intruder attacks.

3.3. Attacks Related to User's Privacy. Unauthorized nodes may attempt to access sensitive data from network and target the privacy of a legitimate user. Some common attacks on the user's privacy and confidentiality requirement in a vehicular network are given as follows:

(1) **Eavesdropping:** This type of attack is a risk to confidentiality of a network. The core objective of this attack is to get sensitive and confidential data for which the attacker is not authorized [24]. It is a passive attack in which an attacker sniffs the data silently to get the confidential information and further use it for his own benefits. Vehicular networks consist of relays that may be corrupted by multiple cochannel interferers, and the information transmitted from the relays to the destination can be overheard by the eavesdropper. Fan et al. investigate the impact of cochannel interference on the security performance of multiple amplify-and-forward (AF) relaying networks [57, 58].

(2) **Location trailing attack:** Location attacks generally target the privacy of a user in vehicular network by continuously tracking the location of a user. In this attack, position of the vehicle at a given moment or path trace along certain period of time can be used to map out the user [52, 59].

(3) **Identity revealing:** Attacker may try to reveal identity of vehicle's owner. As identity of the owner represents the driver, it can be latter used by the attacker for its own illegal benefits.

4. Cryptographic Techniques for Automotive Security

Safety has a long practice in history of automotive industry. Cryptography has played a key role in securing vehicular systems. Cryptography in vehicles was introduced in Remote Key-less Entry (RKE) in the middle of 1990s, which was followed by electronic immobilizers. We have a lot of solutions in isolated systems, such as single car. Developments in automotive technology such as connected cars and vehicular networks set up new security challenges. Although security in these networks depends more than just on cryptographic algorithms, still cryptographic schemes are the basic building blocks of security solutions in automotive industry. The embedded security applications in vehicular networks tend to be computationally hard and memory constrained due to their unique characteristics as described in Section 2.2.

We present an overview of the existing cryptographic schemes with respect to their complexity. Firstly, asymmetric cryptography is mainly used for digital signatures and key distribution over unsecured channels in vehicular networks. secondly, the symmetric algorithms are used for data encryption and message integrity checks. Recently there have been researches done on the lightweight cryptographic algorithms and dynamic key generation schemes are developed to secure vehicular networks.

4.1. Asymmetric or Public-Key Algorithms. Public-key infrastructure (PKI) based algorithms involve complex mathematical computations with large numbers and hard theoretical problems (commonly in the range of 1024-4048 bits), depending on the security level of selective algorithm. However, they provide advanced functions for data encryption and integrity check. Digital signatures and key distribution schemes are used for privacy preservation in unsecured channels. Asymmetric cryptographic techniques are projected in order to protect transmitted messages and also support mutual authentication between network nodes [60]. Table 2 presents some of the common asymmetric cryptographic solutions with security requirements support and their limitations.

A security protocol based on PKI was introduced by Raya et al. in which every vehicle is equipped with several private keys and their corresponding certificates [43]. The above security scheme is inefficient and apparently cannot manage to facilitate large vehicle populations due to its computational hardness. Efficient Conditional Privacy Preservation (ECPP) protocol is proposed by Liu et al. [61]. Instead of storing many anonymous keys and certificates, ECPP protocol generates short-time anonymous keys and certificates to reduce storage requirement. However, this protocol involves complex processing to generate anonymous keys, which results in serious computational overhead. Lin et al., Studer et al., and Ying et al. proposed hash chains based authentication protocol to deal with the overhead issue [21, 62, 63].

ID-based signatures are proposed to hide real identities of vehicles [51, 64]. Biswas et al. proposed an ID-based proxy method by using signatures [64]. This authentication technique is effective but vulnerable to reveal private key. Lo et al. also proposed similar authentication protocol which is based on elliptic curve cryptography [51]. Privacy preservation schemes also use ID-based signatures to provide anonymity [62, 65, 66]. In above schemes, public keys are used as vehicles' identity, so there is no need to store certificates. However, the scheme is vulnerable to replay attack [67]. Zhang et al. showed that the above technique is also vulnerable to impersonation attacks [62]. In order to enhance the security, Zhang et al. presented a privacy preservation scheme by using improved ID-based authentication process which generates digital signatures for vehicles' anonymity. However this scheme is vulnerable to the modification attacks [68]. Moreover, the above ID-based signature techniques lead to computational overhead because of bilinear pairing calculations. Recently Qun et al. proposed linearly homomorphic signature schemes that allow performing linear computations on authenticated data [69]. Qun et al. also proposed a short homomorphic proxy signature scheme. Proxy signature schemes permit an original signer to hand over his/her signing authority to a proxy signer, so that the proxy signer can sign on behalf of the original signer [70].

Zhang et al. presented another asymmetric technique based group signature method which allows RSUs to authenticate messages from vehicles [41, 62]. Zhang et al. also proposed RSU-aided authentication method using Hash Message Authentication Codes for secure vehicular communication [71]. In the above scheme, RSU provides a symmetric key to each vehicle by a key agreement protocol. Jung et al. also presented an RSU-aided privacy preservation technique

Table 2: Asymmetric cryptographic solutions for VANET.

Asymmetric Cryptographic solutions	Security Requirements support	Limitations
Anonymous keys and certificates.	Authentication/Availability/Privacy Preservation	Computational hardness
ID-based proxy signatures	Authentication/Privacy Preservation/Non-Repudiation	Vulnerable to reveal private key
Elliptic Curve Cryptography	Authentication/Availability/Privacy Preservation	Vulnerable to replay attack
RSU-aided Authentication Methods	Authentication/Privacy Preservation	Compromise on an RSU can result in disclosure of information
Smart Cards for identification	Message authentication/privacy preserving	Storage

Table 3: Symmetric ciphers for vehicular network with respect to security requirements and attack mitigation.

Cryptographic Ciphers	Security Requirements support	Attack Mitigation
Blowfish	Authentication/Availability	Differential related-key attacks/Brute-force attack
PBAS	Authentication/Availability/Confidentiality	DoS attack, Impersonation attack
Camellia	Authentication/Availability/Privacy Preservation	Impersonation attack/DoS/Sybil attacks
CAST	Authentication/Availability/Confidentiality	Sybil/Impersonation attack/routing attacks

that assigns anonymous certificates to vehicles which helps to minimize system overhead [72]. RSU-aided schemes however become easy targets for the attackers because they are semitrusted authorities. Compromise on an RSU can result in disclosure of information.

Use of smart cards has also been suggested for authentication and identification of vehicle under active attack scenarios. Paruchuri et al. proposed smart cards in vehicular networks for message authentication [73]. Smart cards can store users private/public keys, real identity, and the related certificates. However there are limitations regarding the storage. Smart cards can only store small amount of data whereas the data required to store private/public keys, real identity, and the related certificates may exceed the capacity.

4.2. Symmetric Algorithms. Symmetric algorithms often require less memory resources and tend to run comparatively faster than asymmetric algorithms. A wealth of established symmetric algorithms exists; among those the most prominent representatives are the block ciphers: Advanced Encryption Standard (AES) and Data Encryption Standard (DES). Other than block ciphers, there also exit several symmetric stream ciphers, which prove to be even more efficient as compared to block ciphers. Stream ciphers sometimes are preferred for embedded applications; however block ciphers are still more secure. We present the list of symmetric ciphers that are proposed to meet the security requirements of vehicular networks as shown in Table 3.

(1) **Blowfish:** Blowfish is a symmetric block cipher which was designed by Bruce Schneier in 1993 [87]. It provides an efficient encryption rate in software based embedded devices. It is equipped with variable length keys, which allows user to trade off between security and speed. A simple encryption algorithm makes it fast and efficient. Blowfish is a license-free and unpatented cipher that is available for free for almost all applications. However Blowfish cipher is vulnerable to attacks on a class of keys known to be weak [74]; therefore Blowfish users must select keys carefully. Although it suffers from weak keys attacks, there is no attack on S-boxes and subkeys generated by cipher itself. If the private key is large enough then brute-force key search is not possible. It is also secure against differential related-key attack.

(2) **PBAS:** Proxy-based Authentication Scheme (PBAS) allows proxy vehicles to authenticate multiple messages from other vehicles by using its computational capacity. This scheme helps to reduce the load on RSUs [75]. It also provides RSUs with an independent and systematic mechanism to authenticate messages from the proxy vehicle. In addition to this, PBAS is also able to negotiate session key with other vehicles to make the sensitive information confidential. PBAS scheme continues working properly, even if few proxy vehicles are compromised in the network, which makes it fault tolerant. It is an effective security scheme for efficient authentication in VANET.

(3) **Camellia:** Nippon Telegraph and Mitsubishi Electric Corporation in 2000 joined together to develop a symmetric cipher called Camellia [47, 76]. It has the same security level and processing capacity as compared to AES. It is compatible for both hardware and software implementations on common 8-bit processors as well as 32-bit processors, for instance, cryptographic hardware, smart cards, and embedded systems. Camellia provides high level security on multiple platforms for embedded systems.

(4) **CAST:** Carlisle Adams and Stafford Tavares in 1996 created a symmetric cipher which was named as CAST [77]. It is commonly a 64-bit block cipher which also allows key sizes up to 128 bits and 256 bits. CAST is used in applications of GPG and PGP as the default symmetric cipher [74]. Canadian government has approved it for the use of Secure Communication Establishment. CAST cipher has the ability to survive against linear and differential cryptanalysis attacks.

4.3. Lightweight Protocols. Based on asymmetric and symmetric cryptography, the following lightweight protocols have been designed to enhance future automotive security and meet the VANET security requirements as shown in Table 4:

(1) **ARAN:** Authenticated Routing for Ad hoc Network (ARAN) is based on Ad hoc On-demand Distance Vector (AODV) routing protocol in which the third-party CA presents signed certificate to vehicular nodes [78]. Every new node joining the network has to send request certificate to CA. All authorized nodes are provided with the public-key of CA. ARAN uses timestamps for route freshness and asymmetric cryptographic technique for secure route discovery authentication.

(2) **SEAD:** Secure and Efficient Ad hoc Distance (SEAD) vector protocol is based on dynamic destination-sequenced distance vector routing (DSDV) [79]. It

Table 4: Lightweight Protocols for vehicular network with respect to security requirements and attack mitigation.

Lightweight Protocols	Security Requirements	Attack Mitigation
ARAN	Message authentication/Integrity	Impersonation/Eavesdropping/Replay
SEAD	Authentication/Availability/Privacy Preservation	Routing/DoS/Impersonation Attacks
Ariadne	Availability/Privacy Preservation	DoS/Routing/Replay attacks
SAODV/A_SAODV	Authentication/Availability/Privacy Preservation	Impersonation/Bogus/information/Routing attack
OTC	Availability	Session hijacking
ECDSA	Authentication	Bogus information/Impersonation Attacks
RobSAD	Confidentiality/Authentication/Integrity	Sybil Attack
Holistic	Authentication/Confidentiality	Impersonation Attacks

works on one-way hash function for authentication purpose. This protocol shields against incorrect routing. Destination-sequence number is used to avoid long-lived route and ensure route freshness. The protocol applies intermediate node hashing to guarantee the authenticity of each route.

(3) **ARIADNE:** This protocol is based on Dynamic Source Routing (DSR) on-demand routing protocol [80]. Ariadne works very efficiently with symmetric cryptographic operations. It uses one-way hash function and MAC authentication for secure communication between nodes. Authorization is done by using shared key. TESLA broadcast authentication technology is source of Ariadne protocol that uses TESLA time interval for authentication and route discovery process.

(4) **SAODV:** This protocol was projected to embed security in AODV [81]. Hash functions are used to protect hop count and all messages are digitally signed to ensure authenticity of routes. However, this approach prevents the intermediate node to send any route reply even if it knows the fresh route. This problem can be solved by using Double Signature, but at the cost of system complexity increase.

(5) **A-SAODV:** An extension of secure ad hoc on-demand distance vector (SAODV) protocol was proposed that has features of adaptive reply decisions. Depending on the threshold conditions and queue length, each intermediate node can make decision to reply to source node [82].

(6) **OTC:** Generally, cookies are allotted per session for session management purpose. One time cookie (OTC) protocol is proposed to protect the system from theft of SID and session hijacking [83]. This protocol generates tokens for every request and attaches them to the request by using HMAC to avoid the reuse of token.

(7) **ECDSA:** As the name suggests Elliptical Curve Digital Signature (ECDS) Algorithm uses digital signature [84]. Asymmetric cryptographic operations with hash function provide security and authenticity to the system. The sender and receiver both require agreeing upon elliptical curve parameters.

(8) **RobSAD:** This protocol provides an efficient method for Sybil attack detection [85]. Sybil node is identified if two or more nodes have similar motion trajectories. Two different vehicles driven by different drivers cannot hold same motion patterns, because each person drives according to his own need and comfort.

(9) **Holistic protocol:** In this protocol, the authentication of every vehicle is done by RSU [86]. Vehicles are registered to RSU by sending a "Hello" message. In response, the RSU sets up a registration ID (consisting vehicle registration number and licence number) and sends it to the vehicle. Further authentication is made through certificate supplied by RSU. Data can only be shared if the node is authenticated by RSU or else the node is blocked.

4.4. Physical Layer Key Generation Schemes. There are certain attacks that may attempt to extract the private key from security devices. These types of attacks are known as side channel attacks [88]. They are performed by observing the electromagnetic radiations, power consumption, or the timing behavior of an embedded device. After collecting this information, the attacker attempts to extort the secret key by utilizing signal processing techniques. Side channel attacks approve being severe threat in the real world unless some extraordinary countermeasures are applied to generate dynamic secret keys based on physical layer. The basic advantage of generating dynamic key on physical layer is that there is no direct key distribution process involved. In the ideal condition, an eavesdropper cannot obtain any information related to the secret key [89]. Secret keys can be generated dynamically for two terminals by using random characteristics of the communication channel such as received signal strength (RSS), frequency phase information, or secrecy wiretap channel codes as shown in Figure 2. These random characteristics of channel are known as channel state information (CSI). Recently there has been focus on extracting similar feature from the channel which can be used to generate dynamic secret keys on physical layer.

The theoretical secrecy extraction characteristics from the correlation of random source were first considered in open literature in 1993 [90]. These schemes exploit random characteristics of the physical layer to share secret keys. It

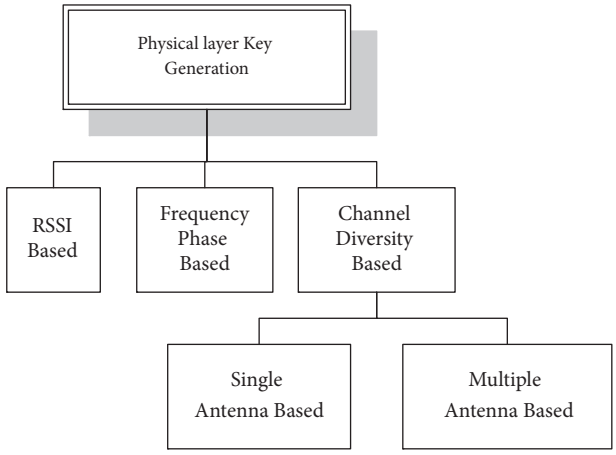


Figure 2: Physical layer key generation schemes for vehicular security.

is clearly shown that the correlated information of random sources can be used to extract secret keys by communicating over a shared channel, whereas the leaked information rate is arbitrarily low to the eavesdropper.

The best attainable secret key generation rate is defined as secret key capacity. Physical layer key generation scheme has gained significant attention in recent years due to its lightweight and information theoretic security features [88, 91–99].

The main challenge in physical layer key generation is to find a proper random source for high key generation rate. It is shown that there is a tradeoff between the public communication rate and secret key generation rate in the key agreement process [92, 93]. The random source is provided by an artificial signal and secret key is generated by the quasi-static fading channel [91]. Signals are sent if the channel state of legitimate node has better correlation than that of eavesdropper. However, the above approach contains certain assumptions that are difficult to realize in practice. Recently key generation in fast fading channels is a challenging issue and limits the application related to vehicular communications. Physical layer based key generation schemes are designed with the vehicle's maximum speed up to 50 mph but their key generation rates are limited to 5 bit/s [100, 101]. Much attention is needed in this area to develop certain schemes to improve the key generation rates. Moreover new random characteristics of the fading channel need to be explored in order to achieve higher key generation rates with more security.

4.5. Comparison. In Table 5, we present a summary of asymmetric, symmetric, and lightweight cryptographic techniques for attack mitigation and security requirements support. We also present the related references for the reader to understand these security protocols that are a foundation towards future automotive security. All protocols have their own advantages and disadvantages. A designer may select these protocols according to their own preferences. For instance some protocols provide good authentication but they are vulnerable to location based attacks; on the other hand, some protocols provide strong privacy but they are computationally complex. So there is a need to trade off for the choice of best suitable algorithms for securing the network. New standards can be developed by combining the existing protocols or use in parallel with the techniques presented in Table 5 to enhance the vehicular security.

5. Conclusion

Information technology has achieved vital significance for many new applications and services for automotive industry. The majority of innovations in cars are mainly based on software and electronic technology. Intelligent transportation systems are developed, aiming to improve the efficiency and safety of transportation systems. Security of these systems is a pivotal concern for next generation automotive technology. Conventional cryptographic algorithms such as public-key infrastructure, elliptic curve cryptography, HASH functions, and symmetric key cryptography may not be applied directly in vehicular networks due to their high mobility and dynamic network topology. Vehicular networks require real-time response and cannot tolerate delay in communication. Therefore, the conventional protocols that are developed for traditional networks fail to provide high throughput performance, low latency, and reliability for vehicular networks. So, there is a need to implement secure lightweight cryptographic algorithms on small embedded devices at acceptable execution time.

We argue that lightweight cryptographic protocols play a vital role in order to tackle the upcoming security challenges in future automotive technology, especially regarding vehicular safety and traffic efficiency. Security concerns for the future automotive industry act as a barrier in the way of extensive deployment of vehicular networks commercially. There is a need for understanding security threats and finding a solution to secure automotive technology by either building new lightweight cryptographic protocols or even using already existing algorithms in an efficient way. The public acceptance for new technology in vehicular networks can only be ensured by optimizing the security and privacy of users.

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Key Research and Development Program (no. 2016YFB0800602), the National Natural Science Foundation of China (NSFC) (no. 61502048), and Shandong provincial Key Research and Development Program of China (2018CXGC0701, 2018GGX106005).

Table 5: Summary of cryptographic techniques for attack mitigation and security requirements support.

	Cryptographic Solutions	Security Requirements	Attack Mitigation	References
Asymmetric Algo	Anonymous keys and certificates.	Authentication/Availability/Privacy Preservation	Eavesdropping/Replay/Impersonation attacks	[21, 63, 66, 68]
	ID-based proxy signatures	Authentication/Privacy Preservation/Non-Repudiation	Routing/DoS/Impersonation attacks	[46, 51, 62, 64]
	Elliptic Curve Cryptography	Authentication/Availability/Privacy Preservation	DoS/Routing/Replay attacks	[51]
	RSU-aided authentication method	Privacy Preservation/Authentication	Impersonation/Bogus/information Routing attack	[66, 72]
	Smart cards for identification	Message authentication/integrant/privacy preserving	Impersonation/Sybil attack	[71, 73]
Symmetric Algo	Blowfish	Authentication/Availability	Differential related-key attacks/broot force attack	[74]
	PBAS	Authentication/Availability/Confidentiality	DoS attack, Impersonation attack	[75]
	Camellia	Authentication/Availability/Privacy Preservation	Impersonation attack/DoS/Sybil attacks	[76]
	CAST	Authentication/Availability/Confidentiality	Sybil/Impersonation attack/routing attacks	[77]
Lightweight Protocols	ARAN	Message authentication/Integrity	Impersonation/Eavesdropping/Replay	[78]
	SEAD	Authentication/Availability	Routing/DoS	[79]
	Ariadne	Availability/Privacy Preservation	DoS/Routing/Replay attacks	[80]
	SAODV/A_SAODV	Authentication/Availability/Privacy Preservation	Impersonation/Bogus/information/Routing attack	[81, 82]
	One Time Cookie	Availability	Session hijacking	[83]
	ECDSA	Authentication	Bogus information/Impersonation Attacks	[84]
	RobSAD	Confidentiality/Authentication/Integrity	Sybil Attack	[85]
	Holistic	Authentication/Confidentiality	Impersonation Attacks	[86]

References

[1] T. Vaa, M. Penttinen, and I. Spyropoulou, "Intelligent transport systems and effects on road traffic accidents: state of the art," *IET Intelligent Transport Systems*, vol. 1, no. 2, pp. 81–88, 2007.

[2] J. Shen, T. Zhou, X. Chen, J. Li, and W. Susilo, "Anonymous and traceable group data sharing in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 912–925, 2018.

[3] R. Stanica, E. Chaput, and A.-L. Beylot, "Properties of the MAC layer in safety vehicular ad hoc networks," *IEEE Communications Magazine*, vol. 50, no. 5, pp. 192–200, 2012.

[4] B. Jarupan and E. Ekici, "A survey of cross-layer design for VANETs," *Ad Hoc Networks*, vol. 9, no. 5, pp. 966–983, 2011.

[5] H. Trivedi, P. Veeraraghavan, S. Loke, A. Desai, and J. Singh, "Routing mechanisms and cross-layer design for vehicular ad hoc networks: a survey," in *Proceedings of the IEEE Symposium on Computers and Informatics*, pp. 243–248, March 2011.

[6] J. Shen, Z. Gui, S. Ji, J. Shen, H. Tan, and Y. Tang, "Cloud-aided lightweight certificateless authentication protocol with anonymity for wireless body area networks," *Journal of Network and Computer Applications*, vol. 106, pp. 117–123, 2018.

[7] K. N. Qureshi and A. H. Abdullah, "A survey on intelligent transportation systems," *Middle East Journal of Scientific Research*, vol. 15, no. 5, pp. 629–642, 2013.

[8] S. Ahmed and S. S. Kanere, "SKVR: Scalable knowledge-based routing architecture for public transport networks," in *Proceedings of the Third ACM International Workshop on Vehicular Ad Hoc Networks (VANET '06)*, pp. 92-93, September 2006.

[9] Y. L. Morgan, "Notes on DSRC & WAVE standards suite: its architecture, design, and characteristics," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 504–518, 2010.

[10] Y. Toor, P. Muhlethaler, and A. Laouiti, "Vehicle ad hoc networks: applications and related technical issues," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 3, 2008.

[11] C.-C. Huang-Fu, Y.-B. Lin, and N. Alrajeh, "Mobility management of unicast services for wireless access in vehicular environments," *IEEE Wireless Communications Magazine*, vol. 19, no. 2, pp. 88–95, 2012.

[12] M. Torrent-Moreno, J. Mittag, P. Santi, and H. Hartenstein, "Vehicle-to-vehicle communication: fair transmit power control for safety-critical information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3684–3703, 2009.

[13] T. Tielert, D. Jiang, Q. Chen, L. Delgrossi, and H. Hartenstein, "Design methodology and evaluation of rate adaptation based congestion control for vehicle safety communications," in *Proceedings of the IEEE Vehicular Networking Conference (VNC '11)*, pp. 116–123, November 2011.

[14] G. Bansal, J. B. Kenney, and C. E. Rohrs, "LIMERIC: a linear adaptive message rate algorithm for DSRC congestion control," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4182–4197, 2013.

[15] J. Xu, L. Wei, Y. Zhang, A. Wang, F. Zhou, and C. Gao, "Dynamic Fully Homomorphic encryption-based Merkle Tree for lightweight streaming authenticated data structures," *Journal of Network and Computer Applications*, vol. 107, pp. 113–124, 2018.

[16] A. Dhamgaye and N. Chavhan, "Survey on security challenges in VANET 1," *International Journal of Computer Science and Network*, vol. 2, no. 1, pp. 88–96, 2013.

[17] S. Zeadally, R. Hunt, Y.-S. Chen, A. Irwin, and A. Hassan, "Vehicular ad hoc networks (VANETS): status, results, and challenges," *Telecommunication Systems*, vol. 50, no. 4, pp. 217–241, 2012.

[18] J. Blum and A. Eskandarian, "The Threat of Intelligent Collisions," *IT Professional*, vol. 6, no. 1, pp. 24–29, 2004.

[19] T. S. Rappaport, *Wireless Communications: Principles and Practice*, vol. 2, Prentice Hall PTR, New Jersey, NJ, USA, 1996.

[20] R. G. Engoulou, M. Bellaïche, S. Pierre, and A. Quintero, "VANET security surveys," *Computer Communications*, vol. 44, pp. 1–13, 2014.

[21] A. Studer, F. Bai, B. Bellur, and A. Perrig, "Flexible, extensible, and efficient VANET authentication," *Journal of Communications and Networks*, vol. 11, no. 6, pp. 574–588, 2009.

[22] X. Lin and X. Li, "Achieving efficient cooperative message authentication in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3339–3348, 2013.

[23] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, "The TESLA broadcast authentication protocol," *Rsa Cryptobytes*, vol. 5, no. 2, pp. 2–13, 2005.

[24] M. N. Mejri, J. Ben-Othman, and M. Hamdi, "Survey on VANET security challenges and possible cryptographic solutions," *Vehicular Communications*, vol. 1, no. 2, pp. 53–66, 2014.

[25] F. Qu, Z. Wu, F.-Y. Wang, and W. Cho, "A security and privacy review of VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 2985–2996, 2015.

[26] Z. Ye, S. Krishnamurthy, and S. Tripathi, "A framework for reliable routing in mobile ad hoc networks," in *Proceedings of the Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM '03)*, pp. 270–280, San Francisco, Calif, USA, 2003.

[27] M. K. Marina and S. R. Das, "On-demand multipath distance vector routing in ad hoc networks," in *Proceedings of the 9th International Conference on Network Protocols (ICNP '01)*, pp. 14–23, November 2001.

[28] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, "Privacy in inter-vehicular networks: why simple pseudonym change is not enough," in *Proceedings of the IEEE/IFIP International Conference on Wireless On-Demand Network Systems and Services (WONS '10)*, pp. 176–183, February 2010.

[29] K. Sampigethaya, M. Li, L. Huang, and R. Poovendran, "AMOEBA: robust location privacy scheme for VANET," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1569–1589, 2007.

[30] L. Buttyán, T. Holczer, A. Weimerskirch, and W. Whyte, "SLOW: a practical pseudonym changing scheme for location privacy in VANETs," in *Proceedings of the IEEE Vehicular Networking Conference (VNC '09)*, pp. 1–8, IEEE, Tokyo, Japan, October 2009.

[31] Z. Huang, S. Liu, X. Mao, K. Chen, and J. Li, "Insight of the protection for data security under selective opening attacks," *Information Sciences*, vol. 412-413, pp. 223–241, 2017.

[32] Y. Wei and Y. Chen, "Safe distance based location privacy in vehicular networks," in *Proceedings of the 71st Vehicular Technology Conference*, pp. 1–5, Taipei, Taiwan, May 2010.

[33] B. Palanisamy and L. Liu, "Attack-resilient mix-zones over road networks: architecture and algorithms," *IEEE Transactions on Mobile Computing*, vol. 14, no. 3, pp. 495–508, 2015.

[34] R. Yu, J. Kang, X. Huang, S. Xie, Y. Zhang, and S. Gjessing, "MixGroup: accumulative pseudonym exchanging for location privacy enhancement in vehicular social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 1, pp. 93–105, 2016.

[35] X. Zhang, Y. Tan, C. Liang, Y. Li, and J. Li, "A covert channel over VoLTE via adjusting silence periods," *IEEE Access*, vol. 6, pp. 9292–9302, 2018.

[36] H. Li, H. Zhu, and D. Ma, "Demographic information inference through meta-data analysis of wi-fi traffic," *IEEE Transactions on Mobile Computing*, vol. 17, no. 5, pp. 1033–1047, 2018.

[37] T. Li, J. Li, Z. Liu, P. Li, and C. Jia, "Differentially private Naive Bayes learning over multiple data sources," *Information Sciences*, vol. 444, pp. 89–104, 2018.

[38] X. F. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, "New algorithms for secure outsourcing of modular exponentiations," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 9, pp. 2386–2396, 2014.

[39] J. Li, Z. Liu, X. Chen, F. Xhafa, X. Tan, and D. S. Wong, "L-EncDB: a lightweight framework for privacy-preserving data queries in cloud computing," *Knowledge-Based Systems*, vol. 79, pp. 18–26, 2015.

[40] B. Li, Y. Huang, Z. Liu, J. Li, Z. Tian, and S. Yiu, "HybridORAM: practical oblivious cloud storage with constant bandwidth," *Information Sciences*, 2018.

[41] J. R. Douceur, "The sybil attack," in *Peer-to-Peer Systems*, vol. 2429 of *Lecture Notes in Computer Science*, pp. 251–260, Springer, Berlin, Germany, 2002.

[42] G. Guette and B. Ducourthial, "On the Sybil attack detection in VANET," in *Proceedings of the Internatonal Conference on Mobile Adhoc and Sensor Systems (MASS '07)*, pp. 1–6, October 2007.

[43] M. Raya and J.-P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[44] M. S. Al-Kahtani, "Survey on security attacks in vehicular ad hoc networks (VANETs)," in *Proceedings of the 6th International Conference on Signal Processing and Communication Systems (ICSPCS '12)*, pp. 1–9, Queensland, Australia, December 2012.

[45] S. Sharma, "A review: analysis of various attacks in VANET," *International Journal of Advanced Research in Computer Science*, vol. 7, no. 3, pp. 249–253, 2006.

[46] D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing," in *Advances in Cryptology—(CRYPTO 2001)*, vol. 2139 of *Lecture Notes in Computer Science*, pp. 213–229, Springer, 2001.

[47] S. Park, B. Aslam, D. Turgut, and C. C. Zou, "Defense against Sybil attack in the initial deployment stage of vehicular ad hoc network based on roadside unit support," *Security and Communication Networks*, vol. 6, no. 4, pp. 523–538, 2013.

[48] T. W. Chim, S. M. Yiu, L. C. Hui, and V. O. Li, "Security and privacy issues for inter-vehicle communications in VANETs," in *Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops (SECON Workshops' 09)*, pp. 1–3, Rome, Italy, June 2009.

[49] Y. Kim, I. Kim, and C. Y. Shim, "A taxonomy for DOS attacks in VANET," in *Proceedings of the 14th International Symposium on Communications and Information Technologies (ISCIT '14)*, pp. 26-27, September 2014.

[50] A. Rathod and S. Patel, "A survey on black hole & gray hole attacks detection scheme for vehicular ad-hoc network," *International Research Journal of Engineering and Technology*, vol. 04, no. 11, pp. 1508–1511, 2017.

[51] N.-W. Lo and J.-L. Tsai, "An efficient conditional privacy-preserving authentication scheme for vehicular sensor networks without pairings," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1319–1328, 2016.

[52] A. Prado, S. Ruj, and A. Nayak, "Enhanced privacy and reliability for secure geocasting in VANET," in *Proceedings of the 2013 IEEE International Conference on Communications (ICC '13)*, pp. 1599–1603, June 2013.

[53] Q. G. Fan, L. Wang, Y. N. Cai et al., "VANET routing replay attack detection research based on SVM," *Matec Web of Conferences*, vol. 63, p. 05020, 2016, EDP Sciences.

[54] H. Zhu, C. Fang, Y. Liu, C. Chen, M. Li, and X. S. Shen, "You can jam but you cannot hide: defending against jamming attacks for geo-location database driven spectrum sharing," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 10, pp. 2723–2737, 2016.

[55] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: a review," *IEEE Access*, vol. 6, pp. 10179–10188, 2018.

[56] J. Li, L. Sun, Q. Yan, Z. Li, W. Srisa-an, and H. Ye, "Significant permission identification for machine learning based android malware detection," *IEEE Transactions on Industrial Informatics*, no. 99, 2018.

[57] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secure multiple amplify-and-forward relaying with cochannel interference," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 8, pp. 1494–1505, 2016.

[58] X. Lai, W. Zou, D. Xie, X. Li, and L. Fan, "DF relaying networks with randomly distributed interferers," *IEEE Access*, vol. 5, pp. 18909–18917, 2017.

[59] H. Li, H. Zhu, S. Du, X. Liang, and X. Shen, "Privacy leakage of location sharing in mobile social networks: attacks and defense," *IEEE Transactions on Dependable and Secure Computing*, no. 99, 2016.

[60] J.-P. Hubaux, S. Capkun, and J. Luo, "The security and privacy of smart vehicles," *IEEE Security & Privacy*, vol. 2, no. 3, pp. 49–55, 2004.

[61] R. Lu, X. Lin, H. Zhu, P.-H. Ho, and X. Shen, "ECPP: efficient conditional privacy preservation protocol for secure vehicular communications," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 1229–1237, April 2008.

[62] C. Zhang, R. Lu, X. Lin, P.-H. Ho, and X. Shen, "An efficient identity-based batch verification scheme for vehicular sensor networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 246–250, April 2008.

[63] B. Ying, D. Makrakis, and H. T. Mouftah, "Privacy preserving broadcast message authentication protocol for VANETs," *Journal of Network and Computer Applications*, vol. 36, no. 5, pp. 1352–1364, 2013.

[64] S. Biswas and J. Misic, "A cross-layer approach to privacy-preserving authentication in WAVE-enabled VANETs," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 2182–2192, 2013.

[65] 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.

[66] X. Lin, X. Sun, X. Wang, C. Zhang, P.-H. Ho, and X. Shen, "TSVC: timed efficient and secure vehicular communications

with privacy preserving," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4987–4998, 2008.

[67] C.-C. Lee and Y.-M. Lai, "Toward a secure batch verification with group testing for VANET," *Wireless Networks*, vol. 19, no. 6, pp. 1441–1449, 2013.

[68] J. K. Liu, T. H. Yuen, M. H. Au, and W. Susilo, "Improvements on an authentication scheme for vehicular sensor networks," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2559–2564, 2014.

[69] Q. Lin, H. Yan, Z. Huang, W. Chen, J. Shen, and Y. Tang, "An ID-based linearly homomorphic signature scheme and its application in blockchain," *IEEE Access*, vol. 6, pp. 20632–20640, 2018.

[70] Q. Lin, J. Li, Z. Huang, W. Chen, and J. Shen, "A short linearly homomorphic proxy signature scheme," *IEEE Access*, vol. 6, pp. 12966–12972, 2018.

[71] C. Zhang, X. Lin, R. Lu, and P.-H. Ho, "RAISE: an efficient RSU-aided message authentication scheme in vehicular communication networks," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 1451–1457, May 2008.

[72] Y. Jiang, M. Shi, X. Shen, and C. Lin, "BAT: a robust signature scheme for vehicular networks using Binary Authentication Tree," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1974–1983, 2009.

[73] V. Paruchuri and A. Durresi, "PAAVE: protocol for anonymous authentication in vehicular networks using smart cards," in *Proceedings of the 53rd IEEE Global Communications Conference (GLOBECOM '10)*, pp. 1–5, December 2010.

[74] M. Alimohammadi and A. Pouyan, "Performance analysis of cryptography methods for secure message exchanging in VANET," *International Journal of Scientific &Engineering Research*, vol. 5, no. 2, pp. 911–917, 2014.

[75] W. Huang, Y. Xiong, and D. Chen, "DAAODV: a secure ad-hoc routing protocol based on direct anonymous attestation," in *Proceedings of the 7th IEEE/IFIP International Conference on Computational Science and Engineering (CSE '09)*, pp. 809–816, August 2009.

[76] S. Chang, Y. Qi, H. Zhu, J. Zhao, and X. Shen, "Footprint: detecting Sybil attacks in urban vehicular networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 6, pp. 1103–1114, 2012.

[77] C. Adams, "The CAST-128 encryption algorithm," 1997, https://dl.acm.org/citation.cfm?id=RFC2144.

[78] K. Sanzgiri, B. Dahill, B. N. Levine, C. Shields, and E. M. Belding-Royer, "A secure routing protocol for ad hoc networks," in *Proceedings of the 10th IEEE International Conference on Network Protocols (ICNP '02)*, pp. 78–87, IEEE, Paris, France, November 2002.

[79] P. Mutalik, S. Nagaraj, J. Vedavyas, R. V. Biradar, and V. G. C. Patil, "A comparative study on AODV, DSR and DSDV routing protocols for Intelligent Transportation System (ITS) in metro cities for road traffic safety using VANET route traffic analysis (VRTA)," in *Proceedings of the 2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT '16)*, pp. 383–386, December 2016.

[80] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Ariadne: a secure on-demand routing protocol for ad hoc networks," *Wireless Networks*, vol. 11, no. 1-2, pp. 21–38, 2005.

[81] M. G. Zapata and N. Asokan, "Securing ad hoc routing protocols," in *Proceedings of the 1st ACM Workshop on Wireless Security*, pp. 1–10, Atlanta, GA, USA, September 2002.

[82] D. Cerri and A. Ghioni, "Securing AODV: the A-SAODV secure routing prototype," *IEEE Communications Magazine*, vol. 46, no. 2, pp. 120–125, 2008.

[83] I. Dacosta, S. Chakradeo, M. Ahamad, and P. Traynor, "One-time cookies: preventing session hijacking attacks with stateless authentication tokens," *ACM Transactions on Internet Technology*, vol. 12, no. 1, pp. 1–24, 2012.

[84] S. S. Manvi, M. S. Kakkasageri, and D. G. Adiga, "Message authentication in vehicular ad hoc networks: ECDSA based approach," in *Proceedings of the International Conference on Future Computer and Communication (ICFCC '09)*, pp. 16–20, April 2009.

[85] C. Chen, X. Wang, W. Han, and B. Zang, "A robust detection of the sybil attack in urban VANETs," in *Proceedings of the 29th IEEE International Conference on Distributed Computing Systems Workshops (ICDCS Workshops '09)*, pp. 270–276, Montreal, Quebec, Canada, June 2009.

[86] K. S. TamilSelvan and R. Rajendiran, "A holistic protocol for secure data transmission in VANET," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 6, pp. 2278-1021, 2013.

[87] B. Schneier, "Description of a new variable-length key, 64-bit block cipher (Blowfish)," in *Fast Software Encryption*, vol. 809 of *Lecture Notes in Computer Science*, pp. 191–204, Springer, Berlin, Germany, 1994.

[88] S. Watanabe and Y. Oohama, "Secret key agreement from correlated gaussian sources by rate limited public communication," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 93, no. 11, pp. 1976–1983, 2010.

[89] B. Yang and J. Zhang, "Physical layer secret-key generation scheme for transportation security sensor network," *Sensors*, vol. 17, no. 7, p. 1524, 2017.

[90] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. I. Secret sharing," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.

[91] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

[92] T.-H. Chou, S. C. Draper, and A. M. Sayeed, "Key generation using external source excitation: capacity, reliability, and secrecy exponent," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 58, no. 4, pp. 2455–2474, 2012.

[93] S. Watanabe and Y. Oohama, "Secret key agreement from vector Gaussian sources by rate limited public communication," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 541–550, 2011.

[94] S. Nitinawarat and P. Narayan, "Secret key generation for correlated Gaussian sources," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 58, no. 6, pp. 3373–3391, 2012.

[95] T. Shimizu, H. Iwai, and H. Sasaoka, "Physical-layer secret key agreement in two-way wireless relaying systems," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 650–660, 2011.

[96] H. Zhou, L. M. Huie, and L. Lai, "Secret key generation in the two-way relay channel with active attackers," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 476–488, 2014.

[97] T.-H. Chou, S. C. Draper, and A. M. Sayeed, "Secret key generation from sparse wireless channels: ergodic capacity and secrecy outage," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1751–1764, 2013.

[98] M. Wilhelm, I. Martinovic, and J. B. Schmitt, "Secure key generation in sensor networks based on frequency-selective channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1779–1790, 2013.

[99] V. M. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via sources and channels," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 58, no. 11, pp. 6747–6765, 2012.

[100] J. Wan, A. B. Lopez, and M. A. Al Faruque, "Exploiting wireless channel randomness to generate keys for automotive cyber-physical system security," in *Proceedings of the ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS '16)*, pp. 1–10, Vienna, Austria, April 2016.

[101] X. Zhu, F. Xu, E. Novak, C. C. Tan, Q. Li, and G. Chen, "Using wireless link dynamics to extract a secret key in vehicular scenarios," *IEEE Transactions on Mobile Computing*, vol. 16, no. 7, pp. 2065–2078, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 7107295, 17 pages
https://doi.org/10.1155/2018/7107295

Research Article

Biometrics Based Privacy-Preserving Authentication and Mobile Template Protection

Wencheng Yang,[1] **Jiankun Hu**,[2] **Song Wang**,[3] **and Qianhong Wu**[4]

[1]*Security Research Institute, School of Science, Edith Cowan University, WA 6027, Australia*
[2]*School of Engineering and Information Technology, University of New South Wales at the Australian Defence Force Academy (UNSW@ADFA), Canberra, ACT 2600, Australia*
[3]*School of Engineering and Mathematical Sciences, La Trobe University, VIC 3086, Australia*
[4]*School of Electronic and Information Engineering, Beihang University, Beijing, China*

Correspondence should be addressed to Jiankun Hu; j.hu@adfa.edu.au

Received 27 December 2017; Revised 26 April 2018; Accepted 10 May 2018; Published 12 June 2018

Academic Editor: Kim-Kwang Raymond Choo

Smart mobile devices are playing a more and more important role in our daily life. Cancelable biometrics is a promising mechanism to provide authentication to mobile devices and protect biometric templates by applying a noninvertible transformation to raw biometric data. However, the negative effect of nonlinear distortion will usually degrade the matching performance significantly, which is a nontrivial factor when designing a cancelable template. Moreover, the attacks via record multiplicity (ARM) present a threat to the existing cancelable biometrics, which is still a challenging open issue. To address these problems, in this paper, we propose a new cancelable fingerprint template which can not only mitigate the negative effect of nonlinear distortion by combining multiple feature sets, but also defeat the ARM attack through a proposed feature decorrelation algorithm. Our work is a new contribution to the design of cancelable biometrics with a concrete method against the ARM attack. Experimental results on public databases and security analysis show the validity of the proposed cancelable template.

1. Introduction

Nowadays, mobile devices, e.g., smartphone, have become one of our daily necessities. They are used to store personal data and handle private communication. Unfortunately, there are some privacy and security issues along with the use of smartphones; for example, a user's private data, e.g., photo, contacts, and bank details, can be compromised, if his/her smartphone without any protection is stolen or lost. Secret knowledge-based approaches, e.g., password or PIN, are commonly used for authentication. However, these methods rely heavily on the user to ensure continued validity, and poor use of password or PIN may lead to great security breaches [1].

Fingerprint recognition has been extensively used in various applications, e.g., authentication on mobile devices. With good recognition accuracy and high convenience, fingerprint-based authentication systems hold more markets than other biometrics, e.g., face, iris, and voice. Fingerprint authentication is usually composed of two stages, the enrollment stage and the verification stage [2]. In the enrollment stage, feature data extracted from a captured fingerprint image are considered as a template, which is stored in a database. In the verification stage, the query feature data, extracted in the same way as template feature data, are compared with the stored template, and a match or nonmatch verdict will be made.

The use of fingerprint authentication systems eliminates the need of remembering long passwords or PIN, which is the disadvantages of traditional password- or PIN-based authentication schemes. However, fingerprint-based biometric systems have one main drawback; that is, once a fingerprint is compromised, it will be lost forever. To address this issue, cancelable fingerprint templates [3], among other biometric template protection schemes, have been proposed. Instead of storing raw fingerprint data as a template, it is distorted through a one-way transformation function in the enrollment stage. Such a transformation is intentional and repeatable. One important property of cancelable fingerprint templates is noninvertibility, which means that it should be

computationally hard to recover the raw fingerprint data from the transformed fingerprint template [4]. In the verification stage, the same transformation is applied to the query data. Matching between the transformed template and query is conducted in the transformed domain. In this way, if the stored template is compromised, a new version of it can be generated by changing the transformation parameter(s) [5].

1.1. Related Work

1.1.1. Mobile Biometrics. A lot research effort has been devoted to the design of more precise, usable, and secure biometric authentication schemes on mobile devices. For instance, in [16], Clarke and Furnell introduced a method for authenticating users by getting them to input telephone numbers or write text messages. This method is called biometric keystroke analysis. In [17], Kim and Hong proposed to use teeth together with voice to authenticate users, which is the first research work using the teeth and voice combination. The matching scores of each individual trait are calculated and fused using a weighted-summation operation. The experiments are conducted by using a dataset that contains one thousand teeth images and voices collected by smartphones. Later they proposed an enhanced multimodal authentication system [18], which adds another biometric trait, face, on top of teeth and voice to achieve better results.

In [19], Lee et al. designed a mobile multimodal biometric system based on finger-vein and fingerprint. The proposed system can obtain fingerprint and finger-vein images simultaneously and also is able to overcome some limitations of unimodal biometric systems, e.g., lack of accuracy. In [20], Tao and Veldhuis developed a face-based biometric authentication system on mobile devices, which contained detailed information about the process, including face detection, registration, illumination normalization, verification, and information fusion.

In [21], Chen et al. proposed a fingerprint-based remote authentication method using mobile devices. In their method, both fingerprint and password are involved to improve the security level of the system. Moreover, hashing functions are used to implement mutual authentication. In [22], a palmprint-based recognition system is proposed for mobile devices. Specifically, a hand-shaped guide window is introduced for fast image acquisition and an enhanced competitive code is used to cope with image variation.

In [23], Rattani et al. investigated gender prediction from ocular images acquired by smartphones so as to enhance the accuracy of the integrated biometric authentication and mobile healthcare system. In [24], Marsico et al. compared the performance of several participant methods in the Mobile Iris Challenge Evaluation-1 contest. Furthermore, some analysis is given to image covariate and interoperability.

Several survey papers, e.g., [25, 26], discussed the biometric authentication methods on mobile devices, including current development, trends, and challenges. The significance of template protection on mobile devices is described in [26]. However, we notice that most of above-mentioned biometric authentication methods do not protect the biometric template on mobile devices, which potentially put important personal information at risk. With good recognition accuracy and high convenience offered by fingerprint, in this paper, we propose a fingerprint-based authentication system using the cancelable technique to provide template protection on mobile devices.

1.1.2. Cancelable Biometrics. The concept of cancelable biometrics was initiated by Ratha et al. in [27]. Later, they constructed a practical cancelable fingerprint authentication system [4] by using three different transformation functions, namely, Cartesian transformation, polar transformation, and functional transformation. The transformation functions are able to distort the fingerprint minutiae feature into a new data format. This method is registration-based and hence relies on precise detection of the reference points, e.g., singular, core, or delta points. However, fingerprint uncertainty caused by displacement, nonlinear distortion, and rotation during the process of fingerprint capturing is unavoidable, thus making accurate registration hard to achieve. A matching error can possibly be caused by a registration error [28].

To relinquish the process of global registration and also reduce the impact of nonlinear distortion [29, 30], registration-free local structure based methods have been proposed; see, e.g., [5–11, 13, 14, 31–40]. In [5], Yang et al. proposed a cancelable template design based on geometric transformation. Each local Delaunay triangle-based structure instead of each single minutia acts as a unit to be transformed under the guidance of two transformation matrices. In [31], Farooq et al. presented a cancelable fingerprint template based on a set of triangles derived from any three minutiae. The features extracted from the set of triangles are further converted into the binary format. Then this binary string was randomly permuted into a different feature representation under the conduct of a user specific key. The cancelability of the feature representation can be achieved by applying different keys. Lee et al. [32] proposed an approach to calculating a rotation- and translation-invariant value from the orientation information of neighboring local regions around each minutia. The invariant value is then utilized as the input of two transformation functions to generate transformed features. The cancelable template is governed by these two transformation functions. In [33], in order to avoid global alignment, the authors used localized matching, which consists of matching minutia triplets constructed by each minutia and its two nearest neighbors. Invariant features extracted from these triplets are varied and secured by the symmetric hash functions. A major drawback of this approach is the assumption that the genuine query sample minutiae, being described as locations in a complex plane, are linearly transformed from the template minutia set. However, the most challenging issue for fingerprint minutia-based matching is the nonlinear elastic distortion of minutia locations. Therefore, this assumption is not realistic. Ahn et al. [6] applied geometrical properties, e.g., local relation, from minutiae triplets to hiding the minutiae information. Generation of these geometrical features is conducted via a transformation function in an attempt to keep the discriminating capability.

In [7], Yang et al. introduced the geometrically aligned and protected minutia vicinity for template protection. The proposed method transforms the original minutia vicinity by adding some parameter guided offsets into each minutiae group so as to destroy the original local topological relationship among those minutiae. The original minutia vicinity is defined by a minutia together with its M closest neighboring minutiae. In [34], minutia pairs applying redundant combinations of two minutiae points are formed to counter some image noise. A bit-string cancelable template is derived from the minutia pairs. Similarly to [31], the same user specific tokens are used to guarantee that the bit-string features are permuted in the same manner during both enrollment and verification stages. A different local structure represented by a 3D array was proposed in [35]. In this structure, each minutia is chosen as the reference point and other minutiae are rotated and translated based on the orientation and position of the reference point so as to map the minutiae into the 3D array. Each minutia after transformation falls into a specific cell of the 3D array and each cell is marked as 1 if more than one minutia locates in it; otherwise it is marked as 0. A resultant bit-string is then permuted by using a user specific PIN. Yang et al. [8] proposed to use a dynamic random projection method to protect the biometric features extracted from local structures composed of a reference minutia and three closest minutiae around it. Although this method can dynamically choose a projection matrix from a set of candidate projection matrices, it is a primitive direct biometric key generation technique, which has poor error tolerance for low quality fingerprint images.

Ahmad et al. [9] built a pair-polar coordinate-based alignment-free structure. The pair-polar structure uses the relative position of each minutia to other minutiae in a polar coordinate space. Three local features are extracted from any two minutiae and then a functional transformation is applied to these local features to achieve the resultant cancelable template. In [36], a circular region is constructed around each minutia and the circular region can be divided into different levels according to different radiuses. Then the circular regions are encrypted by two transformation functions and stored as the cancelable template. In [37], Jin et al. generated a revocable fingerprint template via a polar grid-based method. For each reference minutia, the polar transform is performed first to align the remaining minutiae. After that, a 3-tuple quantization technique is utilized to generate the local feature in the form of a bit-string. The user specific token based permutation technique is also utilized for feature transformation. Das et al. [38] proposed to use the minimum distance graphs, which involve a set of interminutia minimum distance vectors starting from the core point, as rotation- and translation-invariant features to conduct hash transformation so as to protect the original template.

In contrast to the pair-polar based structure proposed in [9], Wang and Hu [10–12, 41] mainly considered the noninvertible transformation functions. In [10], a densely infinite-to-one mapping method is presented to accomplish the transformation; in [11], a curtailed circular convolution is used to achieve noninvertible transformation; in [41], the identifiability condition in blind channel estimation is deliberately violated to protect the source input—the binary string's frequency samples; and, in [12], the binary biometric representations are securely protected by the partial Hadamard transformation, which transforms them into complex vectors. In [39], Wong et al. designed a multiline code for generating the cancelable fingerprint template. The multiline code is a string-based minutia descriptor extracted from a set of minutiae surrounding a virtual line within a specified range. Then a user specific secret key guided permutation is performed to achieve feature transformation. In [13], Jin et al. developed a two-dimensional random projection technique to secure the minutiae-based fingerprint template generated from minutia vicinities. Each minutia vicinity is formed by a minutia and three nearest minutiae around it. The feature matrix generated from a set of minutia vicinities is transformed/mapped onto a random subspace determined by an external orthogonal random matrix, which is generated by a user specific token. In [14], Jin et al. adopted the same local structure as [13] but incorporated a different noninvertible transformation method named randomized graph-based hamming embedding (RGHE) to protect the original features. In [40], Zhang et al. presented two methods, which are a designed combo plate and a functional transformation, to produce cancelable templates based on the MCC code, which associates a local structure to each minutia [42]. Note that the method named P-MCC in [42] is not a cancelable template as it does not provide the property of revocability. Subsequently, a partial permutation based scheme named 2P-MCC [15] was proposed to add revocability to P-MCC. However, 2P-MCC suffers from the ARM. P-MCC is not revocable, which means that the feature vector $\widehat{V}$ that contains k binary values in P-MCC remains unchanged in different applications. The 2P-MCC scheme uses a user specific key s as the index and chooses c ($0 < c \leq k$) binary values from those k binary values of $\widehat{V}$ to create a new feature vector $\ddot{V}$. When unlinkability is required, at least one element of the c binary values from any two templates, e.g., $\ddot{V}$, should be different. By combining templates from at most (k-c+1) applications, the feature vector $\widehat{V}$can be restored. Once feature vector $\widehat{V}$ is obtained by the adversary and due to $\widehat{V}$ being the same in different applications, the adversary can verify the correctness of the inverted feature generated from template $\ddot{V}$ and the user specific key s by comparing it with $\widehat{V}$.

In [43], Kaur and Khanna presented a cancelable biometric method, named random slop, which can reduce the feature dimensions by up to 75%. This method has been tested on many biometric databases, such as face, palmprint, palmvein, and fingervein, except fingerprint, although it can be implemented to fingerprint features. In order to further enhance recognition accuracy and security of cancelable unibiometric systems, in [44], Paul et al. developed a cancelable biometric template creation algorithm using random biometric feature fusion, random projection, and selection based on face and ear. In [45], Yang et al. proposed a fingerprint and fingervein based cancelable multibiometric system, in which a feature-level fusion strategy with three fusion options are designed. In the meantime, an enhanced partial discrete

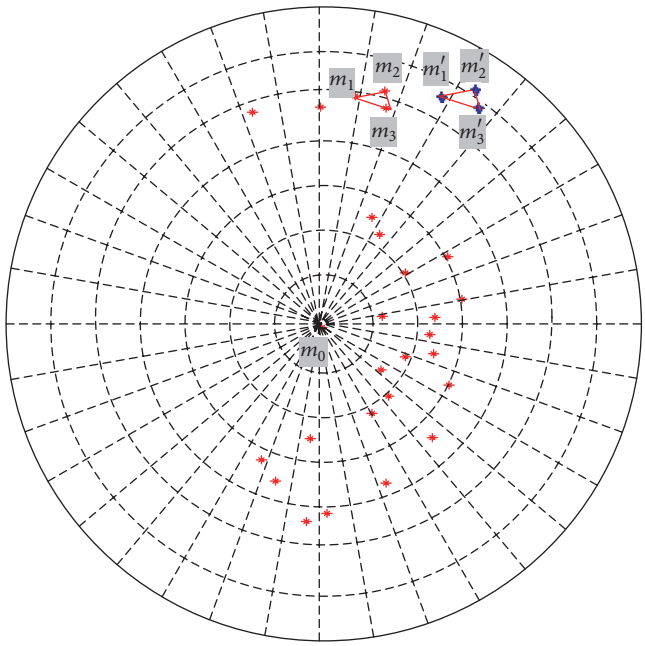


FIGURE 1: A polar space centering around minutia m_0.

Fourier transform based noninvertible transformation is applied.

1.2. Motivation and Contributions. From the above-mentioned cancelable fingerprint template systems, it is observed that the rotation- and transformation-invariant local region based structures are employed and the main difference between these local structures is the region size. For example, small-sized local structures are used in [6–8, 13, 14, 33]. In these systems, the local structures are only composed of one reference minutia and two or three of its neighboring minutiae. Small-sized structures can resist nonlinear distortion to some extent. However, since these local structures only contain a moderate number of minutiae, the feature data extracted from them are not discriminative enough. By contrast, relatively large local structures are employed in [9–11, 31, 34, 35, 37–40]. Large-sized structures consist of more minutiae and tend to include relatively more information that can be exploited for matching. However, the positional change of minutiae for large-sized structures is likely to occur under nonlinear distortion, especially when minutiae are located far away from the reference point. Because of nonlinear distortion, for example, minutiae m_1, m_2, and m_3 in the template image and their corresponding minutiae m_1', m_2', and m_3' in the query image are possibly located in different cells in a polar coordinate system, as shown in Figure 1. To strike a balance and achieve good matching performance, some authors [32, 36] proposed to use parameters to control the region size. However, testing different parameter settings in a practical authentication process would create additional computational load and resource burden or even be infeasible.

Another issue that troubles the existing cancelable template systems is the security concern. The methods in [5, 31, 34, 35, 37, 39] use permutation matrices for feature transformation. Since the permutation functions are invertible [10], original fingerprint features are not secured safely if the permutation matrices are compromised. Although the methods in [7, 9–15, 40, 41] that use the noninvertible many-to-one mapping strategy can avoid the above issue, they suffer from the attacks via record multiplicity (ARM) [46, 47], if multiple transformed templates and their corresponding transformation parameters are acquired by an adversary. Readers can refer to [46, 47] for more details about the ARM.

To address the above issues, in this paper, we propose a new cancelable fingerprint template system. Not only can the proposed system reduce the impact of nonlinear distortion on those minutiae that are located far away from the reference point, but it also can defend the system against the ARM attack. In particular, the new cancelable fingerprint template system processes both local and global structures and fuses two different schemes at the score level, so that the overall system achieves better recognition accuracy than a single scheme only. The overall processing flow of the proposed system is shown in Figure 2. Specifically, in the enrollment stage, two feature sets, the polar coordinate-based feature set (T1) and the Delaunay triangulation-based feature set (T2), are extracted from the polar coordinate-based local structure and Delaunay triangulation-based global structure, respectively. Next, feature set T1 is processed by the proposed feature decorrelation algorithm and then varied by a random projection transformation function. Feature set T2 is permutated under the guidance of the feature codes. Both transformed feature sets, T1 and T2, are stored in the database. In the verification stage, the same transformations used in the enrollment stage are applied to the query feature sets, Q1 and Q2. Matching between the transformed template and query feature sets is conducted in the transformed domain.

The main contributions of this paper are highlighted as follows.

(1) Two schemes, polar coordinate-based scheme and Delaunay triangulation-based scheme, are utilized and fused on the score level. Specifically, the polar coordinate-based scheme uses the feature set from the polar coordinate space and acts as our basic scheme, which can provide reasonable feature discrimination. The Delaunay triangulation-based scheme that uses the feature set from Delaunay triangulation serves as a complement to the basic scheme. It can effectively reduce the negative impact of nonlinear distortion on the minutiae that are located far away from the reference minutia. As a result, compared with the system that only uses the polar coordinate-based scheme, higher recognition accuracy is achieved.

(2) A main reason that cancelable biometrics suffer the ARM attack is due to feature correlation which exists among cancelable templates stored across multiple applications but actually derived from the same biometric features. To address the issue, we propose a feature decorrelation algorithm so that the feature vectors, which are generated from the same feature set, are uncorrelated in different applications. Without feature correlation, the adversary would not have adequate information to determine the original feature data.

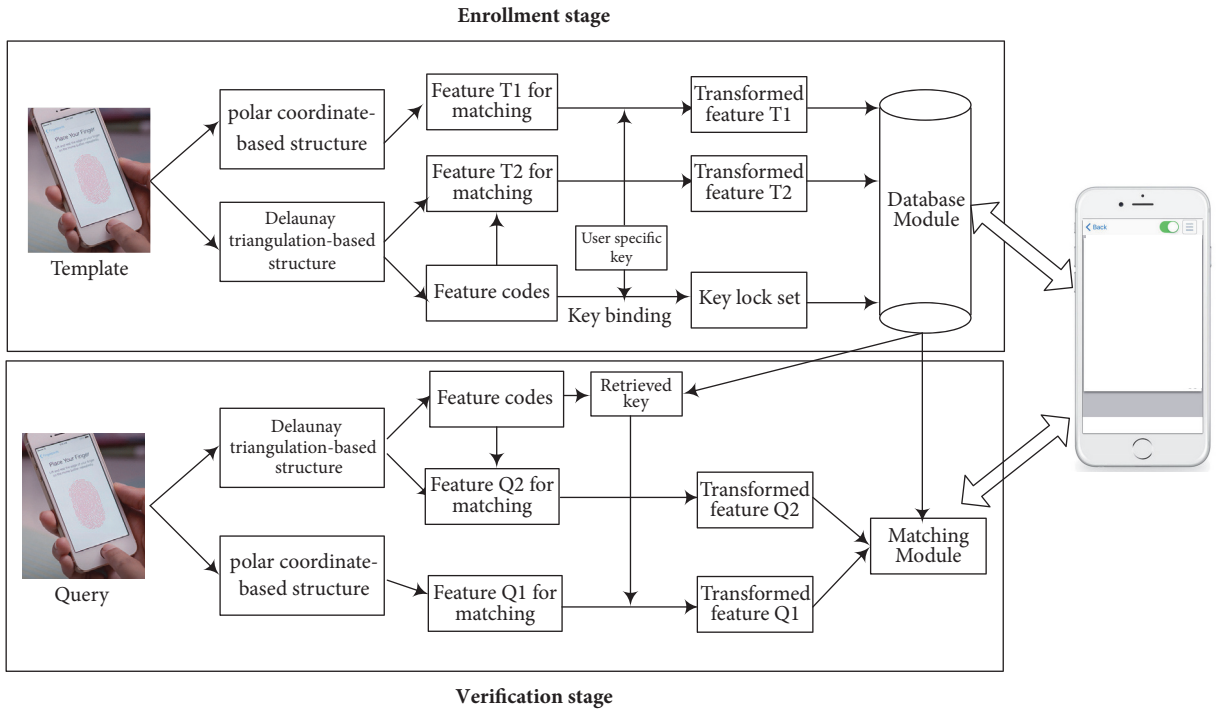


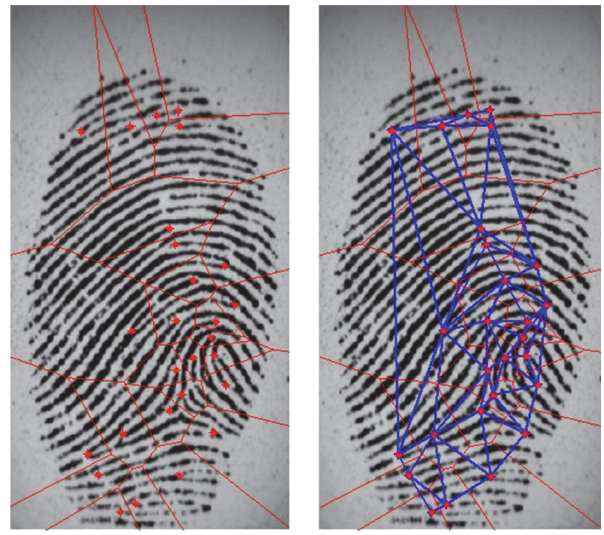
FIGURE 2: The overall processing flow of the proposed system (Features T1, T2, Q1, and Q2 represent the feature sets from the polar coordinate-based structure and Delaunay triangulation-based structure of the template image and query image, respectively).

The rest of the paper is organized as follows. In Section 2, two feature sets, polar coordinate-based feature set and Delaunay triangle-based feature set, are introduced. The proposed cancelable fingerprint authentication system, which can defeat the ARM attack, is presented in Section 3. In Section 4, experimental results and security analysis are demonstrated and discussed. The conclusion is given in Section 5.

2. Generation of Two Feature Sets

In this section, we introduce two feature sets that are extracted from the polar coordinate-based and Delaunay triangulation-based structures, respectively. These two structures are derived from the same minutiae set of a fingerprint image. Each polar coordinate-based structure is composed of a reference minutia and other minutiae in a predefined range. The use of the polar coordinate system allows the relationships between the reference minutia and other minutiae to be readily defined and measured. The Delaunay triangulation-based structure formed by a set of minutiae is a triangulation where no minutia in that set of minutiae is inside the circumcircle of any triangle in the triangulation. Specifically, given a set of minutiae $M = (m_0, m_1, m_2, \ldots, m_{N-1})$, where N is the number of minutiae, each minutia $m_{i\in[0,N-1]}$ can be represented by a vector $(x_i, y_i, \theta_i, t_i)$, where x_i and y_i are the x, y coordinates in the Cartesian coordinate system, θ_i is the orientation in the range of $[0, 2\pi)$, and t_i is the minutia type. For each fingerprint image, the following two feature sets are generated from the above two structures.

2.1. Polar Coordinate-Based Feature Set. In the polar coordinate-based structure, if the minutia, e.g., m_0, is considered as the origin of the polar coordinate and the remaining minutiae in range of R (=300 pixels) are rotated and translated with respect to m_0 such that the orientation of m_0 equals 0 degree in the polar coordinate system, then any minutia $m_{i\in[1,N-1]}$ can be converted and expressed as a triplet $(\rho_i, \alpha_i, \beta_i)$, where $0 < \rho_i \leq 300$ is the radial distance, $0 < \alpha_i \leq 2\pi$ is the radial angle, and $0 < \beta_i \leq 2\pi$ is the relative orientation of minutia m_i to m_0. An example of the polar space centered around minutia m_0 is shown in Figure 1. In order to tolerate small distortion, polar grid-based quantization [35] is performed on all the minutiae in the range of R. We assume that the step sizes of ρ_i, α_i, and β_i are s_ρ, s_α, and s_β, respectively ($5 \leq s_\rho \leq 20$ and $\pi/12 \leq s_\alpha, s_\beta \leq 2\pi/9$). Then the polar space centered around m_0 can be quantized into a 3D cube containing $l_C = L \times S \times H$ cells, where $L = \lfloor R/s_\rho \rfloor$, $S = \lfloor 2\pi/s_\alpha \rfloor$, and $H = \lfloor 2\pi/s_\beta \rfloor$. The cell where the minutia m_i is located in the 3D cube is $(\rho_i^q, \alpha_i^q, \beta_i^q)$, where $\rho_i^q = \lfloor \rho_i/s_\beta \rfloor$, $\alpha_i^q = \lfloor \alpha_i/s_\alpha \rfloor$, and $\beta_i^q = \lfloor \beta_i/s_\beta \rfloor$. By this means, we obtain a vector $P(m_0)$ of length l_C containing only '0's and '1's, in which '1' means the appearance of one or more minutiae in the corresponding cell. l_C varies under different parameter settings of s_ρ, s_α, and s_β. It is obvious that when these parameter settings change, the value of l_C changes accordingly. In our application, different parameter settings are chosen for different databases in order to achieve best performance. Within the value range of these parameter settings, the largest value of l_C that can be obtained is 34560. Therefore, we use a fixed value of $l_C = 34560$ for all the chosen databases and pad '0' to those feature vectors with elements



(a) (b)

Figure 3: An example of (a) Voronoi diagram, (b) Delaunay triangulation (bold line).

less than 34560. By applying the above approach to every minutia in the set of minutiae $M = (m_0, m_1, m_2, \ldots, m_{N-1})$, a polar coordinate-based feature set $C = \{P(m_i)\}_{i=0}^{N-1}$ is generated.

2.2. Delaunay Triangulation-Based Feature Set. As a complement to the polar coordinate-based feature set, we propose the second feature set, which is extracted from the Delaunay triangulation-based structure and takes advantage of the desirable features [48–50] of Delaunay triangulation. In the presence of nonlinear distortion, Delaunay triangulation has a stable local neighborhood structure. Minutiae will keep the same neighboring structure if the nonlinear distortion does not move minutiae out of the tolerance region. Also, noise in fingerprint images influences the Delaunay triangulation only locally. Spurious or missing minutiae affect only those local Delaunay structures which contain them. A brief description about the generation of a Delaunay triangulation is given below; readers can refer to [51] for more details.

For the set of minutiae $M = (m_0, m_1, m_2, \ldots, m_{N-1})$, a Voronoi tessellation, which divides the whole fingerprint image region into several smaller regions centering on each minutia, is created first as shown in Figure 3(a). All the points in the region around m_i are closer to m_i than to any other minutia. The Delaunay triangulation is generated by connecting the centers of every neighboring region as shown in Figure 3(b).

Assume that there are N_1 Delaunay triangles generated from N minutiae; several invariant features can be defined from each Delaunay triangle. Taking triangle $\Delta m_1 m_2 m_3$ as an example, four rotation- and translation-invariant features are defined as follows:

(i) $o_{m_1 m_2}$ is the orientation differences between m_1 and m_2.

(ii) $l_{m_2 m_3}$ is the length of edge $m_2 m_3$.

(iii) α_{m_3} is the angle between edge $m_1 m_3$ and $m_2 m_3$.

(iv) $t_{m_1 m_2 m_3}$ is the concatenation of minutia type of m_1, m_2, and m_3.

$0 < o_{m_1 m_2} \leq 2\pi$, $0 < l_{m_2 m_3} \leq 300$, and $0 < \alpha_{m_3} \leq 2\pi$. A feature data set, e.g., $f_{m_1 m_2 m_3} = (o_{m_1 m_2}, l_{m_2 m_3}, \alpha_{m_3}, t_{m_1 m_2 m_3})$, can be extracted from each Delaunay triangle. To tolerate the variation caused by nonlinear distortion that is inherent in fingerprint images, quantization is applied to each feature set. The quantization step sizes are set to be s_o, s_l, and s_α for $o_{m_1 m_2}$, $l_{m_2 m_3}$, and α_{m_3}, respectively ($15 \leq s_l \leq 25$, $\pi/12 \leq s_o, s_\alpha \leq \pi/9$). Each element of $f_{m_1 m_2 m_3}$, after quantization, can be expressed as $o^q_{m_1 m_2}$, $l^q_{m_2 m_3}$, $\alpha^q_{m_3}$, and $t_{m_1 m_2 m_3}$. If $o^q_{m_1 m_2}$, $l^q_{m_2 m_3}$, $\alpha^q_{m_3}$, and $t_{m_1 m_2 m_3}$ are represented by a_1, a_2, a_3, and a_4 bits, respectively, then the triangle $\Delta m_1 m_2 m_3$ can be represented by a bit-string $f^q_{m_1 m_2 m_3}$ by putting them in sequence together. The integer value of this bit-string falls in the range of $[0, l_D - 1]$, where $l_D = 2^{a_1 + a_2 + a_3 + a_4}$. With N_1 Delaunay triangles constructed from a given fingerprint image, each of them should match a value in $[0, l_D - 1]$ and so the corresponding bin is indexed by 1, as shown in [34]. As a result, the Delaunay triangulation-based feature set can be represented by a binary vector D of length l_D.

We now explain why we add the second feature set. In the Delaunay triangulation-based feature set, each Delaunay triangle rather than each minutia is treated as a feature unit because a Delaunay triangle is more robust against nonlinear distortion than each individual minutia. For example, as mentioned in Section 1, minutiae m_1, m_2, and m_3 in the template image may not seem to match their corresponding minutiae m'_1, m'_2, and m'_3 in the query image in a polar coordinate system since they fall in different cells due to nonlinear distortion, as shown in Figure 1. However, if we consider the triangle $\Delta m_1 m_2 m_3$ constituted by minutiae m_1, m_2, and m_3 as a unit, no matter where $\Delta m_1 m_2 m_3$ is moved, features such as the edge length and angle, extracted from it,

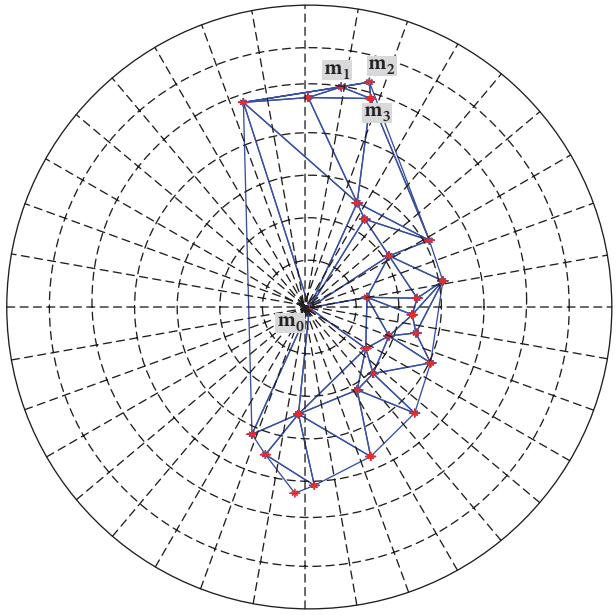


Figure 4: Delaunay triangulation-based structure centering around m_0.

still remain invariant. Thus, using these features, $\Delta m_1 m_2 m_3$ and $\Delta m_1' m_2' m_3'$ can match each other.

3. Design of the Cancelable Fingerprint Authentication System

In this section, we present the two general stages included in the proposed cancelable fingerprint authentication system.

3.1. Enrollment Stage. In the enrollment stage, the original feature sets are transformed. The detail of the enrollment stage is explained below.

Given a template image F^T (letter T means template), the untransformed feature sets C_T and D_T as introduced in Section 2 can be extracted from the polar coordinate-based and Delaunay triangulation-based structures as shown in Figure 4. However, feature sets C_T and D_T are vulnerable. Because if the adversary acquires the vector C_T, he/she would be able to know the minutiae's location in the 3D cube. Subsequently the feature triplet (ρ, α, β) of each minutia can be recovered. Similarly, it is not difficult to reveal the original minutiae information from D_T. Hence, it is crucial to protect the feature sets C_T and D_T. To achieve this, the polar coordinate-based scheme and the Delaunay triangulation-based scheme are utilized to protect each element of C_T and D_T, respectively.

3.1.1. The Polar Coordinate-Based Scheme. In the polar coordinate-based scheme, the feature set is processed by a feature decorrelation algorithm and then varied by a random projection matrix $\mathcal{M}$ generated under the participation of a user specific key k_{pm}. The loss of key k_{pm} means the loss of the projection matrix $\mathcal{M}$. The feature decorrelation algorithm can eliminate feature correlation in different applications, thus making it robust against the ARM attack.

INPUT: the original feature $P(m_i)$
Process:
Step 1: Each position V_p of value 1 that is located in the binary string $P(m_i)$, is input into a folding function $f_P = \mathrm{mod}(V_p, L_C)$, where L_C is an application-specific parameter and set to be smaller than l_C. mod(·) is a modulo operation, for example, mod(7,5) = 2 and mod(12,5) = 2.
Step 2: With each value of 1 in $P(m_i)$ relocated to a new position depending on f_P, a new binary string $\mathcal{P}(m_i)$ of length L_C is generated.
Step 3: $\mathcal{P}(m_i)$ is converted into a real-value vector through the Discrete Fourier Transform (DFT) as $\mathbb{P}(m_i) = DFT(\mathcal{P}(m_i))$.
OUTPUT: decorrelated feature $\mathbb{P}(m_i)$

Algorithm 1: Feature decorrelation algorithm.

For an element, e.g., $P(m_0)$, of the feature set $C_T = \{P(m_i)\}_{i=0}^{N^T-1}$, where N^T is the number of minutiae in the template image F^T, before performing random projection, it is first processed by the feature decorrelation algorithm as shown in Algorithm 1. This feature decorrelation algorithm is useful because $P(m_i)$ only includes values of 1 and 0 and has a sparse distribution, which might restrict the search space if the random projection is directly implemented on it. Furthermore, if C_T is applied in multiple applications, its feature correlation can be utilized by the adversary to launch the ARM attack. For the above reasons, Algorithm 1 is proposed to eliminate feature correlation in different applications.

It follows from Steps 1 and 2 in Algorithm 1 that the new binary string $\mathcal{P}(m_i)$, generated from the original feature $P(m_i)$, can be varied by simply adjusting the application-specific parameter L_C in different applications. Note that even if one bit in $\mathcal{P}(m_i)$ is changed, the output vector $\mathbb{P}(m_i)$ will be totally different due to the nature of the Discrete Fourier Transform (DFT) in Step 3. Therefore, feature correlation does not exist in the real-value vector $\mathbb{P}(m_i)$ used in different applications.

The modulo operation in Step 1 is a many-to-one mapping and controlled by the parameter L_C. Different L_C values may lead to varying matching performance, which is discussed in Case 3 of Section 4.1. Under the assumption that the random projection based transformation is conquered, the modulo operation might not be strong enough to safeguard $P(m_i)$ from the compromised $\mathcal{P}(m_i)$ if L_C is set to be a large value. For example, when $l_C = 34560$ and L_C is set to be 20000, for values of '1' that appear in $\mathcal{P}(m_i)$ from positions 1 to 14560, e.g., 123, its original position in $P(m_i)$ only has three possible locations, either 123 or 20123 or both. To increase security, we further propose Algorithm 2 as an enhancement of Algorithm 1.

Algorithm 2 produces a binary code based segment permutation before carrying out the steps in Algorithm 1.

INPUT: the original feature $P(m_i)$
Process:
Step 1: $P(m_i)$ is divided into two parts, P_1 and P_2, where P_1 contains the first L_C elements of $P(m_i)$, while P_2 contains the remaining elements.
Step 2: P_1 is evenly divided into N_S segments, so $P_1 = \{S_j\}_{j\in 1}^{Ns}$ and each segment S_j is of length $L_S = L_C/N_S$.
Step 3: For each segment S_j, calculate $\phi_j = sum(k * S_j(k))$, where $1 \le k \le L_S$. ϕ_j is a weighted sum of position k and its binary value for segment S_j. Then all the elements in S_j are circularly shifted left by ϕ_j bits to generate a permutated segment S_j'.
Step 4: All the permutated segments $\{S_j'\}_{j\in 1}^{Ns}$ are concatenated into a new binary vector P_1', which is further concatenated with P_2 to create a binary vector $P'(m_i) = P_1' \parallel P_2$
Step 5: Perform Steps 1 to 3 in Algorithm 1 with input $P'(m_i)$.
OUTPUT: decorrelated feature $\mathbb{P}(m_i)$

Algorithm 2: Enhanced feature decorrelation algorithm.

The segment permutation is guided by an on-the-fly binary code $\phi_j = sum(k * S_j(k))$, which is calculated using the elements' positions and values instead of a user specific key, so its security does not suffer from the lost key attack. Moreover, the binary code ϕ_j is calculated based on just L_S binary elements in S_j, which is only a portion of the original feature $P(m_i)$. This means that the error in one ϕ_j does not affect the permutation of other segments of $P(m_i)$. The parameter L_S is the length of each segment, which impacts on the matching performance and security strength. A larger value of L_S means that it would require more computational effort to invert S_j' to S_j, but it would be more likely to cause errors in ϕ_j, thus having a negative effect on matching accuracy. The matching performance and security related to parameter L_S are discussed in Case 3 of Section 4.1 and Section 4.4, respectively. We remark that Algorithm 2 offers an alternative pathway for tighter security at the expense of performance, as shown in Section 4.1 and that when $L_S = 1$, Algorithm 2 reverts to Algorithm 1.

The output vector $\mathbb{P}(m_i)$ of Algorithm 1 or 2 is then transformed by projecting onto a random space with the help of the random projection matrix $\mathscr{M}$ of size $X \times Y$, where $X = L_C$. The transformation of $\mathbb{P}(m_i)$ using the random matrix $\mathscr{M}$ can be compactly expressed by

$$\widehat{P}(m_i) = \mathbb{P}(m_i) \times \mathscr{M}, \tag{1}$$

where $\widehat{P}(m_i)$ is the inner product of $\mathbb{P}(m_i)$ and $\mathscr{M}$. The above transformation makes the dimension of transformed feature $\mathbb{P}(m_i)$ reduced to Y so that the feature set $\mathbb{P}(m_i)$ is protected. Different values of Y affect the performance of the proposed system, which is discussed in Section 4. Moreover, revocability is achieved—a new template can be issued easily by just changing the user specific key k_{pm}, which is analyzed in Section 4. By applying Algorithm 1 or 2, and random projection to each element of C_T, C_T can be transformed to be $\widehat{C}_T = \{\widehat{P}(m_i)\}_{i=0}^{N^T-1}$.

3.1.2. The Delaunay Triangulation-Based Scheme. In the Delaunay triangulation-based scheme, the feature set D_T, which is a binary vector of length l_D, is transformed by permutation guided by feature codes. Specifically, for each triangle, e.g., $\Delta m_1 m_2 m_3$, a feature code is calculated first. Some stable local features can be used to obtain the feature code. These features are defined as follows:

(i) $o_{m_2m_3}$ is the orientation differences between m_2 and m_3.
(ii) $l_{m_1m_2}$ is the length of edge m_1m_2.
(iii) $l_{m_1m_3}$ is the length of edge m_1m_3.
(iv) α_{m_1} is the angle between edges m_1m_2 and m_1m_3.

$0 < o_{m_2m_3} \le 2\pi$, $0 < l_{m_1m_2}, l_{m_1m_3} \le 300$, and $0 < \alpha_{m_1} \le 2\pi$. Quantization is applied to each of these features to tolerate small variation. Accordingly, a quantized four-element array $f^q_{m_1m_2m_3} = [o^q_{m_1m_2}, l^q_{m_1m_2}, l^q_{m_1m_3}, \alpha^q_{m_1}]$ can be extracted from triangle $\Delta m_1 m_2 m_3$ which is further input into (2), based on [52], to generate the feature code $f^c_{m_1m_2m_3}$, which corresponds to the triangle $\Delta m_1 m_2 m_3$, as follows:

$$\begin{aligned} f^c_{m_1m_2m_3} = {} & \Upsilon^3 f^q_{m_1m_2m_3}(4) + \Upsilon^2 f^q_{m_1m_2m_3}(3) \\ & + \Upsilon^1 f^q_{m_1m_2m_3}(2) + \Upsilon^0 f^q_{m_1m_2m_3}(1), \end{aligned} \tag{2}$$

where $\Upsilon = f(f^q_{m_1m_2m_3}(1), \ldots, f^q_{m_1m_2m_3}(4), \Phi)$. Υ is the output of function $f(\cdot)$ corresponding to inputs $f^q_{m_1m_2m_3}(1)$ to $f^q_{m_1m_2m_3}(4)$ and a random parameter Φ. Φ is set to be different in different applications, which enables Υ to be diverse. In this way, one set of feature codes $\{f_i^c\}_{i=1}^{N_1^T}$ are

computed as such from F^T, where N_1^T is the number of Delaunay triangles in the template image F^T. After the feature code $f^c_{m_1m_2m_3} \in \{f_i^c\}_{i=1}^{N_1^T}$ is obtained, the original bin to which $\Delta m_1m_2m_3$ is matched in D_T, which is decided by its corresponding integer value $f^{it}_{m_1m_2m_3}$, can be permutated under the guidance of the feature code $f^c_{m_1m_2m_3}$. So the new bin becomes $f^{it}_{m_1m_2m_3}+f^c_{m_1m_2m_3}$. To increase the permutation randomness, an extra key guided permutation could be added on top of above permutation. By the same token, bins of other Delaunay triangles can be shuffled to new locations determined by their corresponding feature codes in $\{f_i^c\}_{i=1}^{N_1^T}$. The vector array D_T, after permutation, is transformed into a new version $\widehat{D}_T$. Since each feature code in $\{f_i^c\}_{i=1}^{N_1^T}$ is generated from its corresponding Delaunay triangle feature and not saved in the database or on the smart card, it is hard to figure out the original bin location, even if the new bin position is acquired by the adversary from $\widehat{D}_T$.

The reason why permutation is used to protect the feature set D_T rather than random projection, as in the case of the feature set C_T, is that permutation does not reduce feature dimension, which can help minimize the impact on the system's matching performance. Moreover, feature code set is not saved in the database or on the smart card, which greatly improves the security of the original feature set D_T. Most importantly, noise in fingerprint images influences the Delaunay triangulation only locally. Spurious or missing minutiae affect only those local Delaunay structures which contain them. Each Delaunay triangle is formed by only three minutiae, which means that if a feature code, e.g., $f^c_{m_1m_2m_3}$, is incorrect, it only influences the permutated position of that Delaunay triangle, from which the feature code is generated. On the contrary, each polar coordinate-based structure is composed of a reference minutia and a number of other minutiae (usually more than ten minutiae) in a certain range. Obviously, the chance of generating an incorrect feature code from the polar coordinate-based structure is much higher than that from the Delaunay triangle, which only includes three minutiae. If we choose fewer minutiae from the polar coordinate-based structure, say, only three minutiae are chosen, which is the same number as that in the Delaunay triangle, then how to correctly determine the three minutiae from a query image and their corresponding minutiae from a template image is a tricky issue, especially under the presence of biometric uncertainty, e.g., spurious or missing minutiae. Therefore, compared with permutation, random projection is a better option for the protection of the feature set C_T.

3.2. Verification Stage. In the verification stage, the same projection matrix used in the enrollment stage is used to transform the polar coordinate-based feature set extracted from the query image, and the Delaunay triangulation-based feature set can also be transformed under the guidance of the feature codes. Finally, matching between the template image F^T and query image F^Q is conducted using the transformed features in the transformed domain. The detailed steps of the verification stage are explained below.

Given a query image F^Q (letter Q means query), untransformed query feature sets C_Q and D_Q together with a set of feature codes $\{f_j^c\}_{j=1}^{N_1^Q}$ are extracted first, where $C_Q = \{P(m_i)\}_{i=0}^{N^Q-1}$, N^Q is number of minutiae, and N_1^Q is the number of Delaunay triangles in the query image F^Q.

In the polar coordinate-based scheme, suppose that $\widehat{P}^Q(m_j)$ denotes the j^{th} transformed local feature of the query and $\widehat{P}^T(m_i)$ denotes the i^{th} transformed local feature of the template. Then the similarity score between them is calculated by

$$S_C = 1 - \frac{\left\|\widehat{P}^T(m_i) - \widehat{P}^Q(m_j)\right\|_2}{\left\|\widehat{P}^T(m_i)\right\|_2 + \left\|\widehat{P}^Q(m_j)\right\|_2}, \tag{3}$$

where $\|\cdot\|_2$ denotes the 2-norm. Each polar coordinate-based local feature in the query is compared with each polar coordinate-based local feature in the template to output a similarity score. Upon the completion of the comparison, there will be a score matrix of size $N^T \times N^Q$. The maximum value $S_{C\max}$ in this score matrix is considered to be the matching score of the first feature set between the template and query images.

In the Delaunay triangulation-based scheme, the set D_Q of the query is permuted under the guidance of the feature codes to generate the transformed version $\widehat{D}_Q$. The permutation procedure is the same as that in the enrollment stage. The similarity score between $\widehat{D}_Q$ and $\widehat{D}_T$ can be calculated as

$$S_D = \frac{\sum_{k=1}^{2^{l_D}} \left(\widehat{D}_{Q,k} - \overline{\widehat{D}_Q}\right)\left(\widehat{D}_{T,k} - \overline{\widehat{D}_T}\right)}{\sqrt{\sum_{k=1}^{2^{l_D}} \left(\widehat{D}_{Q,k} - \overline{\widehat{D}_Q}\right)^2 \sum_{k=1}^{2^{l_D}} \left(\widehat{D}_{T,k} - \overline{\widehat{D}_T}\right)^2}}, \tag{4}$$

where $\overline{\widehat{D}}$ represents the mean value.

The final score between the template image F^T and query image F^Q is calculated using both $S_{C\max}$ and S_D from the polar coordinate-based scheme and the Delaunay triangulation-based scheme, respectively, that is,

$$\begin{aligned} final\_score = {} & \rho_C \times norm\left(S_{C\max}\right) + (1 - \rho_C) \\ & \times norm\left(S_D\right), \end{aligned} \tag{5}$$

where $norm(\cdot)$ is a score normalization function; ρ_C is the weight of score $S_{C\max}$, which is set to be 0.7 in our application so as to give more relevance to the use of our basic scheme, the polar coordinate-based scheme. If the final matching score $final\_score$ is larger than a predefined threshold S_t, then the template image F^T and query image F^Q are considered to be matching.

4. Experimental Results and Security Analysis

Three databases (DB1, DB2, and DB3) of FVC2002 and one database (DB2) of FVC2004 were used to evaluate the proposed fingerprint cancelable template design. Detailed

Table 1: Detailed information about the databases used in our experiments.

Parameter	2002DB1	2002DB2	2002DB3	2004DB2
Resolution	500 dpi	569 dpi	500 dpi	500 dpi
Number of fingers	100	100	100	100
Number of images per finger	8	8	8	8
Sensor Type	Optical Sensor	Optical Sensor	Capacitive Sensor	Optical Sensor
Image size	388 × 374	560 × 296	300 × 300	328 × 364
Image quality	Medium	Medium	Low	Very low

information about these four databases is listed in Table 1. To extract minutiae from fingerprint images, a software package called VeriFinger 4.0 from Neurotechnology [53] was utilized.

The performance of the proposed fingerprint cancelable template is evaluated by four performance indices, namely, genuine acceptance rate (GAR), false acceptance rate (FAR), false rejection rate (FRR), and equal error rate (EER). GAR is defined as the ratio of successful genuine tests to the total number of genuine tests. FAR is defined as the ratio of successful imposter tests to the total number of imposter tests, FRR (=1-GAR) is defined as the ratio of failed genuine tests to the total number of genuine tests, and EER is defined as the error rate when FRR and FAR are the same. For all the databases, the 1VS1 matching protocol and standard FVC matching protocol [42] were utilized in the experiments. In the 1VS1 matching protocol, the first two images of each finger were chosen for testing, while, in the FVC matching protocol, all eight images of each finger were used.

4.1. Performance of the Proposed System under the Lost Key Attack. The lost key attack was tested in the experiments by allocating the same key k_{pm} to all the genuine and imposter tests. Three different cases were evaluated as follows.

Case 1 (performance of the system under two different instances). The performance of the system under three different instances is compared:

(i) Instance 1: using only the basic polar coordinate-based scheme

(ii) Instance 2: using only the Delaunay triangulation-based scheme

(iii) Instance 3: combining the polar coordinate-based scheme and Delaunay triangulation-based scheme

The polar coordinate-based scheme in both instances 1 and 3 uses Algorithm 1 and the same parameter settings $L_C = 20000$. The comparison between Instances 1, 2, and 3 is conducted over the database FVC2002 DB2 using the 1VS1 matching protocol. ROC curves are drawn in Figure 5. It can be observed from Figure 5 that Instance 3 (EER = 0.64%) performs better than Instance 1 (EER = 1.00%) and Instance 2 (EER = 2.00%), which only uses the single structure under the same parameter setting ($Y = 300$).

Case 2 (performance of the proposed system using the untransformed feature sets and the transformed feature sets).

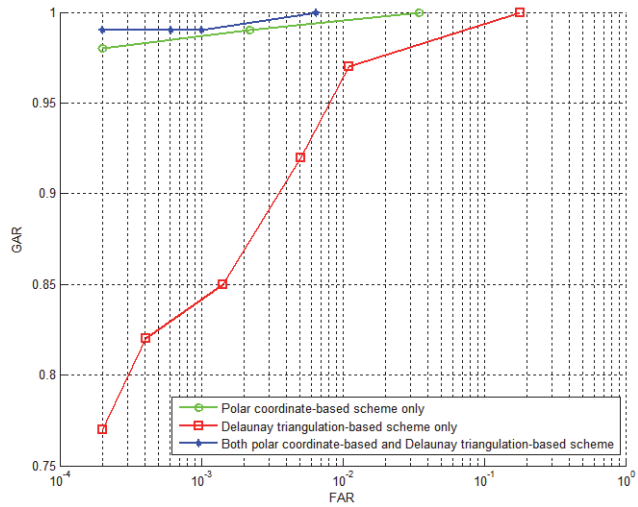


Figure 5: Performance of Instance 1 (the polar coordinate-based scheme only), Instance 2 (the Delaunay triangulation-based scheme only), and Instance 3 (both the polar coordinate-based scheme and Delaunay triangulation-based scheme).

Performance of the proposed system that uses the untransformed feature sets and the transformed feature sets was tested to evaluate the effect of feature transformation using Algorithm 1 and parameter settings $L_C = 20000$. This test was conducted on two databases, FVC2002 DB2 and DB3 with the 1VS1 matching protocol. For database DB2, we obtained EER = 0.62% using the untransformed feature sets and EER = 0.64% using the transformed feature sets. For database DB3, EER = 4% was obtained using the untransformed feature sets, while EER = 4.57% was obtained with the transformed feature sets. In this case, matching performance becomes worse than that before feature transformation.

Case 3 (effect of different parameter settings on the performance of the proposed system). The different parameter settings of L_C and L_S in Algorithms 1 and 2 result in a trade-off between security and matching performance [54]. Here, we investigate the effect of different parameter settings of L_C and L_S on the system's matching performance over databases FVC2002 DB2 and DB3 using the 1VS1 matching protocol. As shown in Table 2, choosing a smaller value of L_C in Algorithm 1 makes matching performance worse. This is because the modulo operation in Algorithm 1 is a many-to-one mapping and smaller L_C increases the possibility of

Table 2: EER (%) of the proposed system using different parameter settings in Algorithms 1 and 2.

		Using Algorithm 1	
L_C	20000	500	
FVC2002 DB2	0.64	1.00	
FVC2002 DB3	4.57	6.00	
		Using Algorithm 2 and $L_C = 20000$	
L_S	50	5	1
FVC2002 DB2	3.37	1.00	0.64
FVC2002 DB3	14.00	8.07	4.57

multiple '1's folded to the same position, leading to a decrease in feature discriminative ability. Algorithm 2 is proposed to increase the difficulty of obtaining the original feature vector. Under the same setting of $L_C = 20000$, we can see from Table 2 that a larger value of L_S reduces matching accuracy, for example, on database FVC2002DB2, EER is 3.37% when $L_S = 50$, as opposed to EER is 1.00% when $L_S = 5$. Both EERs are worse than EER = 0.64% when $L_S = 1$ (which is equivalent to just using Algorithm 1) because the feature-dependent segment permutation is controlled by the binary code ϕ_j and larger L_S, which represents a longer portion of the original feature vector, is more likely to make ϕ_j incorrect.

From the above analysis, we can see that smaller L_C or larger L_S can decrease matching performance. However, such a parameter setting makes the retrieval of the original feature vector harder under the assumption that random projection based transformation is conquered. Detailed security analysis is discussed in Section 4.4.

We also evaluated the performance of the proposed system using different values of Y over databases FVC2002 DB2 and DB3 with the 1VS1 matching protocol under the parameter setting $L_C = 20000$ in Algorithm 1. Security of the first feature set is provided by the random projection matrix $\mathscr{M}$ which is of size $X \times Y$, $(Y \leq X)$. The feature vector $P(m_i)$, extracted from the polar coordinate-base d structure, is transformed using Algorithm 1 and (1). The smaller the value of Y is set, the more the dimensions are reduced. A lower-dimensional transformed feature vector is more secure as less information of the original feature vector is kept. In the tests, we set the value of Y to be 300 and 50, respectively. The EER performance is listed in Table 3. It can be observed that the smaller the value of Y, the higher the EER. This is because less information about the original features is preserved with more dimension cut (smaller Y), leading to performance degradation. Moreover, under the same parameter setting ($Y = 300$), the proposed system demonstrates better performance on database FVC2002 DB2 with EER = 0.64% than the performance on FVC2002 DB3 with EER = 4.57%. The main reason for this is the vast difference of image quality between these two databases. The first two images from database FVC2002 DB2 have much better quality than the first two images of FVC2002 DB3 [10, 11].

4.2. Performance Comparison with the Similar Work. In this section, we compare the performance of the proposed system with other similar systems. The EER comparison under the lost key scenario is reported in Table 4. It can be observed that the proposed method outperforms all the existing similar methods under the 1VS1 matching protocol. Under the FVC matching protocol, the performance of proposed method is worse than that of original 2P-MCC$_{64,64}$; however, 2P-MCC suffers from the ARM, as analyzed in Section 1.1. By contrast, the proposed scheme is resilient to this attack. In [8] transformation parameters are derived directly from biometrics via quantization, causing them rarely to be identical for the same user when large minutia variance exists in fingerprint images, which leads to poor performance. The structures in [9–11] are constructed by each minutia with all other minutiae in the image, resulting in more computational complexity. By contrast, in the proposed method structures are only formed by each minutia and its local neighbors in the range of R. More importantly, the methods [7, 9–14] are vulnerable to the ARM attack [47].

Table 3: EER (%) of proposed system using different values of Y.

Y	300	50
FVC2002 DB2	0.64	1.00
FVC2002 DB3	4.57	8.00

4.3. Revocability and Unlinkability. Revocability is an essential property that a qualified cancelable template design should possess. Once a template is compromised, another transformed template should be generated and the newly generated template should be totally different from the compromised template and suffer no performance degradation. To measure revocability of the proposed system, we generated 50 different templates from the 1st image of each finger by 50 different user specific keys. Then the transformed templates were matched against the original ones. The imposter distribution (with different keys) and pseudo-imposter distribution on database FVC2002 DB2 are shown in Figure 6. It can be seen that the imposter distribution and pseudo-imposter distribution are similar, which means that even if multiple templates are generated from the same image, they are distinct from the original template and there is no performance degradation. The mean and standard derivations of the imposter distribution are 0.2457 and 0.0408, respectively, while the values of these two indicators of the pseudo-imposter distribution are 0.2419 and 0.0042, respectively.

Table 4: EER (%) comparison between the proposed system and similar systems under the lost-key scenario.

Methods	2002 DB1		2002 DB2		2002 DB3		2004 DB2	
	1VS1	FVC	1VS1	FVC	1VS1	FVC	1VS1	FVC
Yang et al. [5]	5.93	-	4	-	-	-	-	-
Ahn et al. [6]	-	7.18	-	3.61	-	11.80	-	-
Yang and Busch [7]	-	-	13	-	-	-	-	-
Yang et al. [8]	-	-	0.85	-	-	-	-	-
Ahmad et al.[9]	9	-	6	-	27	-	-	-
Wang and Hu [10]	3.50	-	4	5	7.5	-	-	-
Wang and Hu [11]	2	-	2.30	3	6.12	-	-	-
Wang and Hu [12]	1	-	2	-	5.20	-	13.30	-
Jin et al. [13]	3.07	-	1.02	-	-	-	-	-
Jin et al.[14]	4.36	-	1.77	-	-	-	21.82	-
Ferrara et al. [15] 2P-MCC$_{64,64}$	-	3.30	-	1.80	-	7.80	-	-
Instance 3 (proposed system)	0.32	5.75	0.64	4.71	4.57	10.22	9.90	12.00

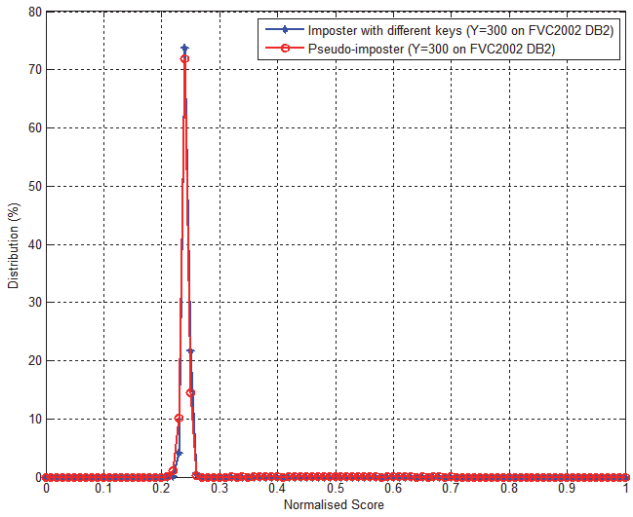


Figure 6: Imposter (with different keys) and pseudo-imposter distributions on FVC2002 DB2 when $Y = 300$.

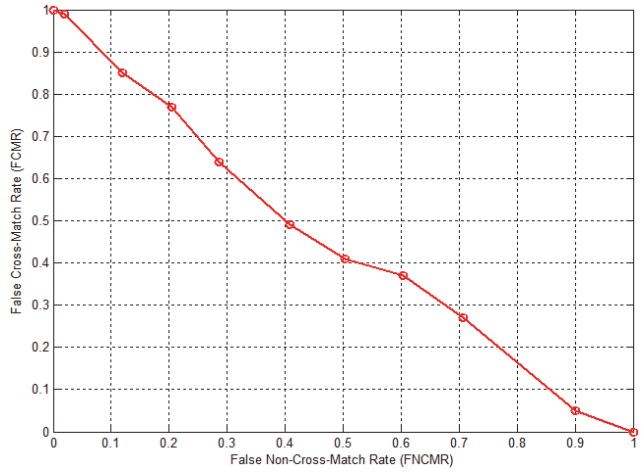


Figure 7: FCMR versus FNCMR tested on FVC2002 DB2 when $Y = 300$.

To protect users' privacy, unlinkability is another essential property of cancelable fingerprint templates. It requires that the transformed templates generated from the same finger using different keys should be different to one another, as if they were transformed from different fingers. In this way, templates from the same individual used in different applications cannot be cross-matched. To verify the unlinkability property of the proposed method, two indicators [55], the false cross-match rate (FCMR) and the false non-cross-match rate (FNCMR), are used in two cases over database FVC2002 DB2. Case 1: FNCMR is the ratio of unsuccessful matching attempts between the transformed templates of the first and second images of each finger. Feature transformation is based on different keys. Case 2: FCMR is the ratio of successful matching attempts between the transformed templates of the first image from each finger and the first image of different fingers. The FCMR and FNCMR curves are shown in Figure 7, from which it is clear that $FCMR + FNCMR \approx 1$, which conforms to the expected behavior of a cross-comparator [55].

4.4. Security Analysis. Given a template image F^T, we obtain feature sets C_T and D_T from the polar coordinate-based and Delaunay triangulation-based structures. The polar coordinate-based feature set $C_T = \{P(m_i)\}_{i=0}^{N^T-1}$ is protected by algorithms included in the polar coordinate-based scheme. Specifically, each element $P(m_i)$ in C_T first goes through the proposed feature decorrection algorithm, Algorithm 1 or 2, which transforms it into another format $\mathbb{P}(m_i)$, making the transformed feature sets uncorrelated in different applications, governed by application-specific parameters, L_C and L_S. Since the DFT is invertible, $\mathbb{P}(m_i)$ is further protected by random projection through the projection matrix $\mathcal{M}$, as shown in (1), which is essentially a many-to-one mapping. $\mathcal{M}$ is of size $X \times Y$ and Y is set to be smaller than X. Hence, the transformed feature vector $\widehat{P}(m_i)$ has a reduced dimension compared with $\mathbb{P}(m_i)$. The adversary can only

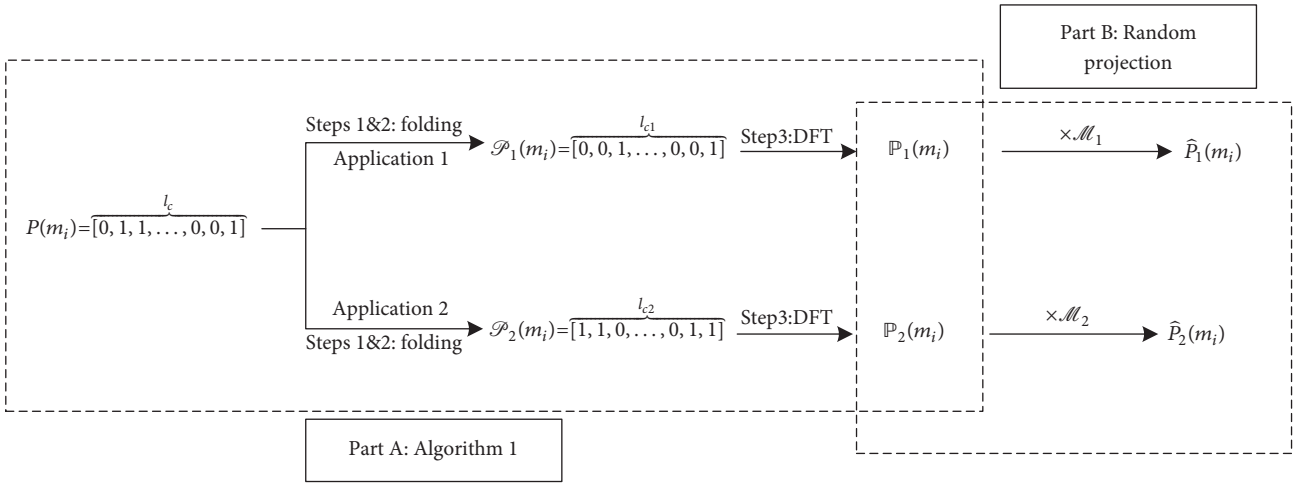


Figure 8: An example of the proposed feature transformation method for the first feature set under the scenario of ARM.

launch the ARM attack by obtaining multiple transformation matrices and transformed feature sets from the same original feature set. Thanks to the feature decorrelation algorithm, Algorithm 1 or 2, the ARM attack can be defended by the proposed method, because the transformed feature set $\mathbb{P}(m_i)$ is uncorrelated in different applications, which is achieved by varying the application-specific parameters, L_C and L_S.

Here we give an example to demonstrate how the proposed method defends the ARM. The whole transformation process is divided into two parts: A and B, as shown in Figure 8. Part A represents the procedure of Algorithm 1 and Part B represents the random projection based transformation. In Part A, with the original feature vector $P(m_i)$ of length l_C, which only contains values of 0 and 1, under the ARM scenario, we assume that the same $P(m_i)$ is used in two applications. Without loss of generality, L_{C1} is chosen for Application 1 and L_{C2} for Application 2, where $L_{C1} \neq L_{C2}$. The folding function in Algorithm 1 is expected to change the input bit stream, yielding two different binary strings for Applications 1 and 2, respectively. The purpose of the DFT (Step 3 in Algorithm 1) is to convert the two new (folded) binary strings into different real-value vectors, i.e., $\mathbb{P}_1(m_i)$ for Application 1 and $\mathbb{P}_2(m_i)$ for Application 2. In Part B, the feature vectors $\mathbb{P}_1(m_i)$ and $\mathbb{P}_2(m_i)$ are, respectively, transformed into $\widehat{P}_1(m_i)$ and $\widehat{P}_2(m_i)$ by random projection. Due to the absence of correlation between $\mathbb{P}_1(m_i)$ and $\mathbb{P}_2(m_i)$, even if the adversary acquires the transformed feature vectors $\widehat{P}_1(m_i)$ and $\widehat{P}_2(m_i)$ as well as the projection matrices $\mathscr{M}_1$ and $\mathscr{M}_2$, he/she cannot launch the ARM to obtain either $\mathbb{P}_1(m_i)$ or $\mathbb{P}_2(m_i)$, because it would be impossible for the ARM to find sufficient number of relevant systems equations matching the number of independent unknown variables.

It is worth noting that the outputs, $\mathbb{P}_1(m_i)$ and $\mathbb{P}_2(m_i)$, of Part A are the inputs to Part B as shown in the overlapped area of Part A and Part B in Figure 8. $\mathbb{P}_1(m_i)$ and $\mathbb{P}_2(m_i)$ are intermediate products, which are not stored in the database or on the smart card, because only the resultant feature vectors $\widehat{P}_1(m_i)$ and $\widehat{P}_2(m_i)$ are needed for similarity score calculation. Therefore, the adversary does not know $\mathbb{P}_1(m_i)$ and $\mathbb{P}_2(m_i)$ unless he/she can recover them. One may argue that if the random projection based transformation (or Part B) is conquered, then $\mathbb{P}(m_i)$ can be retrieved. We now show why this is computationally infeasible. The random projection based transformation in (1) effectively constitutes an underdetermined system of linear equations. Since the projection matrix $\mathscr{M}$ is of size $X \times Y$ with Y being smaller than X, rank($\mathscr{M}$) is no greater than Y, which is less than the number of unknowns, namely, elements of $\mathbb{P}(m_i)$. It is a well-known result in linear algebra [56] that when the coefficient and augmented matrices of (1) have the same rank, (1) has an infinite number of solutions. Clearly, $\mathbb{P}(m_i)$ is just one solution among so many solutions, making the search for $\mathbb{P}(m_i)$ tremendously hard, especially with $\mathbb{P}(m_i)$ having a relatively flat spectrum due to the DFT.

Let us analyze the security provided by Algorithms 1 and 2, under the assumption that $\mathbb{P}(m_i)$ is obtained by adversary. The DFT in Algorithm 1 is invertible and not meant to protect $\mathscr{P}(m_i)$. The DFT is applied for the purpose of rendering a dense data representation of the frequency samples of $\mathscr{P}(m_i)$ so that the search space for $\mathscr{P}(m_i)$ cannot be narrowed down [57]. The modulo operation in Algorithm 1 can provide certain protection to the original feature vector $P(m_i)$ depending on parameter settings. With an element of value '1' found at position f_P in $\mathscr{P}(m_i)$, where $f_P < L_C$, the number of positions in $P(m_i)$ that could yield the value of '1' at position f_P in $\mathscr{P}(m_i)$ is $N_p = \lceil l_C/L_C \rceil$. The number of possibilities that could result in the value of '1' at position f_P in $\mathscr{P}(m_i)$ is $\sum_{k=1}^{N_p} C_{N_p}^k$ in theory, where $C_{N_p}^k$ means choosing k from N_p positions, but, in reality, it is unlikely that a majority of these N_p positions in $P(m_i)$ contain '1'. Without loss of generality, we assume that at most two of these N_p positions are '1' for the rest of our discussions unless stated otherwise; then the computational complexity is $S_e = \log_2(C_{N_p}^1 + C_{N_p}^2)$

bits in order to search all the possible combinations by brute force attack. Suppose that there are N_e '1's in $\mathscr{P}(m_i)$; then the computational complexity for brute force search will be $N_e \times S_e$. When $N_e = 30$, under the parameter settings $l_C = 34560$ and $L_C = 500$, $N_p = \lceil 34560/500 \rceil = 70$, the computational complexity is $30 \times \log_2(C_{70}^1 + C_{70}^2) \approx 330$ bits. However, if L_C is set to be a large value, e.g., $L_C = 20000$, for elements of value '1' found in $\mathscr{P}(m_i)$ from positions 1 to 14560, say for value '1' found at position 123, its original position in $P(m_i)$ only has three possibilities, either position 123 or position 20123 or both. In this case, the number of possibilities provided by the modulo operation is considerably reduced. To strengthen the security of Algorithm 1 for the case that L_C is set to be a large value, e.g., 20000, we propose Algorithm 2 on top of Algorithm 1.

In Algorithm 2, the first part P_1 of original feature vector $P(m_i)$ is first divided into N_S segments and each segment, e.g., S_j, is permutated by a segment-based binary code, e.g., ϕ_j. Take the segment S_j of length L_S as an example. For convenience, trivial segments containing full '0's or full '1's are excluded in the following discussion as permutation does not change the segment bit distribution. For a permuted segment S_j', there exists $L_S - 1$ number of different segments originating from S_j through the circular shift permutation. One-bit difference in one segment implies a different entire feature vector P_1 which contains N_S segments. The computational effort to determine all possible P_1 due to the circular shift permutation will be $(L_S - 1)^{N_S}$ under the brute force attack. Assume there are $N_c < N_S$ nontrivial segments that contain both zero and nonzero elements, then the actual computational complexity reduces to be $(L_S - 1)^{N_c}$ under the brute force attack. Take, for example, $L_S = 5$ and $N_c = 60$, which has been observed in many of our experiments. In this case, the number of security bits imposed by Algorithm 2 is $\log_2((5-1)^{60}) = 120$ bits, which is a substantial amount. Clearly, the added security from Algorithm 2 is at the expense of matching performance, as shown in Section 4.1, which reflects the fact that there is always a compromise between security and recognition accuracy. Note that, for a given permuted segment, it is hard to tell whether it is produced by an actual permutation or it has not been permuted at all. Therefore L_S instead of $L_S - 1$ can be used to account for all possible shift combinations which can help increase the security strength.

Regarding the security of the second feature set D_T, since the feature $\widehat{D}_T$ is obtained from D_T under the guidance of feature codes $\{f_i^c\}_{i=1}^{N_1^T}$, which are generated from their corresponding triangle feature and not saved in the database or on the smart card, the adversary has no idea about the original location of a triangle feature, even if $\widehat{D}_T$ is hacked. Under this situation, the adversary may try to guess the feature code through the brute force attack. The number of security bits for calculating one feature code, e.g., $f_{m_1m_2m_3}^c$, which is given by $f_{m_1m_2m_3}^q = [o_{m_1m_2}^q, l_{m_1m_2}^q, l_{m_1m_3}^q, \alpha_{m_1}^q]$ through (2), can be expressed by $H(\Delta) = \log_2((2\pi/s_o) \times (300/s_l)^2 \times (\pi/s_\alpha))$ because the minutiae orientation range is $[0, 2\pi)$, the edge length is $(0, 300]$, and the range of an angle is $(0, \pi)$. In the experiments, the quantization steps for calculating the feature code were set to be $s_o = 5\pi/36$, $s_l = 25$ pixels and $s_\alpha = 5\pi/36$ on database FVC2002 DB2. With these quantization settings, the number of security bits $H(\Delta)$ is about 14 bits for one feature code. However, this feature code can only be used to find the original location of one triangle in D_T and the adversary cannot verify the correction of the original position. To exactly restore D_T, the original locations of N_1^T triangles have to be determined at the same time, which means that the security of D_T is $N_1^T \times 14$ bits. The average value of N_1^T is 32 on database FVC2002 DB2. Since a certain degree of error tolerance is allowed in the matching process, the real security is lower than $N_1^T \times 14$ bits, depending on the matching score threshold S_t. Note that, even if D_T is revealed, the original location of minutiae in the fingerprint image remains unavailable, because D_T only contains the relative information of the three minutiae forming the Delaunay triangle rather than the absolute coordinates on the image. This means that compromising D_T would not threaten the security of C_T.

Some other issues are discussed here: (1) In our application, polar coordinate-based local structures are employed. Assume that m_i and m_j are two neighbor minutiae and are considered as the origin of the polar coordinate of two polar coordinate-based local structures S_{m_i} and S_{m_j}. Since m_i and m_j are neighbor minutiae, some of the minutiae points included in local structure S_{m_i} are also included in S_{m_j}. The feature vectors, e.g., $P(m_i)$ and $P(m_j)$, extracted from local structures, S_{m_i} and S_{m_j}, respectively, may have some correlation; however, they are different. Here 'correlation' indicates the existence of some common minutiae in different polar coordinate-based structures, but it does not mean that the features extracted from different polar coordinate-based structures are the same, because the features extracted from different polar coordinate-based structures are based on different reference minutiae and the information of the reference minutiae is nonpublic. For instance, according to Section 2.1, the positions of '1' in $P(m_i)$ and $P(m_j)$ are decided by the triplet values $(\rho_{ij}, \alpha_{ij}, \beta_{ij})$ and $(\rho_{ji}, \alpha_{ji}, \beta_{ji})$. The radial distance ρ_{ij} is equal to ρ_{ji}, but α_{ij} and β_{ij} are different from α_{ji} and β_{ji}, respectively. A moderate variation on the orientation or/and its related line angle can change the values of $(\rho_{ij}, \alpha_{ij}, \beta_{ij})$ or $(\rho_{ji}, \alpha_{ji}, \beta_{ji})$. So even if minutiae m_i and m_j are neighbor minutiae, the triplet values $(\rho_{ij}, \alpha_{ij}, \beta_{ij})$ are different from $(\rho_{ji}, \alpha_{ji}, \beta_{ji})$, which leads to different feature vectors $P(m_i)$ and $P(m_j)$. As the radial distance ρ_{ij} is equal to ρ_{ji}, so $P(m_i)$ and $P(m_j)$ may have some correlation that can be utilized by the adversary. However, any correlation based attack needs at least a compromised feature vector, e.g., $P(m_i)$ or $P(m_j)$, as a base, which unfortunately is infeasible because $P(m_i)$ or $P(m_j)$ has been decorrelated and protected by the folding, permutation, and random projection. (2) Assume the number of '1's, in feature vector $P(m_i)$, is N_{1s}. According to our statistics on the databases, N_{1s} is 34, which is much smaller than the length l_C (e.g., $l_C = 34560$) of the feature vector $P(m_i)$. If these N_{1s}'1's are uniformly distributed on $P(m_i)$, the computation complexity is about $\log_2(C_{34560}^{34}) =$

384 bits in order to correctly recover all the 34 '1's in $P(m_i)$ by brute force attack. One may claim that these binary '1's are not uniformly distributed. We give an example here. Assume that all the 34 '1's are only located in the first 1000 bins out of a total of 34560 bins of the feature vector $P(m_i)$. Although this is a highly unlikely case, computational complexity for this example is about $\log_2(C_{1000}^{34}) = 210$ bits under the brute force attack, which demonstrates that the proposed method is still secure enough. Actually, it is nearly impossible if one can generate a binary string that is absolutely uniform. Our statistics of the probability distribution of bit '1' over unfolded raw features, e.g., $P(m_i)$, shows that the features are with a good degree of uniform distribution. For example, the maximum probability of being '1' of a bin is 10^{-4}, which shows that dominant probability does not exist. Moreover, even after removing bins with probability being '1' that are equal to or less than 20% of the maximum probability, the remaining bins are still more than a thousand. The estimated figures, e.g., 330, 120, and 384 bits, in our security analysis, may be not tight. However, these could be the best analytical results one can get unless an accurate distribution model can be developed, which is far beyond the scope of this paper as such distribution model can be a research topic by its own. (3) There is a type of statistical attack, named Moore-Penrose inverse [58], to find possible solutions of the linear system (random projection is a linear system). If the linear system $b = Ax$ has any solutions, they are given by $x' = A^+b$, where A^+ is the Moore-Penrose inverse of A. It is well known that a Moore-Penrose inverse based solution x' has the minimum Euclidean norm $\|x'\|_2$. There is no theoretic basis for any deterministic relationship between the Moore-Penrose inverse based solution x' and the ground truth x, if A is random. Some probabilistic relationship might exist for some cases, for example, based on the values of the Moore-Penrose inverse based solution x', one may estimate the position of '1's in binary vector in x by general statistical relation that large values in x' corresponding to '1's in x. However, for a specific application, the defender can select suitable random projection matrix, e.g., A, so that this statistical relation can be lessened. We have experimentally verified that a solution x' that is based on the Moore-Penrose inverse incorporating the statistical relation can find only one correct position of '1' in x.

4.5. Suitability/Feasibility on Mobile Devices. In this section, the template size and computational complexity of the proposed system are discussed. In our application, the template consists of two parts. One is the transformed polar coordinate-based feature set $\widehat{C}_T = \{\widehat{P}(m_i)\}_{i=0}^{N^T-1}$ and the other one is the transformed Delaunay triangulation-based feature set $\widehat{D}_T$. The size of each element $\widehat{P}(m_i)$ in $\widehat{C}_T$ depends on the parameter setting Y of the transformation matrix. If Y is set to be 300, the size of $\widehat{P}(m_i)$ is 4.7 KB, and then the size of $\widehat{C}_T$ is $4.7 \times N^T$ KB. The feature set $\widehat{D}_T$ is of size $l_D = 20000$ bits $\approx$ 2.5KB. In terms of matching time, it takes about 0.000171 seconds to match an element $\widehat{P}^T(m_i)$ from $\widehat{C}_T$ against an element $\widehat{P}^Q(m_j)$ from $\widehat{C}_Q$. Assume that there are N^T elements in template feature set $\widehat{C}_T$ and N^Q elements in query feature set $\widehat{C}_Q$, then the matching time between $\widehat{C}_T$ and $\widehat{C}_Q$ is $N^Q \times N^T \times 0.000171$ seconds. The matching time between $\widehat{D}_T$ and $\widehat{D}_Q$ is about 0.00872 seconds.

Take the fingerprint image 1_1.tif and image 1_2.tif in FVC2002 DB2 for example. Assume that image 1_1.tif is template image and 1_2.tif is the query image. The value of N^T is 31 in the template, so the template size is 148.2 ($=4.7 \times 31 + 2.5$) KB. The value of N^Q is 33 in the query, so the matching time between the template and query is 0.18263 ($=31 \times 33 \times 0.00017 + 0.00872$) seconds. The above experiment is conducted using MATLAB on a laptop with Intel processor: i5-2450M dual-core CPU of 2.50 GHz, 2.50 GHz, RAM of 8GB, and Operation System of 64-bit Win 7. With the powerful storage and computing capability of today's mobile devices, e.g., smartphones, the imposed storage size and computational load are never an issue.

5. Conclusion

In this paper, a new fingerprint cancelable template system has been proposed for mobile device authentication. The new cancelable template can mitigate the negative impact of nonlinear distortion by combining multiple feature sets. Since the proposed system with two feature sets contains more feature information than most existing cancelable templates with just a single feature set, the discriminative power of the proposed system gets increased and thus recognition performance is enhanced. Furthermore, the proposed method can defeat the ARM attack through eliminating the feature correlation in different applications, which is a clear advantage over those existing cancelable templates that are vulnerable to the ARM attack. In the future work, the proposed scheme will be applied to other biometrics such as cancelable palmprint [59].

Conflicts of Interest

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

Acknowledgments

This paper is supported by the Natural Science Foundation of China through Projects 61672083, 61370190, 61532021, 61472429, and 61402029 and by Defence Science and Technology Group (DST) of Australia through Project CERA 221.

References

[1] R. Spolaor, Q. Li, M. Monaro, M. Conti, L. Gamberini, and G. Sartori, "Biometric authentication methods on smartphones: A survey," *PsychNology Journal*, vol. 14, no. 2-3, pp. 87–98, 2016.

[2] A. K. Jain, L. Hong, S. Pankanti, and R. Bolle, "An identity-authentication system using fingerprints," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1365–1388, 1997.

[3] V. M. Patel, N. K. Ratha, and R. Chellappa, "Cancelable biometrics: A review," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 54–65, 2015.

[4] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, "Generating cancelable fingerprint templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 561–572, 2007.

[5] W. Yang, J. Hu, S. Wang, and J. Yang, "Cancelable Fingerprint Templates with Delaunay Triangle-Based Local Structures," in *Cyberspace Safety and Security*, vol. 8300 of *Lecture Notes in Computer Science*, pp. 81–91, Springer International Publishing, Cham, 2013.

[6] D. Ahn, S. G. Kong, Y. Chung, and K. Y. Moon, "Matching with Secure Fingerprint Templates Using Non-invertible Transform," in *Proceedings of the 2008 Congress on Image and Signal Processing*, pp. 29–33, Sanya, China, May 2008.

[7] B. Yang and C. Busch, "Parameterized geometric alignment for minutiae-based fingerprint template protection," in *Proceedings of the IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems, BTAS 2009*, usa, September 2009.

[8] B. Yang, D. Hartung, K. Simoens, and C. Busch, "Dynamic random projection for biometric template protection," in *Proceedings of the 4th IEEE International Conference on Biometrics: Theory, Applications and Systems, BTAS 2010*, usa, September 2010.

[9] T. Ahmad, J. Hu, and S. Wang, "Pair-polar coordinate-based cancelable fingerprint templates," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2555–2564, 2011.

[10] S. Wang and J. Hu, "Alignment-free cancelable fingerprint template design: A densely infinite-to-one mapping (DITOM) approach," *Pattern Recognition*, vol. 45, no. 12, pp. 4129–4137, 2012.

[11] S. Wang and J. Hu, "Design of alignment-free cancelable fingerprint templates via curtailed circular convolution," *Pattern Recognition*, vol. 47, no. 3, pp. 1321–1329, 2014.

[12] S. Wang, G. Deng, and J. Hu, "A partial Hadamard transform approach to the design of cancelable fingerprint templates containing binary biometric representations," *Pattern Recognition*, vol. 61, pp. 447–458, 2017.

[13] Z. Jin, B.-M. Goi, A. Teoh, and Y. H. Tay, "A two-dimensional random projected minutiae vicinity decomposition-based cancellable fingerprint template," *Security and Communication Networks*, vol. 7, no. 11, pp. 1691–1701, 2014.

[14] Z. Jin, M.-H. Lim, A. B. J. Teoh, and B.-M. Goi, "A non-invertible Randomized Graph-based Hamming Embedding for generating cancelable fingerprint template," *Pattern Recognition Letters*, vol. 42, no. 1, pp. 137–147, 2014.

[15] M. Ferrara, D. Maltoni, and R. Cappelli, "A two-factor protection scheme for MCC fingerprint templates," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG)*, pp. 1–8, 2014.

[16] N. L. Clarke and S. M. Furnell, "Advanced user authentication for mobile devices," *Computers & Security*, vol. 26, no. 2, pp. 109–119, 2007.

[17] D.-J. Kim and K.-S. Hong, "Multimodal biometric authentication using teeth image and voice in mobile environment," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1790–1797, 2008.

[18] D.-J. Kim, K.-W. Chung, and K.-S. Hong, "Person authentication using face, teeth and voice modalities for mobile device security," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2678–2685, 2010.

[19] H. C. Lee, K. R. Park, B. J. Kang, and S. J. Park, "A new mobile multimodal biometric device integrating finger vein and fingerprint recognition," in *Proceedings of the 4th International Conference on Ubiquitous Information Technologies and Applications, ICUT 2009*, jpn, December 2009.

[20] Q. Tao and R. Veldhuis, "Biometric authentication system on mobile personal devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 763–773, 2010.

[21] C.-L. Chen, C.-C. Lee, and C.-Y. Hsu, "Mobile device integration of a fingerprint biometric remote authentication scheme," *International Journal of Communication Systems*, vol. 25, no. 5, pp. 585–597, 2012.

[22] J. S. Kim, G. Li, B. Son, and J. Kim, "An empirical study of palmprint recognition for mobile phones," *IEEE Transactions on Consumer Electronics*, vol. 61, no. 3, pp. 311–319, 2015.

[23] A. Rattani, N. Reddy, and R. Derakhshani, "Convolutional neural networks for gender prediction from smartphone-based ocular images," *IET Biometrics*.

[24] M. De Marsico, M. Nappi, F. Narducci, and H. Proença, "Insights into the results of MICHE I - Mobile Iris CHallenge Evaluation," *Pattern Recognition*, vol. 74, pp. 286–304, 2018.

[25] T. Neal and D. Woodard, "Surveying Biometric Authentication for Mobile Device Security," *Journal of Pattern Recognition Research*, vol. 11, no. 1, pp. 74–110, 2016.

[26] A. Wojciechowska, M. Choraś, and R. Kozik, "The overview of trends and challenges in mobile biometrics," *Journal of Applied Mathematics and Computational Mechanics*, vol. 16, no. 2, pp. 173–185, 2017.

[27] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

[28] Y. Wang, J. Hu, and D. Phillips, "A fingerprint orientation model based on 2D fourier expansion (FOMFE) and its application to singular-point detection and fingerprint indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 573–585, 2007.

[29] X. Chen, J. Tian, X. Yang, and Y. Zhang, "An algorithm for distorted fingerprint matching based on local triangle feature set," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 169–177, 2006.

[30] A. Gago-Alonso, J. Hernández-Palancar, E. Rodríguez-Reina, and A. Muñoz-Briseño, "Indexing and retrieving in fingerprint databases under structural distortions," *Expert Systems with Applications*, vol. 40, no. 8, pp. 2858–2871, 2013.

[31] F. Farooq, R. M. Bolle, T.-Y. Jea, and N. Ratha, "Anonymous and revocable fingerprint recognition," in *Proceedings of the 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'07*, usa, June 2007.

[32] C. Lee, J.-Y. Choi, K.-A. Toh, S. Lee, and J. Kim, "Alignment-free cancelable fingerprint templates based on local minutiae information," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 4, pp. 980–992, 2007.

[33] S. Tulyakov, F. Farooq, P. Mansukhani, and V. Govindaraju, "Symmetric hash functions for secure fingerprint biometric systems," *Pattern Recognition Letters*, vol. 28, no. 16, pp. 2427–2436, 2007.

[34] Z. Jin, A. B. J. Teoh, T. S. Ong, and C. Tee, "A Revocable Fingerprint Template for Security and Privacy Preserving," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 6, pp. 1327–1342, 2010.

[35] C. Lee and J. Kim, "Cancelable fingerprint templates using minutiae-based bit-strings," *Journal of Network and Computer Applications*, vol. 33, no. 3, pp. 236–246, 2010.

[36] H. Chen and H. Chen, "A novel algorithm of fingerprint encryption using minutiae-based transformation," *Pattern Recognition Letters*, vol. 32, no. 2, pp. 305–309, 2011.

[37] Z. Jin, A. B. Jin Teoh, T. S. Ong, and C. Tee, "Fingerprint template protection with minutiae-based bit-string for security and privacy preserving," *Expert Systems with Applications*, vol. 39, no. 6, pp. 6157–6167, 2012.

[38] P. Das, K. Karthik, and B. Chandra Garai, "A robust alignment-free fingerprint hashing algorithm based on minimum distance graphs," *Pattern Recognition*, vol. 45, no. 9, pp. 3373–3388, 2012.

[39] W. J. Wong, A. B. J. Teoh, M. L. D. Wong, and Y. H. Kho, "Enhanced multi-line code for minutiae-based fingerprint template protection," *Pattern Recognition Letters*, vol. 34, no. 11, pp. 1221–1229, 2013.

[40] N. Zhang, X. Yang, Y. Zang, X. Jia, and J. Tian, "Generating registration-free cancelable fingerprint templates based on Minutia Cylinder-Code representation," in *Proceedings of the 6th IEEE International Conference on Biometrics: Theory, Applications and Systems, BTAS 2013*, usa, October 2013.

[41] S. Wang and J. Hu, "A blind system identification approach to cancelable fingerprint templates," *Pattern Recognition*, vol. 54, pp. 14–22, 2016.

[42] M. Ferrara, D. Maltoni, and R. Cappelli, "Noninvertible minutia cylinder-code representation," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1727–1737, 2012.

[43] H. Kaur and P. Khanna, "Random Slope method for generation of cancelable biometric features," *Pattern Recognition Letters*, 2018.

[44] P. P. Paul, M. Gavrilova, and S. Klimenko, "Situation awareness of cancelable biometric system," *The Visual Computer*, vol. 30, no. 9, pp. 1059–1067, 2014.

[45] W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, "A fingerprint and finger-vein based cancelable multi-biometric system," *Pattern Recognition*, vol. 78, pp. 242–251, 2018.

[46] F. Quan, S. Fei, C. Anni, and Z. Feifei, "Cracking cancelable fingerprint template of Ratha," in *Proceedings of the International Symposium on Computer Science and Computational Technology, ISCSCT 2008*, pp. 572–575, chn, December 2008.

[47] C. Li and J. Hu, "Attacks via record multiplicity on cancelable biometrics templates," *Concurrency and Computation: Practice and Experience*, vol. 26, no. 8, pp. 1593–1605, 2014.

[48] C. Wang and M. L. Gavrilova, "Delaunay triangulation algorithm for fingerprint matching," in *Proceedings of the 3rd International Symposium on Voronoi Diagrams in Science and Engineering 2006, ISVD 2006*, pp. 208–216, can, July 2006.

[49] R. Soleymani and M. Chehel Amirani, "A hybrid fingerprint matching algorithm using Delaunay triangulation and Voronoi diagram," in *Proceedings of the 20th Iranian Conference on Electrical Engineering, ICEE 2012*, pp. 752–757, irn, May 2012.

[50] A. Muñoz-Briseño, A. Gago-Alonso, and J. Hernández-Palancar, "Fingerprint indexing with bad quality areas," *Expert Systems with Applications*, vol. 40, no. 5, pp. 1839–1846, 2013.

[51] D.-T. Lee and B. J. Schachter, "Two algorithms for constructing a Delaunay triangulation," *International Journal of Computer & Information Sciences*, vol. 9, pp. 219–242, 1980.

[52] W. K. Gu, "Matching Perspective Views of a Polyhedron Using Circuits," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 3, pp. 390–400, 1987.

[53] S. D. K. VeriFinger, *Neuro Technology*, 2010, http://www.neurotechnology.com/verifinger.html.

[54] K. Nandakumar and A. K. Jain, "Biometric template protection: Bridging the performance gap between theory and practice," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 88–100, 2015.

[55] K. Simoens, B. Yang, X. Zhou et al., "Criteria towards metrics for benchmarking template protection algorithms," in *Proceedings of the 5th IAPR International Conference on Biometrics, ICB '12*, pp. 498–505, India, April 2012.

[56] E. Kreyszig, *Advanced engineering mathematics*, John Wiley and sons, Inc., New York, NY, USA, 2nd edition, 2010.

[57] M. Unser, "On the approximation of the discrete Karhunen-Loeve transform for stationary processes," *Signal Processing*, vol. 7, no. 3, pp. 231–249, 1984.

[58] "Moore–Penrose inverse," https://en.wikipedia.org/wiki/Moore%E2%80%93Penrose_inverse.

[59] L. Leng and A. B. J. Teoh, "Alignment-free row-co-occurrence cancelable palmprint Fuzzy Vault," *Pattern Recognition*, vol. 48, no. 7, pp. 2290–2303, 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 7978027, 13 pages
https://doi.org/10.1155/2018/7978027

Research Article

Secure Certificateless Authentication and Road Message Dissemination Protocol in VANETs

Haowen Tan[1], **Dongmin Choi**[2], **Pankoo Kim**[1], **Sungbum Pan**[3] and **Ilyong Chung**[1]

[1]*Department of Computer Engineering, Chosun University, 309 Pilmun-daero, Seonam-dong, Dong-gu, Gwangju 61452, Republic of Korea*
[2]*Division of Undeclared Majors, Chosun University, 309 Pilmun-daero, Seonam-dong, Dong-gu, Gwangju 61452, Republic of Korea*
[3]*Department of Electronic Engineering, Chosun University, 309 Pilmun-daero, Seonam-dong, Dong-gu, Gwangju 61452, Republic of Korea*

Correspondence should be addressed to Ilyong Chung; iyc@chosun.ac.kr

Received 17 January 2018; Revised 21 March 2018; Accepted 10 April 2018; Published 20 May 2018

Academic Editor: Ding Wang

As a crucial component of Internet-of-Thing (IoT), vehicular ad hoc networks (VANETs) have attracted increasing attentions from both academia and industry fields in recent years. With the extensive VANETs deployment in transportation systems of more and more countries, drivers' driving experience can be drastically improved. In this case, the real-time road information needs to be disseminated to the correlated vehicles. However, due to inherent wireless communicating characteristics of VANETs, authentication and group key management strategies are indispensable for security assurance. Furthermore, effective road message dissemination mechanism is of significance. In this paper, we address the above problems by developing a certificateless authentication and road message dissemination protocol. In our design, certificateless signature and the relevant feedback mechanism are adopted for authentication and group key distribution. Subsequently, message evaluating and ranking strategy is introduced. Security analysis shows that our protocol achieves desirable security properties. Additionally, performance analysis demonstrates that the proposed protocol is efficient compared with the state of the art.

1. Introduction

Vehicular ad hoc networks (VANETs) are distributed, self-organized wireless networks constructed by vehicles and nearby road-side units (RSUs). The real-time dynamic communication enables efficient and durative information exchange between vehicles and RSUs. Hence, intelligent transportation system (ITS) is achievable with the widely implementation of VANETs [1, 2]. A variety of VANET-based applications, which can be mainly classified into safety-related applications and commercial-oriented applications, not only enhance the driving safety but also provide better driving experience. Typical safety-related applications include emergency vehicle warnings, traffic congestion report, road accident informing, and speed monitoring [3, 4]. Commercial-oriented applications provide convenience service and entertainment applications such as weather forecast, information broadcasting of nearby petrol stations and restaurants, navigation, and Internet access.

In general, a basic VANET consists of three important components: the trusted authority (TA), road-side units (RSUs), and vehicles [5–7]. Considered as both the application provider and key server, TA is responsible for providing various services to vehicles through RSUs. Moreover, pivotal secret key assignment, along with the user management for correlated vehicles, is conducted by TA. The RSUs are deployed by the road sides one after another. Commonly, RSUs are built aside the road in every kilometer. Thus the effective range of VANET system can cover each section of the road. In this case, RSUs are considered as the communication bridge connecting TA and vehicles, which provides timely transmission of vital personal data. To a certain degree, RSUs have the capability of conducting computation and storing essential information in its memories [8]. The vehicle

performs as both terminal customer and information collector. In other words, useful information including traffic congestion and emergency road condition is forwarded to the corresponded RSU. Each vehicle is equipped with an onboard unit (OBU), which conducts all the computation and communication [9, 10]. Compared with regular wireless sensor networks [11], vehicle's high mobility is the unique characteristic of VANETs.

In VANETs, the data exchange between TA and RSUs are via secure wired connection, where the adopted cryptographic strategies guarantee transmission security and message confidentiality. Meanwhile, vehicle-to-vehicle (V2V) communication and vehicle-to-RSU (V2R) communication are conducted through open wireless channel, which employs the dedicated short-range communications (DSRC) [12–14]. On the one hand, the moving vehicle can carry out interactive data exchange with specific RSU through V2R communication. On the other hand, one vehicle is capable of sharing essential messages with other vehicles in its vicinity through V2V communication. In this way, a VANET with high connectivity can be built accordingly [15].

As a particular variant of wireless sensor networks, apparently the VANETs suffer from multiple charted and uncharted security attacks [9, 14]. In V2V and V2R communications, the transmitted messages may be eavesdropped, blocked, or even forged by malicious devices. Hence, significant user information is revealed to the attacker accordingly, which compromises the whole VANET and brings severe user privacy disclosure issue [1, 13]. Under this circumstance, proper authentication strategies are required so as to provide security and privacy assurance. Moreover, high mobility feature of the vehicles brings uncertainty to the communication process, which should also be taken into consideration.

Among the aforementioned safety-related applications, road message dissemination is one of the essential functions for VANET [16]. With the assistance of RSUs and remote server, the vehicles of the same VANET could share necessary driving-related information with each other. By analyzing the acquired traffic information of current areas, the driver is able to make better driving decisions such as choosing the best navigation route ahead of time. Furthermore, occurrences of road accidents and traffic jams can be drastically reduced [9]. Thus, the drivers' driving experience is improved.

In VANETs, typical road message management strategies are mainly composed of information collection and dissemination. First, the road messages are reported by the participating vehicles through OBUs [17]. Afterwards, the acquired messages will be processed and then disseminated to the legitimate vehicles. In some VANETs scenarios, TA arranges all the road messages collected from RSUs via wired transmission [18]. Meanwhile, in decentralized VANETs scenarios, most of the computation and storing are done in the RSU side [7], while TA performs as the key generation center (KGC). Consequently, in particular dense scenarios with large amounts of emerging vehicles, the decentralized architecture could reduce the computation overload and storage complexity in TA.

As described above, authentication strategies are necessary during road message dissemination [4, 19]. Furthermore, the characteristic group communication between RSU and vehicles is indispensable, which enables convenient data exchange. In this case, the group key shared between RSU and all the legitimate vehicles is required. Note that the group key distribution should be conducted after mutual authentication [6, 20].

As for message dissemination in practical VANETs occasions, two channels are required [21], namely, the official channel and normal channel. Official channel is provided by governmental agencies, where the broadcast road information is precise and trustworthy. Note that this channel is assumed to be based on real-time monitoring with satellites and road cameras. Thus it is precise and trustworthy. Meanwhile, normal channel is the more ordinary way, where the road information is gathered from normal vehicles. In this case, some of the vehicles are assumed to be benign devices which transmit precise messages, while the rest are negative vehicles [22]. Note that the negative vehicles may report trivial or even false information to the VANET system. For this consideration, with the purpose of guaranteeing the dissemination exactitude, impartial and effective message evaluation mechanism is necessary [23]. For instance, in extreme scenarios with massive road messages to be disseminated, before dissemination, it is necessary for the RSUs to aggregate and evaluate the acquired road messages before dissemination. Hence the RSUs could broadcast in a particular sequence according to the significance and reliability of each message. Urgent and authentic road messages can be broadcast in the first place.

During the message dissemination of the entire VANETs, the accuracy and efficiency of message dissemination closely depend on the participating vehicles [15, 24]. Hence, it is vital to deploy appropriate rewarding strategy so as to motivate the drivers' enthusiasm on reporting [20, 25]. For example, coupons or discounts on certain commodities can be granted to the trustworthy drivers with timely and precise reporting records. In other words, the drivers are encouraged with incentives, which is of great benefit to the entire VANET.

In this paper, we propose a secure certificateless authentication and road message dissemination protocol in vehicular ad hoc networks. Our nontrivial efforts can be summarized as follows:

(i) Secure Certificateless Authentication Scheme for Group Key Distribution. With the purpose of enhancing transmission security, we adopt the bilinear pairing based on elliptic curve into our authentication scheme. Hence, the active vehicles within the effective range can be identified and then allocated with the group key. The proposed scheme yields desirable security properties.

(ii) Road Message Priority Management and Dissemination Mechanism. The encrypted road messages are delivered to the corresponding RSU. The received road message is evaluated based on both the vehicle priority and the assessment. In this way, accuracy and efficiency of the messages dissemination process are provided. Hence, the drivers can timely arrange their routes according to the delivered road information.

(iii) Security and Performance Analysis. The formal security analysis is provided, involving some necessary proofs on resistance to the existing malicious attacks. Furthermore, performance analysis emphasizing the transmission overload and computation cost is hereby presented.

The remainder of this paper is organized as follows. Section 2 provides brief description of the related research achievements. Section 3 introduces some necessary preliminary works and the designed system model in order for the reader to obtain better understanding of this topic. Section 4 presents the proposed secure certificateless authentication scheme in detail. Section 5 describes the proposed road message dissemination scheme. Section 6 demonstrates the security analysis. Section 7 displays the performance analysis. The conclusion is drawn in Section 8.

2. Related Work

In order to provide enhanced authentication and secure transmission in VANETs, various cryptographic techniques have been deployed in existing researches [2, 3, 6, 16, 26–28]. In 2009, Studer et al. [3] developed a hybrid VANET authentication mechanism (VAST) based on the elliptic curve digital signature algorithm (ECDSA) [29] and TESLA [30] with the purpose of providing fast and extensible authentication and nonrepudiation. Subsequently, emphasizing group authentication and conditional privacy, Zhang et al. [26] proposed a scalable decentralized group authentication protocol, where certain vehicle is able to verify anonymous messages from neighboring vehicles. Motivated by chameleon hash signature based on elliptic curve, in 2011, Huang et al. [27] designed pseudonymous authentication-based conditional privacy protocol (PACP), which adopts the pseudonyms for anonymous communication. Similarly, ABAKA [6] with batch verification was proposed by Huang et al. After that, Lu et al. [28] presented a dynamic privacy-preserving key management scheme (DIKE) enabling both vehicle anonymous authentication and double-registration detection. Guo et al. [2] designed a privacy-preserving anonymous authentication protocol with vehicle unlinkability and authority trackability in 2014, where high efficiency and desired security properties can be achieved accordingly. Afterwards, multiple authentication and group key management protocols in VANETs have been designed recently [8, 31].

Specifically, identity-based encryption, which was first presented by Shamir [33] for certificate management of KGC, has been widely implemented in VANETs authentication protocols. In 2007, Lin et al. [34] combined group signature with identity-based cryptography in the proposed GSIS protocol. Hence, appropriate traceability toward specific vehicle is achieved. After that, Zhang et al. [15] designed an identity-based batch signature verification scheme in VANETs, where multiple signatures can be simultaneously verified in one RSU. Nevertheless, this scheme suffers from replay attack [35]. Subsequently, Sun et al. [5] constructed an identity-based security framework in order to address the misbehavior issue in VANET system. In 2012, Shim [9] developed an identity-based conditional privacy-preserving authentication scheme (CPAS) supporting fast batch verification. However, the proposed protocol is vulnerable to modification attack [36]. Another signature scheme for VANETs, named EIBS [37], was proposed in 2015, where the RSUs perform as the certificate verifiers in order to decrease the computation overload in TA side. Moreover, the anonymity of the legitimate vehicle is provided by using pseudo identity instead of real identity. Hence, the vehicle privacy is preserved. Aiming to decrease the computational complexity, He et al. [10] designed an identity-based conditional privacy-preserving authentication scheme in VANETs. With relatively limited computation and communication requirements, the proposed protocol is suitable for practical VANETs applications.

With the purpose of addressing the key escrow issue in identity-based public key cryptography system (ID-PKC), certificateless public key cryptography (CL-PKC) was first introduced by Al-Riyami and Paterson [38] in 2003. In CL-PKC design, the private partial keys are, respectively, generated by the semitrusted key generation center (KGC) and the user itself. Multiple certificateless authentication protocols were proposed afterwards [25, 39]. Thereafter, Li and Wang [17] presented a fast certificateless authentication scheme (RCS) employing bilinear pairing, where particular vehicles are selected as the assistance to the relevant RSUs. In this case, the transmission overload can be alleviated. Afterwards, Xiong and Qin proposed a certificateless encryption scheme and another certificateless signature scheme with efficient revocation against short-term key exposure in [40]. In 2016, Peng [1] designed an anonymous authentication protocol based on certificateless signature scheme, which provides conditional privacy and mutual authentication.

Furthermore, as the crucial and unique feature of VANETs, message dissemination has been studied due to its promising advantages in both safety-related and commercial-oriented applications. Focusing on commercial advertisements dissemination, Tseng et al. [7] adopted Reed-Solomon Code in the incentive scheme through interactions between vehicles. Similarly, a cooperative message authentication scheme [18] is developed to alleviate the verification overload in the RSU side, where the legitimate vehicles are responsible for message verification in the vicinity. Thereafter, in order to achieve high reliability and low dissemination delay at the same time, density-aware emergency message extension protocol (DEEP) [22] is constructed. As illustrated, emergency warning messages can be timely delivered to all the vehicles within the operating range, which could drastically improve driving safety. As one of the significant services offered by VANETs, RSU-assisted navigation is studied by Chim et al. in [20]. In the assumption, the real-time road conditions are used to compute a better route for the requesting vehicles. The privacy of the drivers can be protected with the advantages of anonymous credential. In [23], Milojevic and Rakocevic developed a location-aware data aggregation mechanism for real-time observation and efficient message dissemination. The communication cost is minimized with the use of intelligent passive clustering and adaptive broadcasting. For improving the accuracy of the delivered message, the aggregated information is arranged by real-time spatiotemporal database refreshing. Recently, Liu et al. presented a cloud-assisted message downlink dissemination scheme (CMDS)

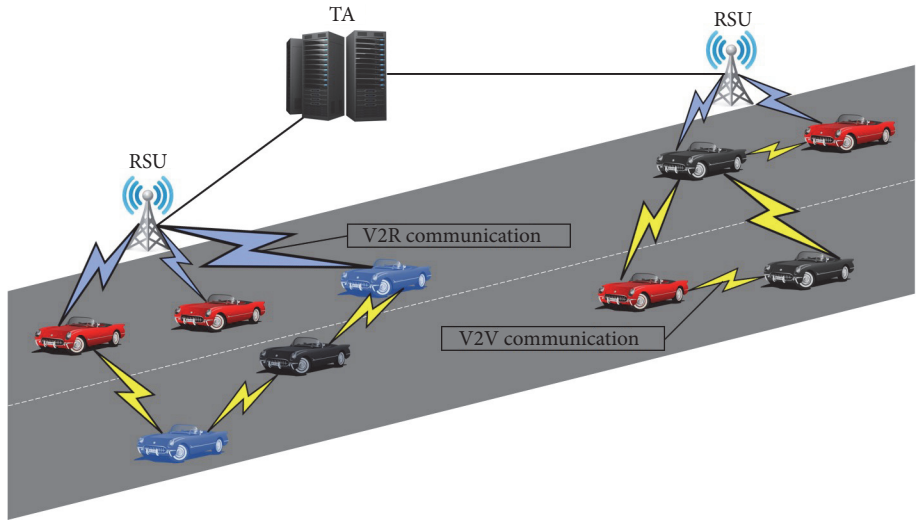


Figure 1: System Model.

under a developed VANET-cellular heterogeneous framework combining cloud computing [21, 41].

3. Model Definition and Preliminaries

In this section, some necessary preliminaries are introduced with the purpose of facilitating the readers' understanding, including the definition of bilinear pairing and hash function. Subsequently, the corresponding notations, the system model, and network assumptions are illustrated.

3.1. Bilinear Pairing. Let $\mathbb{G}_{\mathcal{O}}$ and $\mathbb{G}_{\mathcal{N}}$ be two additive cyclic groups of a large prime order $\mathcal{P}$. A map function $\hat{e} : \mathbb{G}_{\mathcal{O}} \times \mathbb{G}_{\mathcal{O}} \rightarrow \mathbb{G}_{\mathcal{N}}$ is a bilinear pairing if it satisfies the three properties below:

(1) Bilinear: for $\forall M, N \in \mathbb{G}_{\mathcal{O}}$ and $\forall a, b \in \mathbb{Z}^*_{\mathcal{P}}$, there is $\hat{e}(aM, bN) = \hat{e}(M, N)^{ab}$. In addition, for $\forall M, N, Y \in \mathbb{G}_{\mathcal{O}}$, there are $\hat{e}(M + N, Y) = \hat{e}(M, Y)\hat{e}(N, Y)$ and $\hat{e}(M, N + Y) = \hat{e}(M, N)\hat{e}(M, Y)$.

(2) Nondegeneracy: for $\exists M, N \in \mathbb{G}_{\mathcal{O}}$, there is $\hat{e}(M, N) \neq 1$.

(3) Computability: for $\forall M, N \in \mathbb{G}_{\mathcal{O}}$, there is an efficient algorithm to compute $\hat{e}(M, N)$.

In order to prove the security of our schemes, the following intractable problems are briefly presented as

(1) discrete logarithm problem (DLP): for $\forall M, N \in \mathbb{G}_{\mathcal{O}}$, it is difficult to find an integer $a \in \mathbb{Z}^*_{\mathcal{P}}$, such that $M = aN$ holds;

(2) computational Diffie-Hellman problem (CDHP): for $\forall M, aM, bM \in \mathbb{G}_{\mathcal{O}}$, it is difficult to compute abM;

(3) decisional Diffie-Hellman problem (DDHP): for $\forall M, aM, bM, cM \in \mathbb{G}_{\mathcal{O}}$ and $\forall a, b, c \in \mathbb{Z}^*_{\mathcal{P}}$, it is difficult to decide whether $c = ab \bmod \mathcal{P}$ holds;

(4) pairing inversion problem (PIP): for a pairing $\hat{e}$ and $\forall c \in \mathbb{Z}^*_{\mathcal{P}}$, it is difficult to find $M, N \in \mathbb{G}_{\mathcal{O}}$, such that $\hat{e}(M, N) = c$ holds.

3.2. Hash Function. A one-way hash function is considered to be secure if the following properties can be satisfied [42]:

(1) Inputting a message x of arbitrary length, it is easy to compute a message digest of a fixed length output $h(x)$.

(2) Given y, it is difficult to compute $x = h^{-1}(y)$.

(3) Given x, it is computationally infeasible to find $x' \neq x$ such that $h(x') = h(x)$ holds.

3.3. Notations. The notations and the brief description are listed in Notations.

3.4. System Model. The structure of VANET system of our design is shown in Figure 1, where the whole VANET system is composed of three entities: the trusted authority (TA), the road-side units (RSUs), and the vehicles. Descriptions of these entities are, respectively, illustrated below.

Trusted Authority (TA) is a trustworthy management center in charge of all the involving RSUs and vehicles. The vital system operations, including vehicle registration, assignment, and secret key generation, are all conducted by TA. Additionally, TA stores the significant user data in its memory. Hence, TA is assumed to have adequate storage and computing capability. Moreover, performing as both the trustworthy verifier and key generation center, TA is infeasible to be compromised by the adversaries. Thus various services can be securely presented to the designated vehicles. In this case, the group key is necessary for secure message exchange.

Road-side units (RSUs) are vital VANET infrastructure implemented at the roadside, which perform as the sole intermediaries between TA and vehicles. RSUs are responsible for verifying the vehicles. Furthermore, the group key issued by TA will also be delivered by RSUs. Note that RSUs are assumed to have adequate storage in order to manage the acquired data. Hence, in our scheme, the gathered road messages are stored and managed in RSUs. In general, RSUs are connected with TA in a secure wired way. However, since

the RSUs are placed along the roadside far from TA, it is possible that these RSUs may be physically compromised. In this case, the stored user information may be illegally acquired [20]. For this consideration, the RSUs are assumed to be semitrusted devices.

Vehicles are referred to as terminal users of the VANET system. It is designed to be both service receiver and information collector. In other words, with the implemented OBU, each vehicle is able to receive the broadcasting messages. Meanwhile, the vehicle reports real-time road information to RSU wirelessly. In this case, it is essential to adopt effective cryptographic strategies in order to guarantee the secure transmission. Furthermore, each vehicle is equipped with a tamper-proof device (TPD), where the corresponding secrets and derived group key are stored. In our system model, each driver is relevant to certain vehicles during the registration to TA. Every time when the driver activates the vehicles, his/her fingerprint and the assigned certification card are verified. This way, the driver and the correlative vehicle are closely connected. Consequently, for better description, the driver, the OBU, and the vehicle are considered the same entity in this paper.

3.5. Network Assumption. As illustrated in Figure 1, TA manages all the operative RSUs of the VANET system through wired communication. Various safe strategies deployed for TA-RSU communication guarantee the security of the key data exchange. Therefore, the vehicle secret keys can be securely delivered to the correlative RSUs. However, it is possible that some RSUs are compromised physically since they are far from the TA. In this way, the distributed vehicle secret keys are illegally acquired by the adversaries, which could damage both the VANET system and the user privacy. Considered as semitrusted devices, it is not appropriate for the RSUs to manage all the vehicle-related secret keys. As a result, we assume in this paper the TA-RSU communication channel is safe for data transmission, while the RSU itself may be damaged, which results in vehicle information disclosure.

Two types of wireless communication are displayed in the proposed system model, including the vehicle-to-vehicle communication (V2V) and the vehicle-to-RSU communication (V2R). Due to the inherent wireless transmission characteristics, both V2V and V2R communication suffer from charted and uncharted attacks. In a nutshell, the V2V communication is used for information sharing and cooperative data processing between the neighboring vehicles. While the V2R communication emphasizes longitudinal message acquisition and feedback between vehicle and RSU. Note that in our scheme the operative vehicles safely exchange messages with RSUs on road condition using the derived group key.

4. Proposed Secure Certificateless Authentication Scheme

In a nutshell, two principal factors are taken into consideration in this paper: the secure authentication and road message management mechanism, which will be, respectively, discussed in two sections. In this section, we describe the proposed secure certificateless authentication scheme between RSU and vehicles. The proposed scheme can be clarified into three different phases, including *initialization phase*, *authentication phase*, and *group key distribution phase*. Accordingly, some nontrivial preparations are made in the initialization phase. Subsequently, verification on the vehicles is conducted in the following authentication phase. Finally, the generated group key is allocated to the legitimate vehicle.

Our design adopts the certificateless encryption strategy based on elliptic curve cryptography (ECC). Note that the corresponding public keys have been previously revealed to the devices. Meanwhile, the confidential information is assigned to vehicle during registration. Based on this, the adopted cryptographic techniques are available, which could provide adequate security assurance for the VANET system. Emphasizing the authentication between RSU and vehicle, we describe our scheme in the scenario involving single RSU and single vehicle. Note that the scheme for regular VANET scenarios with multiple RSUs and vehicles is similar.

4.1. Initialization Phase. Necessary preliminary works are conducted in the initialization phase, which can be generally classified into user registration and key information allocation. It is desirable that each vehicle should register to TA first. After that, TA assigns the secret information to the corresponding vehicle. Moreover, TA stores the drivers' personal information such as the car plate number, the contact information, and the address. Let $Q_{\mathscr{H}}$ be the generator of a cyclic additive group $\mathbb{G}_{\mathscr{H}}$ and id be the unique identifier for vehicle. Additionally, TA adopts secure hash functions $h : \{0,1\}^* \times \mathbb{G}_{\mathscr{H}} \to \mathbb{Z}^*_{\mathscr{P}}$, where $\mathbb{Z}^*_{\mathscr{P}}$ is defined as a nonnegative integer set less than the large prime number $\mathscr{P}$. Hence, TA generates the secret key R_{id} for each vehicle illustrated as

$$R_{id} = h\left(id, Q_{\mathscr{H}}\right), \tag{1}$$

which is allocated to the relevant vehicle after user registration. Note that the secret keys of all the registered vehicles are securely stored in TA's database. At the same time, TA chooses a random integer $s_{\mathrm{RSU}} \in \mathbb{Z}^*_{\mathscr{P}}$ as the RSU private key. Let $\mathbb{G}_1$ be the cyclic additive group generated by P with the order q. Hence the RSU public key can be computed according to

$$Q_{\mathrm{RSU}} = s_{\mathrm{RSU}} P. \tag{2}$$

It is worth noting that the RSU public key Q_{RSU}, the generator P, the hash function h, and $\mathbb{G}_1$ will be published to all the devices, while the private key s_{RSU} is kept secret during the entire process.

Now we assume that the registered vehicle approaches the working range of a fixed RSU. If certain vehicle wants to receive services from the VANET system, identification and key assignment are essential. In this assumption, the vehicle chooses $s_v \in \mathbb{Z}^*_{\mathscr{P}}$ as its partial private key. Then the corresponding partial public key Q_v is defined as

$$Q_v = s_v P, \tag{3}$$

where P is the system parameter as mentioned above. Subsequently, $\langle Request, Q_v, id \rangle$ are delivered to RSU.

After deriving the partial public key Q_v, RSU requests TA for the secret key R_{id} of vehicle id. Let $H : \{0,1\}^* \times \mathbb{G}_1 \to \mathbb{G}_1$

denote the secure hash function. Related computations can be conducted for partial key generation as follows:

$$\begin{aligned} Q_{id} &= H\left(id, Q_{\mathrm{RSU}}\right) \\ C &= R_{id} s_{\mathrm{RSU}} Q_{id}. \end{aligned} \tag{4}$$

Thereafter, the generated C is delivered to vehicle. Hence the vehicle derives the partial private key s_u according to

$$s_u = R_{id}^{-1} C = s_{\mathrm{RSU}} Q_{id}. \tag{5}$$

At this point, the public key set for vehicle can be displayed as $\langle Q_v, id\rangle$. Meanwhile, the relevant private key set is defined as $\langle s_v, s_u\rangle$. Note that the two partial private keys are, respectively, decided by RSU and vehicle. In other words, RSU has no access to s_v, so that the privacy protection based on certificateless cryptography is achieved even if RSU is compromised by attackers.

4.2. Authentication Phase. After initialization, RSU conducts authentication on the requesting vehicle. In certain time point t, we assume that vehicle starts to use the road message service. Then the following computation is conducted:

$$Q_1 = tQ_v = ts_v P, \tag{6}$$

which combines the current time with the partial public key. In addition, let $\mathbb{G}_2$ be the cyclic group of prime order s and $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$ be the bilinear pairing operator. Hence the vehicle gets the intermediate value Q_{id} by

$$Q_{id} = H\left(id, Q_{\mathrm{RSU}}\right), \tag{7}$$

where the vehicle identity id and RSU public key Q_{RSU} are known to vehicle. Subsequently, two necessary parameters r and v are generated as

$$\begin{aligned} r &= \hat{e}\left(Q_{\mathrm{RSU}}, Q_{id}\right) \\ v &= h\left(t \parallel r, R_{id}\right), \end{aligned} \tag{8}$$

where R_{id} is the secret key previously allocated to vehicle in the initialization. Accordingly, the signature U is generated based on

$$U = tvs_v Q_{id} + s_u. \tag{9}$$

At this point, vehicle sends $\langle id, t, v, U\rangle$ to the RSU. In the RSU side, the validity of the received t and v will be verified first. Then RSU computes whether

$$r \stackrel{?}{=} \frac{\hat{e}\left(P, U\right)}{\hat{e}\left(Q_1, Q_{id}\right)^v} \tag{10}$$

holds. The correctness is elaborated as follows:

$$\begin{aligned} \frac{\hat{e}\left(P, U\right)}{\hat{e}\left(Q_1, Q_{id}\right)^v} &= \frac{\hat{e}\left(P, tvs_v Q_{id} + s_u\right)}{\hat{e}\left(tQ_v, Q_{id}\right)^v} \\ &= \frac{\hat{e}\left(P, h\left(t \parallel r, R_{id}\right) ts_v Q_{id} + s_{\mathrm{RSU}} Q_{id}\right)}{\hat{e}\left(Q_v, tQ_{id}\right)^{h(t\parallel r, R_{id})}} \\ &= \frac{\hat{e}\left(P, h\left(t \parallel r, R_{id}\right) ts_v Q_{id} + s_{\mathrm{RSU}} Q_{id}\right)}{\hat{e}\left(ts_v P, tQ_{id}\right)^{h(t\parallel r, R_{id})}} \\ &= \frac{\hat{e}\left(P, h\left(t \parallel r, R_{id}\right) ts_v Q_{id} + s_{\mathrm{RSU}} Q_{id}\right)}{\hat{e}\left(P, ts_v Q_{id}\right)^{h(t\parallel r, R_{id})}} \\ &= \frac{\hat{e}\left(P, h\left(t \parallel r, R_{id}\right) ts_v Q_{id}\right) \hat{e}\left(P, s_{\mathrm{RSU}} Q_{id}\right)}{\hat{e}\left(P, ts_v Q_{id}\right)^{h(t\parallel r, R_{id})}} \\ &= \frac{\hat{e}\left(P, ts_v Q_{id}\right)^{h(t\parallel r, R_{id})} \hat{e}\left(P, s_{\mathrm{RSU}} Q_{id}\right)}{\hat{e}\left(P, ts_v Q_{id}\right)^{h(t\parallel r, R_{id})}} \\ &= \hat{e}\left(P, s_{\mathrm{RSU}} Q_{id}\right) = \hat{e}\left(s_{\mathrm{RSU}} P, Q_{id}\right) = \hat{e}\left(Q_{\mathrm{RSU}}, Q_{id}\right). \end{aligned} \tag{11}$$

If the delivered signature U passes the above verification, the validity of the requesting vehicle can be guaranteed. Thereafter, the authentication phase is completed.

4.3. Group Key Distribution Phase. After the authentication phase, the generated group key is distributed to the legitimate vehicle. It is worth emphasizing that the group key is assumed to be chosen by TA. Meanwhile, the key is delivered to RSUs in a secure way. In this way, when a certain vehicle travels from the effective range of one RSU to the next, the group key is always effective and can be continuously used.

We assume that the secret $\mathcal{K} \in \mathbb{Z}_{\mathcal{P}}^*$ is randomly chosen by TA and then delivered to RSU. In certain time point T, RSU computes

$$\begin{aligned} W &= \mathcal{K} T Q_v \\ F &= h\left(W \parallel v, s_u\right) \end{aligned} \tag{12}$$

and sends $\langle W, F, T\rangle$ to vehicle. Note that the RSU could generate the partial private key s_u using the known information. In this way, the secret $\mathcal{K}$ is combined with current time stamp T and previously acquired intermediate value v.

Similarly, the vehicle first compares the received value F with the stored one. If F is valid, vehicle derives the secret by computing

$$\begin{aligned} N &= T^{-1} s_v^{-1} W = T^{-1} s_v^{-1} \mathcal{K} T Q_v = T^{-1} T s_v^{-1} s_v \mathcal{K} P \\ &= \mathcal{K} P. \end{aligned} \tag{13}$$

At this point, the final group key gk can be acquired according to

$$gk = h\left(N\right) = h\left(\mathcal{K} P\right). \tag{14}$$

Therefore, the group key gk is successfully allocated to the legitimate vehicle. Note that gk will be used in the subsequent communication such as road message dissemination, reporting, and evaluation. The delivered packet format is as follows:

$$\left\langle type, E_{gk}\left(E_{R_{id}}\left(m\right) \parallel id \parallel TS\right), TS\right\rangle, \tag{15}$$

where the transmitted message m is symmetrically encrypted using both R_{id} and the group key gk. Similarly, the current

time stamp denoted as TS is adopted in the encryption. Note that $E_x(y)$ represents the symmetric encryption on y using secret key x. Additionally, type indicates the type of m. For security consideration, the communication between vehicle and RSU adopts the assigned group key gk for encryption.

5. Proposed Road Message Dissemination Scheme

In this section, we describe the corresponding road message dissemination scheme in detail. Meanwhile, the message evaluation and award mechanism are presented.

5.1. Road Message Reporting. We assume the scenario that a specific vehicle with identity id is within the effective range of a RSU. Note that the vehicle has successfully passed the authentication and acquired the group key gk. Subsequently, in a certain time point TS, the vehicle passes through a particular spot where road event occurs. For example, when the vehicle passes through the road accident scene, the driver could consider this accident as a road message and report it to the RSU. According to (15) in previous section, the RSU gets the packet involving the encrypted message and the detailed time point TS. The decryption process is as follows.

First, the decryption with group key gk is conducted in the RSU side according to

$$D_{gk}\left(E_{gk}\left(E_{R_{id}}(m) \parallel id \parallel TS\right)\right) = E_{R_{id}}(m) \parallel id \parallel TS. \quad (16)$$

Next, RSU checks the time stamp TS with the derived one in order to ensure that the received message is timely and effective. Additionally, id of the vehicle is derived. According to the aforementioned design, the RSU acquires the relevant secret R_{id} in its storage so that the transmitted road message m can be acquired according to

$$m = D_{R_{id}}\left(E_{R_{id}}(m)\right). \quad (17)$$

At this point, the RSU is aware of the identity information of the reporting vehicle. Hence, according to id, RSU requests TA for the vehicle priority parameter, which is considered as the initial element in the message management process. In practical scenario with multiple vehicles existing in one RSU's effective range, it is possible that more than one vehicles report the same road message to TA. For example, two vehicles V_1 and V_2 may successively pass through a certain accident scene. Hence, both of them report this event to RSU. In this case, the RSU stores this road message m in its storage and assigns the broadcast priority $\mathrm{br}(m)$ for m, which can be calculated using the priority of the two reporting vehicles as follows:

$$\mathrm{br}(m) = \frac{1}{2}\left[\mathrm{pr}(id_1) + \mathrm{pr}(id_2)\right], \quad (18)$$

where the identifiers of V_1 and V_2 are denoted as id_1 and id_2, respectively. Meanwhile, the priorities of the two vehicles are $\mathrm{pr}(id_1)$ and $\mathrm{pr}(id_2)$. The calculated $\mathrm{br}(m)$ here represents the broadcast priority right after the two vehicles report the message. Moreover, among all the vehicles, it is assumed that only V_1 and V_2 report m to RSU such that $\mathrm{br}(m)$ is achieved as the average value of all the reporting vehicles' priority.

In a nutshell, we assume that the vehicle set $V = \{V_i \mid i \in [1,n],\ n \in \mathbb{N}^*\}$ denotes the n legitimate vehicles that have already passed the authentication process conducted by the effective RSU. Among these n vehicles, the vehicle subset $RP_m = \{V_j \mid j \in [1,g],\ g \in \mathbb{N}^* \wedge g \le n\}$ consists of all the g vehicles that report road message m. Note that the identity of V_j is denoted as id_j. Hence, the broadcast priority is $\mathrm{pr}(id_j)$ and $RP_m \subseteq V$ holds. It is assumed that the g vehicles report message following the sequence of $V_1, \ldots, V_g$. Moreover, the road event $Event_m^{loc}$ is denoted as follows:

$$Event_m^{loc} = \{loc, type, cont\}, \quad (19)$$

which indicates that $Event_m^{loc}$ happened in location loc and the detailed information is showed in $cont$. Moreover, the event type is defined as $type$. In this way, the road message m contains the essential elements of $Event_m^{loc}$.

In our assumption, after $Event_m^{loc}$ occurred, within certain time interval g vehicles will report the event to the RSU. Hence, after RSU receives the road report from V_1 for the first time, the broadcast priority is computed as $\mathrm{br}(m) = \mathrm{pr}(id_1)$. Similarly, the broadcast priority after RSU receives the road report for k ($1 \le k \le g$) times, and $\mathrm{br}(m)$ can be computed as

$$\mathrm{br}(m) = \frac{1}{k}\sum_{i=1}^{k}\mathrm{pr}(id_i), \quad (20)$$

which is defined as the average vehicle priorities of all the reporting vehicles. Hence after Δt, the broadcast priority for m is

$$\mathrm{br}(m) = \frac{1}{g}\sum_{i=1}^{g}\mathrm{pr}(id_i). \quad (21)$$

Practically, one RSU handles multiple different road messages simultaneously. Therefore, each message will be assigned a broadcast priority and then stored in the storage. In our design, the creditability and accuracy of the road message are highly related to the reporter's previous records. And the vehicle priority is able to reflect this property properly. Consequently, RSU sorts all these messages so that the reliable messages will be broadcast first.

5.2. Road Message Dissemination. As illustrated in the above section, the RSU manages all the road information within its effective range. Periodically, RSU broadcasts the messages in certain sequence. Note that all the road messages are encrypted using the distributed group key gk. Hence, only the registered vehicles can get access to this service. The aforementioned broadcast priority $\mathrm{br}(m)$ is roughly decided by reliability of the reporting vehicles. In this way, a predefined parameter $\mathscr{W}$ is set as the minimum requirement for message dissemination. That is to say, the messages m will be broadcast only if $\mathrm{br}(m) \ge \mathscr{W}$. Otherwise it will be considered as unreliable information and then temporarily disabled from the broadcast list. In this case, in future, if other vehicles

report the same message, $\mathrm{br}(m)$ will be compared with $\mathscr{W}$ again. After a predefined time interval, if $\mathrm{br}(m) < \mathscr{W}$, RSU permanently deletes the message in its storage. Following the above procedure, vehicles could acquire road messages in an accurate way. Thus the driving security can be improved with this service.

5.3. Evaluation and Priority Management. For practical consideration, an appropriate evaluation mechanism towards the road messages is necessary. In this subsection, we describe the proposed evaluation and priority management scheme. The value of the stored road message is decided by not only the reputation of the reporter but also the message itself. In order to achieve this, we assume that the vehicles have the capability of evaluating the received road messages. Following the above assumptions, the vehicle subset $RV_m = \{V_l \mid l \in [1, f],\ f \in \mathbb{N}^*\}$ denotes f vehicles who receive the road message m within time interval Δt_{RC}. After $V_l \in RV_m$ approaches the location loc where the road event $Event_m^{loc}$ happened, V_l could evaluate whether the received road information is correct, which helps improve the road report accuracy. The format of the evaluation message is as follows:

$$\left\langle type, E_{gk}\left(E_{s_u}\left(ifNO \parallel assess\right) \parallel id \parallel TS_{ev}\right), TS_{ev}\right\rangle, \quad (22)$$

where the message type denoted as $type$ here indicates that it is an evaluation message. $ifNO$ denotes the assigned information number for message m. In addition, TS_{ev} is the current time stamp. Note that this evaluation will be sent back to RSU. According to

$$\begin{aligned} &D_{gk}\left(E_{gk}\left(E_{s_u}\left(ifNO \parallel assess\right) \parallel id \parallel TS_{ev}\right)\right) \\ &\quad = E_{s_u}\left(ifNO \parallel assess\right) \parallel id \parallel TS_{ev}, \end{aligned} \quad (23)$$

RSU checks the time stamp TS_{ev} with the derived one in order to ensure that the message is timely and effective. Additionally, id of the vehicle is derived. RSU acquires the relevant secret s_u in its storage so that the transmitted evaluation on message m can be acquired according to

$$D_{s_u}\left(E_{s_u}\left(ifNO \parallel assess\right)\right) = ifNO \parallel assess. \quad (24)$$

In this way, the evaluation can be combined to m according to $ifNO$. As a result, for road message m, the RSU could receive h_m evaluation messages where $h_m \le f$. For practical consideration, the $assess$ can be analyzed using different state parameters such as $sp \in \{-2, -1, 0, 1, 2\}$, where $sp = 2$ means that $assess$ is totally accurate and helpful, while $sp = -2$ means that $assess$ is of no help and thus is considered as the fake message. The h_m state parameters are $\{sp_u \mid u \in [1, h_m],\ u \in \mathbb{N}^*\}$. During every certain period Δt_{RD}, RSU analyzes all the received evaluation messages and updates the broadcast priority following the steps below:

(i) Screening: firstly, RSU checks whether

$$Rate_m^{\Delta t_{RD}}\left[sp_u > 0\right] \ge \mathscr{V} \quad (25)$$

holds, where $Rate_m^{\Delta t_{RD}}[sp_u > 0]$ denotes the proportion of the received $assess$ whose $sp_u > 0$ among all the h_m evaluation on m. Furthermore, $\mathscr{V}$ is the predefined system parameter according to different practical scenarios. In this case, if most of the users give negative assessments, m is considered as invalid information and must be discarded from the storage immediately.

(ii) Updating: secondly, the updating on broadcast priority is conducted as

$$\begin{aligned} \mathrm{br}'(m) &= \mathrm{br}(m) + \frac{\sum_{u=1}^{h_m}\left[\mathrm{pr}\left(id_u\right) \times sp_u\right]}{h_m} \\ &= \frac{\sum_{i=1}^{g}\mathrm{pr}\left(id_i\right)}{g} + \frac{\sum_{u=1}^{h_m}\left[\mathrm{pr}\left(id_u\right) \times sp_u\right]}{h_m}, \end{aligned} \quad (26)$$

where $[\mathrm{pr}(id_u)]_{u\in[1,h_m]}$ denotes the priority of h_m vehicles.

At this point, the updating process for Δt_{RD} is completed. Similarly, after n time periods $n\Delta t_{RD}$, the broadcast priority for m is

$$\begin{aligned} \mathrm{br}^{n\Delta t_{RD}}(m) = &\frac{\sum_{i=1}^{g}\mathrm{pr}\left(id_i\right)}{g} \\ &+ \sum_{i=1}^{n}\left(\frac{\sum_{u=1}^{h_m^i}\left[\mathrm{pr}\left(id_u^i\right) \times sp_u^i\right]}{h_m^i}\right), \end{aligned} \quad (27)$$

where h_m^i, $\mathrm{pr}(id_u^i)$, and sp_u^i are the parameters in ith time periods $i\Delta t_{RD}$. Note that the above process should be conducted for each stored road message in RSU. Hence, the broadcast sequence is updated. In future, after $x\Delta t_{RD}$, m will be deleted if $br^{x\Delta t_{RD}}(m) < \mathscr{N}$, where $\mathscr{N}$ is the preset system parameter.

5.4. Vehicle Priority Management. As illustrated above, the vehicle priority on vehicle id is $\mathrm{pr}(id)$, which is a significant user property in both the broadcast priority computing and updating processes. As a matter of fact, the reporting vehicle plays a crucial role in the message dissemination scheme. Hence, appropriate rewarding strategy is essential to motivate drivers' enthusiasm on road situation reporting. The incentives will be given according to the vehicle priority. To achieve this, $\mathrm{pr}(id)$ will be updated according to the value of his reporting road message.

We assume that road message m is reported by several vehicles in $B = \{vp_i \mid 1 \le i \le n,\ i \in \mathbb{N}^*\}$. After a sufficient time period, for example, twenty-four hours, one road message m has been evaluated by multiple vehicles. According to (25) in the previous section, m is valid if

$$Rate_m^{\Delta t_{RD}}\left[sp_u > 0\right] \ge \mathscr{V} \quad (28)$$

holds. In this way, $\mathrm{pr}(vp_i)$ of all $vp_i \in B$ are updated as

$$\mathrm{pr}\left(vp_i\right) = \mathrm{pr}\left(vp_i\right) + 1. \quad (29)$$

In contrast, if m is evaluated to be invalid and discarded by RSU, $\mathrm{pr}(vp_i)$ of all $vp_i \in B$ are updated as

$$\mathrm{pr}\left(vp_i\right) = \mathrm{pr}\left(vp_i\right) - 1. \tag{30}$$

Note that the driver could change it into incentives such as coupons of cooperative stores or scorecard in the road service area.

6. Security Analysis

In this section, we analyze the security properties of the proposed authentication scheme. The security theorems as well as the corresponding proofs are given below.

6.1. Unforgeability against Chosen Message Attack. We analyze the unforgeability against chosen message attack in the proposed protocol.

Theorem 1. *The proposed certificateless authentication scheme is existentially unforgeable against adaptive chosen message attack under the assumption of random oracle model if and only if the CDHP is hard.*

Proof. The security of unforgeability is formally defined through game $\mathscr{G}_1$. Let $\mathscr{A}_1$ be a probabilistic polynomial time adversary. $\mathscr{C}_1$ denotes the challenger; h and H denote the random oracles. In order to solve CDHP problem, it is assumed that $\mathscr{C}_1$ is able to simulate all the related oracles. In $\mathscr{G}_1$, $\mathscr{A}_1$ can conduct the following corresponding queries to $\mathscr{C}_1$.

h Query. We assume that the adversary $\mathscr{A}_1$ itself does not have the ability to directly compute the hash function $h(\cdot)$. Hence, the response to *h Query* is simulated by maintaining a list $List_h$ initialized to be empty. That is to say, when the oracle is queried with the input values $\langle t, r, R_{id}\rangle$, if the query $\langle t, r, R_{id}\rangle$ already exists in $List_h$, $\mathscr{C}_1$ outputs $v = h(t \parallel r, R_{id})$ to $\mathscr{A}_1$. Otherwise, $\mathscr{C}_1$ chooses a random number $v \in \mathbb{Z}^*$ and forwards it to $\mathscr{A}_1$. After that, $\langle t, r, R_{id}, v\rangle$ will be subsequently added to $List_h$.

Extract Query. The adversary $\mathscr{A}_1$ is able to query the partial private key of any given key set $\langle Q_v, id\rangle$. According to id, $\mathscr{C}_1$ outputs the partial private key s_u to $\mathscr{A}_1$.

H Query. $\mathscr{A}_1$ can query the random oracle H at any time. The response to *H Query* is simulated by maintaining a list $List_H$. Note that $List_H$ is initialized to be empty. When the oracle H is queried with input values $\langle id, Q_{\mathrm{RSU}}\rangle$, if the query $\langle id, Q_{\mathrm{RSU}}\rangle$ already exists in $List_H$, $\mathscr{C}_1$ outputs $Q_{id} = H(id, Q_{\mathrm{RSU}})$ to $\mathscr{A}_1$. Otherwise, $\mathscr{C}_1$ chooses a random number $Q_{id} \in \mathbb{Z}^*$ and forwards it to $\mathscr{A}_1$. After that, $\langle id, Q_{\mathrm{RSU}}, Q_{id}\rangle$ will be subsequently added to $List_H$.

Authenticating Query. $\mathscr{C}_1$ simulates the authenticating oracle by responding to the authenticating query as follows:

(i) $\mathscr{C}_1$ randomly chooses $U \in \mathbb{G}_1$ as the certificate and $v \in \mathbb{Z}^*$ as the intermediate parameter.

(ii) $\mathscr{C}_1$ computes $r = \widehat{e}(P, U)/\widehat{e}(Q_1, Q_{id})^v$. If $\langle t, r, R_{id}, v\rangle$ already exists in $List_h$, $\mathscr{C}_1$ chooses other values and tries again.

(iii) $\mathscr{C}_1$ adds the above $\langle t, r, R_{id}, v\rangle$ to $List_h$.

(iv) $\mathscr{C}_1$ outputs $\langle id, t, v, U\rangle$ as the certificate.

According to the *Forking Lemma* [43], $\mathscr{A}_1$ produces two valid certificates $\langle id, t, v, U\rangle$ and $\langle id, t, v', U'\rangle$ $(v \neq v')$. In this case,

$$\begin{aligned} U &= tvs_v Q_{id} + s_u \\ U' &= tv' s_v Q_{id} + s_u, \end{aligned} \tag{31}$$

hold. Hence we can get

$$s_v Q_{id} = \frac{\left(v - v'\right)^{-1}\left(U - U'\right)}{t}. \tag{32}$$

In this way, we show that the CDHP problem can be solved. In other words, the attacker needs to solve the CDHP problem in order to forge the certificate. However, this contradicts the hardness of the CDHP problem [44, 45]. In conclusion, an attacker cannot forge the certificate in the authentication process [46, 47]. □

6.2. Resistance to Replay Attack. The replay attack is achieved by reusing the previous generated message to pass the current authentication process. The security property against replay attack is discussed in this section.

Theorem 2. *During the certificateless authentication process, replay attack can be prevented. That is, the previous messages of the past authentication sessions cannot pass the current authentication process in RSU side.*

Proof. We discuss the resistance to replay attack through game $\mathscr{G}_2$. Similarly, let $\mathscr{A}_2$ be a probabilistic polynomial time adversary. It is assumed that, in time point $\mathscr{T}_1$, $\mathscr{A}_2$ has access to all the published system parameters as well as the transmitted messages from $\mathscr{T}_0$ to $\mathscr{T}_1$ $(\mathscr{T}_0 < \mathscr{T}_1)$. Randomly, $\mathscr{A}_2$ chooses the message $\langle id, \mathscr{T}_r, v^{\mathscr{T}_r}, U^{\mathscr{T}_r}\rangle$ at time $\mathscr{T}_r \in [\mathscr{T}_0, \mathscr{T}_1]$. At $\mathscr{T}_1$, $\mathscr{A}_2$ sends $\langle id, \mathscr{T}_1, v^{\mathscr{T}_r}, U^{\mathscr{T}_r}\rangle$ as the replaying message. Note that $U^{\mathscr{T}_r} = \mathscr{T}_r v^{\mathscr{T}_r} s_v Q_{id} + s_u$. In this way, $r \neq \widehat{e}(Q_{\mathrm{RSU}}, Q_{id})$. Hence, the previous message cannot pass the current authentication. □

6.3. Forward Security. In this section, we analyze the forward security property of the proposed protocol.

Theorem 3. *The proposed authentication scheme provides forward security against adversary. That is, the adversary cannot pass the authentication with the acquired vehicle secret key from the compromised RSU.*

Proof. We assume that the RSU has already been compromised by brute-force attack and all the stored key information is leaked to the adversary $\mathscr{A}_3$. The secret key set $\{s_u^i \mid i \in [1, n]\}$ denotes the private keys of all the n vehicles. In this case, $\mathscr{A}_3$ is able to use the derived partial private key s_u for

Table 1: Comparison of storage overhead.

Protocol	ICPA [10]	DAKM [4]	ABAP [8]	SAAP [32]	Our protocol
Storage (Vehicle)	2112 bits	2712 bits	3432 bits	2904 bits	2520 bits
Storage (RSU)	$1440n + 1056$ bits	$1676n + 1144$ bits	$1936n + 1048$ bits	$2008n + 2392$ bits	$1592n + 768$ bits

certificate generation. However, the private key contains both s_u and s_v, while s_v is chosen by the vehicle itself. Due to the hardness of the aforementioned DLP problem, the probability that $U = tvs_vQ_{id} + s_u$ can be correctly computed is illustrated as $1/2^{\varpi}$, where ϖ is the size of s_v. In general, the certificateless authentication property guarantees the forward security of the proposed scheme. □

6.4. Session Key Establishment. In the system model of this paper, it is necessary to generate a shared session key between the RSU and vehicle so as to guarantee the data confidentiality and transmission security, which is analyzed as follows.

Theorem 4. *In the proposed protocol, the shared session key is established after successfully authentication between RSU and vehicle.*

Proof. According to the protocol design, $\mathcal{K}$, along with the current time stamp and the previously acquired intermediate value v, is transmitted to the vehicle. In the vehicle side, the below derivation of $\mathcal{K}$ is conducted as $N = T^{-1}s_v^{-1}W = \mathcal{K}P$, where s_v is stored in vehicle already. Note that the security assurance of the message transmission is based on the hardness of DLP problem. The final group key is generated and adopted to the following message transmission. □

6.5. Mutual Authentication. In this section, we analyze the mutual authentication property in the proposed authentication protocol.

Theorem 5. *The proposed protocol can provide mutual authentication between RSU and vehicle if the DLP problem is intractable.*

Proof. During the authentication process, the RSU-to-vehicle security is preserved by the aforementioned certificate $\langle id, t, v, U\rangle$, which has been discussed in the proof of Theorem 1. On the other hand, the vehicle-to-RSU security is based on the hardness of DLP problem. Specifically, $F = h(W \parallel v, s_u)$ is contained in the delivered $\langle W, F, T\rangle$ and will be verified by the RSU with the known key information. Therefore, we could conclude that the proposed authentication scheme provides mutual authentication property. □

7. Performance Analysis

In this section, we present the performance analysis towards the proposed protocol. Our analysis on the performance mainly emphasizes the storage overhead, the computation cost, and the communication cost, which are the dominant factors in the proposed protocol.

7.1. Storage Overhead. In the proposed protocol, storage overhead is a crucial parameter for VANETs, especially for vehicles. Due to the inherent resource restriction, it is impractical for the vehicle to store massive key messages and communication data in its own memory. Moreover, the RSU is designed to handle both the key distribution and road message management simultaneously. As a result, the storage overhead in both the vehicle and RSU sides should be considered.

As for the vehicle in the proposed protocol, some essential key information is previously stored during the registration including $\langle id, R_{id}, Q_v, s_v, s_u\rangle$. The published public key of RSU, namely, Q_{RSU}, as well as the intermediate value $\langle Q_{id}, C\rangle$, is also stored in vehicle. Moreover, the transmitted message $\langle r, v, U\rangle$, and the necessary group key distribution value are stored, respectively, in the group key distribution phase. According to [48], we assume that the length of elements in $\mathbb{G}_{\mathcal{H}}$ and $\mathbb{G}_1$ is 256 bits. The lengths of relevant vehicle secret keys such as s_v, s_u, and gk are 160 bits. Moreover, it is assumed than the lengths of the adopted time stamps and the identity of vehicle id are 32 bits and 24 bits each. In this way, the storage overhead for each vehicle is $24 + 256 \times 5 + 160 \times 7 + 32 \times 3 = 2520$ bits. Similarly, we assume that the number of vehicles in the RSU range is n. Consequently, the storage overhead in RSU side includes key information of RSU itself and secret messages of all the n vehicles. In this way, the storage overhead for the RSU is $256+160\times3+32+n(24+256\times4+160\times3+32\times2) = 1592n+768$ bits. The comparison with the state-of-the-art VANETs authentication protocols ICPA [10], DAKM [4], ABAP [8], and SAAP [32] is illustrated in Table 1.

7.2. Computation Cost. In this section, we analyze the computation cost of the proposed protocol. The computation cost is defined as the time consumption for the group key distribution process. The comparison result with ICPA, DAKM, ABAP, and SAAP is given in Table 2. We denote modulo operation as mod, exponential operation as Ex, and bilinear pairing as $\hat{e}$. Enc and Dec refer to encryption and decryption. Additionally, H, M, D, and A represent one-way hash function, multiplication operation, division operation, and addition operation, respectively. Finally, the point multiplication operation is dented as p.

7.3. Communication Cost. The communication cost refers to the time consumption for message transmission. In this subsection, we consider the required communication passes for RSU to successfully authenticate vehicles. The comparison result on communication cost is given in Table 3.

8. Conclusion

Emphasizing the secure authentication and road message dissemination in VANETs, a secure certificateless authentication and road message dissemination protocol is proposed in this paper. In our design, certificateless cryptographic technique

Table 2: Comparison of computation cost.

Protocol	ICPA [10]	DAKM [4]	ABAP [8]	SAAP [32]	Our protocol
Computation cost (vehicle)	$3p + 3H + 2A + 2M + 2\text{mod}$	$1Dec + \text{mod} + 1Enc$	$1Dec + 4p + 2H + 2A + D$	$10p + 5H + 4A + M$	$4p + 3H + \widehat{e} + A + 3M$
Computation cost (RSU)	$3nM + (2n + 2)p + 2A$	$3Enc + 2nM + nD + (n-1)A$	$2nH + 2np + 2nM + 3n\widehat{e} + nA$	$4nH + (8n + 1)p + nM + 3nA$	$2nH + 2np + 2nM + 2n\widehat{e} + nEx$

Table 3: Comparison of communication cost.

Protocol	ICPA [10]	DAKM [4]	ABAP [8]	SAAP [32]	Our protocol
Computation rounds	$4n + 2$	$4n + 1$	$4n + 1$	$4n + 4$	$4n$

is employed for authentication and key distribution. Subsequently, an appropriate road message dissemination mechanism is designed. The security analysis and performance evaluation are given accordingly. The proposed protocol is suitable for practical VANET scenarios and is capable of providing timely road information services, which improves both the user safety and the driving experience.

Notations

TA, RSU:	Trustworthy authority, road-side units
$\mathbb{G}_{\mathscr{K}}$:	Cyclic additive group
$Q_{\mathscr{K}}$:	Generator of $\mathbb{G}_{\mathscr{K}}$
id:	Unique identifier of vehicle
h:	Secure hash function, $h : \{0,1\}^* \times \mathbb{G}_{\mathscr{K}} \to \mathbb{Z}^*_{\mathscr{P}}$
R_{id}:	Secret key for vehicle
s_{RSU}:	RSU private key
P:	Generator of cyclic additive group $\mathbb{G}_1$
Q_{RSU}:	RSU public key
s_v:	Vehicle partial private key
Q_v:	Vehicle partial public key
H:	Secure hash function, $H : \{0,1\}^* \times \mathbb{G}_1 \to \mathbb{G}_1$
s_u:	Vehicle partial private key
Q_{id}:	Intermediate authenticating value
$\mathscr{K}$:	Secret key generated by TA
gk:	Group key
$E_x(y)$, $D_x(y)$:	Symmetric encryption and decryption on y with x
m:	The disseminated message
$\text{br}(m)$:	Broadcast priority of m
$\text{pr}(id)$:	Vehicle priority
$Event_m^{loc}$:	Road event
g:	Number of reporting vehicles
$\mathscr{W}, \mathscr{V}, \mathscr{N}$:	Predefined system parameters.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIP) (nos. NRF-2016R1A2B4012638 and NRF-2017R1D1A3B03034005) and by the MIST (Ministry of Science & ICT), Korea, under the National Program for Excellence in SW, supervised by the IITP (Institute for Information & Communication Technology Promotion) (2017-0-00137).

References

[1] X. Peng, "A novel authentication protocol for vehicle network," in *Proceedings of the 2016 3rd International Conference on Systems and Informatics, ICSAI 2016*, pp. 664–668, Shanghai, China, November 2016.

[2] S. Guo, D. Zeng, and Y. Xiang, "Chameleon hashing for secure and privacy-preserving vehicular communications," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 11, pp. 2794–2803, 2014.

[3] A. Studer, F. Bai, B. Bellur, and A. Perrig, "Flexible, extensible, and efficient VANET authentication," *Journal of Communications and Networks*, vol. 11, no. 6, pp. 574–588, 2009.

[4] P. Vijayakumar, M. Azees, A. Kannan, and L. J. Deborah, "Dual Authentication and Key Management Techniques for Secure Data Transmission in Vehicular Ad Hoc Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 1015–1028, 2015.

[5] J. Sun, C. Zhang, Y. Zhang, and Y. Fang, "An identity-based security system for user privacy in vehicular ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 9, pp. 1227–1239, 2010.

[6] J.-L. Huang, L.-Y. Yeh, and H.-Y. Chien, "ABAKA: An anonymous batch authenticated and key agreement scheme for value-added services in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 248–262, 2011.

[7] F.-K. Tseng, Y.-H. Liu, J.-S. Hwu, and R.-J. Chen, "A secure reed-solomon code incentive scheme for commercial Ad dissemination over VANETs," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4598–4608, 2011.

[8] S. Jiang, X. Zhu, and L. Wang, "An efficient anonymous batch authentication scheme based on HMAC for VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2193–2204, 2016.

[9] K.-A. Shim, "CPAS: An efficient conditional privacy-preserving authentication scheme for vehicular sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 4, pp. 1874–1883, 2012.

[10] 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.

[11] J. Shen, H. Tan, Y. Zhang, X. Sun, and Y. Xiang, "A new lightweight RFID grouping authentication protocol for multiple tags in mobile environment," *Multimedia Tools and Applications*, vol. 76, no. 21, pp. 22761–22783, 2017.

[12] X. Sun, X. Lin, and P.-H. Ho, "Secure vehicular communications based on group signature and ID-based signature scheme," in *Proceedings of the 2007 IEEE International Conference on Communications, ICC'07*, pp. 1539–1545, Glasgow, UK, June 2007.

[13] K. Ansari, C. Wang, L. Wang, and Y. Feng, "Vehicle-to-vehicle real-time relative positioning using 5.9–GHZ DSRC media," in *Proceedings of the IEEE 78th Vehicular Technology Conference (VTC Fall' 13)*, pp. 1–7, September 2013.

[14] X. Cheng, L. Yang, and X. Shen, "D2D for intelligent transportation systems: A feasibility study," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1784–1793, 2015.

[15] C. Zhang, R. Lu, X. Lin, P.-H. Ho, and X. Shen, "An efficient identity-based batch verification scheme for vehicular sensor networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 816–824, Phoenix, AZ, USA, April 2008.

[16] F. Wang, Y. Xu, H. Zhang, Y. Zhang, and L. Zhu, "2FLIP: A two-factor lightweight privacy-preserving authentication scheme for VANET," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 896–911, 2016.

[17] X. Li and L. Wang, "A rapid certification protocol from bilinear pairings for vehicular ad hoc networks," in *Proceedings of the 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom-2012*, pp. 890–895, Liverpool, UK, June 2012.

[18] Y. Hao, Y. Cheng, C. Zhou, and W. Song, "A distributed key management framework with cooperative message authentication in VANETs," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 616–629, 2011.

[19] H. Tan, D. Choi, P. Kim, S. Pan, and I. Chung, "Comments on "Dual Authentication and Key Management Techniques for Secure Data Transmission in Vehicular Ad Hoc Networks"," *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1–3, 2017.

[20] T. W. Chim, S. M. Yiu, L. C. Hui, and V. . Li, "VSPN: VANET-based secure and privacy-preserving navigation," *Institute of Electrical and Electronics Engineers. Transactions on Computers*, vol. 63, no. 2, pp. 510–524, 2014.

[21] B. Liu, D. Jia, J. Wang, K. Lu, and L. Wu, "Cloud-assisted safety message dissemination in VANET-cellular heterogeneous wireless network," *IEEE Systems Journal*, vol. 11, no. 1, pp. 128–139, 2017.

[22] M.-C. Chuang and M. C. Chen, "DEEP: density-aware emergency message extension protocol for VANETs," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4983–4993, 2013.

[23] M. Milojevic and V. Rakocevic, "Location aware data aggregation for efficient message dissemination in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5575–5583, 2015.

[24] J. Shen, T. Zhou, X. Chen, J. Li, and W. Susilo, "Anonymous and Traceable Group Data Sharing in Cloud Computing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 912–925, 2018.

[25] J. Song, C. He, L. Zhang, S. Tang, and H. Zhang, "Toward an RSU-unavailable lightweight certificateless key agreement scheme for VANETs," *China Communications*, vol. 11, no. 9, pp. 93–103, 2014.

[26] L. Zhang, Q. Wu, A. Solanas, and J. Domingo-Ferrer, "A scalable robust authentication protocol for secure vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1606–1617, 2010.

[27] D. Huang, S. Misra, M. Verma, and G. Xue, "PACP: An efficient pseudonymous authentication-based conditional privacy protocol for VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 736–746, 2011.

[28] R. Lu, X. Lin, X. Liang, and X. Shen, "A dynamic privacy-preserving key management scheme for location-based services in VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 127–139, 2012.

[29] D. Johnson, A. Menezes, and S. Vanstone, "The elliptic curve digital signature algorithm (ECDSA)," *International Journal of Information Security*, vol. 1, no. 1, pp. 36–63, 2001.

[30] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, "The TESLA broadcast authentication protocol," *RSA CryptoBytes*, vol. 5, 2005.

[31] J. Shao, X. Lin, R. Lu, and C. Zuo, "A threshold anonymous authentication protocol for VANETs," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1711–1720, 2016.

[32] H. Xiong, "Cost-effective scalable and anonymous certificateless remote authentication protocol," *IEEE Transactions on Information Forensics & Security*, vol. 9, no. 12, pp. 2327–2339, 2014.

[33] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Proceedings of the Advances in Cryptology*, pp. 47–53, 1984.

[34] X. Lin, X. Sun, P.-H. Ho, and X. Shen, "GSIS: a secure and privacy-preserving protocol for vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6 I, pp. 3442–3456, 2007.

[35] C.-C. Lee and Y.-M. Lai, "Toward a secure batch verification with group testing for VANET," *Wireless Networks*, vol. 19, no. 6, pp. 1441–1449, 2013.

[36] J. K. Liu, T. H. Yuen, M. H. Au, and W. Susilo, "Improvements on an authentication scheme for vehicular sensor networks," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2559–2564, 2014.

[37] Y. Zhang, L. Yang, and S. Wang, "An efficient identity-based signature scheme for vehicular communications," in *Proceedings of the 11th International Conference on Computational Intelligence and Security, CIS 2015*, pp. 326–330, Shenzhen, China, December 2015.

[38] S. S. Al-Riyami and K. G. Paterson, "Certificateless public key cryptography," in *Proceedings of the Advances in Cryptology-ASIACRYPT2003*, pp. 452–473, 2003.

[39] H. Xiong, Z. Chen, and F. Li, "Provably secure and efficient certificateless authenticated tripartite key agreement protocol," *Mathematical and Computer Modelling*, vol. 55, no. 3-4, pp. 1213–1221, 2012.

[40] H. Xiong and Z. Qin, "Revocable and scalable certificateless remote authentication protocol with anonymity for wireless body area networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 7, pp. 1442–1455, 2015.

[41] J. Shen, T. Zhou, F. Wei, X. Sun, and Y. Xiang, "Privacy-Preserving and Lightweight Key Agreement Protocol for V2G

in the Social Internet of Things," *IEEE Internet of Things Journal*, pp. 1-1, 2017.

[42] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *Journal of Computer and System Sciences*, vol. 18, no. 2, pp. 143–154, 1979.

[43] D. Pointcheval and J. Stern, "Security arguments for digital signatures and blind signatures," *Journal of Cryptology*, vol. 13, no. 3, pp. 361–396, 2000.

[44] D. Wang, D. He, P. Wang, and C.-H. Chu, "Anonymous Two-Factor Authentication in Distributed Systems: Certain Goals Are Beyond Attainment," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 428–442, 2015.

[45] C. Wang, G. Xu, and J. Sun, "An enhanced three-factor user authentication scheme using elliptic curve cryptosystem for wireless sensor networks," *Sensors*, vol. 17, no. 12, article no. 2946, 2017.

[46] 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*, no. 99, pp. 1–14, 2016.

[47] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, "Zipfs law in passwords," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

[48] D. He, S. Zeadally, N. Kumar, and J. H. Lee, "Anonymous authentication for wireless body area networks with provable security," *IEEE Systems Journal*, no. 99, pp. 1–12, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 5609195, 16 pages
https://doi.org/10.1155/2018/5609195

Research Article

Attack Potential Evaluation in Desktop and Smartphone Fingerprint Sensors: Can They Be Attacked by Anyone?

Ines Goicoechea-Telleria, Raul Sanchez-Reillo, Judith Liu-Jimenez, and Ramon Blanco-Gonzalo

The University Group for Identification Technologies, Carlos III University of Madrid, 28911 Leganes, Spain

Correspondence should be addressed to Ines Goicoechea-Telleria; igoicoec@ing.uc3m.es

Received 30 November 2017; Revised 22 January 2018; Accepted 1 February 2018; Published 16 April 2018

Academic Editor: Ding Wang

The use of biometrics keeps growing. Every day, we use biometric recognition to unlock our phones or to have access to places such as the gym or the office, so we rely on the security manufacturers offer when protecting our privileges and private life. It is well known that it is possible to hack into a fingerprint sensor using fake fingers made of Play-Doh and other easy-to-obtain materials but to what extent? Is this true for all users or only for specialists with a deep knowledge on biometrics? Are smartphone fingerprint sensors as reliable as desktop sensors? To answer these questions, we performed 3 separate evaluations. First, we evaluated 4 desktop fingerprint sensors of different technologies by attacking them with 7 different fake finger materials. All of them were successfully attacked by an experienced attacker. Secondly, we carried out a similar test on 5 smartphones with embedded sensors using the most successful materials, which also hacked the 5 sensors. Lastly, we gathered 15 simulated attackers with no background in biometrics to create fake fingers of several materials, and they had one week to attack the fingerprint sensors of the same 5 smartphones, with the starting point of a short video with the techniques to create them. All 5 smartphones were successfully attacked by an inexperienced attacker. This paper will provide the results achieved, as well as an analysis on the attack potential of every case. All results are given following the metrics of the standard ISO/IEC 30107-3.

1. Introduction

Biometric recognition has become a commonplace technology in our everyday lives. We use it to unlock our phones, to get in the gym or to enter the office due to its comfort of use [1] and freeing us from remembering passwords [2, 3]. Nevertheless, when sensitive data (personal pictures, documents) or privileges (gym membership, food coupons) are at stake, we need to make sure that we know how secure these systems are.

There have been many tests of many fingerprint sensors' ability to detect attacks in the last decades [4–12]. For that end, different materials were used to create fake fingers and use them on sensors, to check if an attacker would be able to bypass the security. When the first smartphones with embedded fingerprint sensors were released, fake fingers were created to try to attack the system and succeeded [13]. It was rapidly spread on the media that biometric recognition was not secure and that people should not trust it. Soon after, it was proven that an attacker can steal a person's fingerprint by taking a picture of it from a distance [14].

The simulated attackers that performed these tests were, to the best of our knowledge, researchers or proficient in biometrics. Moreover, the released videos only show one attempt, which showed to be successful. However, how many attempts did they need until it worked? How long had they been working on biometrics? What was their expertise? These questions were left unanswered.

To answer them, it is necessary to follow a common standard to have comparable security evaluations and to give a complete understanding of how the systems behave against attacks. This is what was achieved with this work. To fulfill the need of comparable security evaluations, there are several tools like Common Criteria [15] and its evaluation methodology, CEM (Common Methodology for Information Technology Security Evaluation) [16]. These are focused on Information Technology security in general and need some adaptation for the case of biometrics, and more particularly

for the attacks at the sensor capture level, the so-called presentation attacks. Hence, a new ISO standard was created, ISO/IEC 30107 [17], to address this need for Presentation Attack Detection (PAD) evaluation standardization. Also, some works have been done on methodologies and best practices to evaluate security [18, 19] and to evaluate the performance of sensors embedded in smartphones [20]. A methodology unifying both was proposed in [21]. This work gathers 3 studies following thoroughly the said standards, in order to give a complete answer to the attack resistance ability of these systems.

First, an experiment was made with 4 desktop fingerprint sensors: 1 thermal, 2 capacitive, and 1 optical. In total, 4,672 attacks were attempted using 7 different artefact species (fake finger material) [22]. Both cooperative and noncooperative tests were made. Cooperative attacks require the collaboration of the capture subject to get their biometric characteristic on a mold, while, for noncooperative attacks, the attacker gets the fingerprint from a surface without the collaboration of the capture subject.

For the second study, the 3 most interesting artefact species (according to ease of production, success on attacks, or level of resemblance to real finger) from the first study were chosen to attempt attacks on 5 smartphones that have embedded fingerprint sensors, summing a total of 2,669 attacks [21]. This experiment was performed by the same evaluator of the first experiment, thus having gathered knowledge on how to perform attacks. All the artefacts were created in a collaborative manner in this case.

Finally, for the third study, we gave 15 simulated attackers with no background in biometrics one week to attack one smartphone's fingerprint sensor (5 smartphones in total, same ones than in the second study). Each had to, at least, use 3 bona fide capture subjects (but they could use more for extra credit) and use each material at least 120 times on the smartphone sensor, making a total of 5,841 attempts. As more than one week would be needed to create noncooperative fake fingers, the study was focused only on cooperative attacks.

With these 3 studies following the same standard (ISO/IEC 30107) and methodology, it is possible to make a comparison of the attack potential for each case and get results and thus answer if attacking fingerprint sensors is a matter of expertise and how many attempts are needed to successfully attack a sensor, in average.

This paper is divided into 6 sections. Section 1 includes an introduction and Section 2 gives an overview on related work. Section 3 is focused on the methodology carried out during the evaluations and Section 4 analyzes Common Criteria's attack potential for each case. Results are reported and discussed in Section 5 and Section 6 describes the conclusions that were obtained during this work, as well as future work.

2. Related Work

There can be many vulnerable points in a biometric system: at presentation level, identity claim, data transfer, quality and feature extraction, decision thresholds, and so on [23]; and this also applies to mobile devices. For instance, some vulnerabilities were found creating a malicious application that steals the temporary fingerprint image by accessing its memory space or extracting a stored template from the nonvolatile memory and recreating the feature points of the fingerprint [24]. In [25, 26], privacy issues are addressed for biometric user authentication, and they propose some countermeasures for a properly designed secure and privacy-preserving system.

In addition, several security analyses have been made using altered fingerprints [27, 28] and one was performed specifically on mobile devices [29]. Nevertheless, it differs slightly from the topic of this paper, as it is focused only on alterations to fingerprint sample images, feeding them directly to the system, and not on creating artefacts and molds.

Many studies and evaluations have been performed regarding presentation attacks on fingerprint sensors. Already in 1990, several sensors started being tested using artefacts, and the system failed to reject them even from the first attempt [4, p. 15]. On 2000, an evaluation was performed on [10, p. 9] by calculating the acceptance rate of 1 user's finger made with gelatin on 11 sensors, where the artefacts were accepted by the systems in a very high percentage (the lowest being 67% fake finger acceptance rate). On 2002, several more attacks were demonstrated by using latent fingerprint reactivation on 6 capacitive, 2 optical, and 1 thermal scanners [30]. For the case of [7], 10 subjects were used to create gelatin fingers and use them on 3 sensors, getting success rates from 44.6% to 76% success rates. In all experiments, only index fingers were used. In general, nevertheless, these studies do not follow a thorough evaluation procedure nor standard and merely prove when a certain material or technique is effective on specific sensors at least once.

In 2009 [31], Liveness Detection *(LivDet)* competitions started and continued in 2011 [32], 2013 [33], and 2015 [34]. Their goal was to compare different liveness detection (Presentation Attack Detection) mechanisms by using them on a very large database of fake fingers (made of gelatin, latex, ecoflex, Play-Doh, silicone, wood glue, and modasil). Different academic institutions or industries could try their algorithms on the database. Four different sensors were used to acquire the images and the evaluations were done using a common testing protocol.

To the authors' knowledge, there have not been evaluations specifically focused on attacking mobile devices with fake fingers, but there have been reports on found vulnerabilities. In 2013, when the first iPhone with an integrated fingerprint sensor came out, the Chaos Computer Club [13] proved that it was possible to break into the sensor using a white glue fake finger covered with graphite, and the fingerprint could be stolen from the phone screen using a scanner and doing some image processing. Nonetheless, this was only reported once in a video, not in a complete evaluation. In 2016, fake fingers were printed using conductive ink (having a sample of the fingerprint image beforehand), so they could be used directly on the mobile phone sensor without having to create molds previously [35]. This was a technical report to inform about the vulnerability and not an evaluation.

There are several ways to overcome these attacks, divided into two main groups: software and hardware PAD mechanisms [5]. Software PAD mechanisms take the sample captured by the sensor and apply image processing and classification techniques to tell whether the finger is real or not. On the other hand, hardware PAD adds additional sensors (temperature sensor, multispectral cameras, etc. [36]) to make this distinction. Hardware solutions have lower error rates than the software ones [37] but are usually more expensive or bulky due to the additional equipment needed [36, p. 10]. Thus, hardware solutions are usually not included in mobile devices that should be cheaper and smaller.

3. Materials and Methods

The methodology was as homogeneous as possible across all studies, following evaluation methods from Common Criteria's CEM [16] and ISO/IEC 30107 [17], which were unified as a methodology on [21]. As for every biometrics evaluation, three steps are needed: planning, execution, and results reporting. In this case, the results will be given later in Section 5.

3.1. Planning the Evaluation. In the case of the desktop fingerprint sensors, an expert performed an evaluation on 1 thermal, 2 capacitive, and 1 optical sensors. The process used to create the artefacts was the usual one seen in many evaluations and research on PAD [10, pp. 5–8], both cooperative (capture subject cooperates in the creation of the mold) and noncooperative (attacker steals biometric characteristic with no help from capture subject). For this study, it was important to try many artefact species (i.e., Play-Doh, gelatin, latex, silicone, white glue, latex with graphite, and silicone with graphite) to check which ones were more threatening to the systems.

For the second study, an expert performed an evaluation on 5 mobile devices with an embedded fingerprint sensor. The artefact species (materials) that turned out to be most threatening for the previous evaluation were used for this one: Play-Doh, gelatin, and latex with graphite.

Lastly, for the case of the third study, we gathered 15 simulated attackers with no background in attacking biometric systems. We prepared a process or "recipe" they had to follow to create the artefacts and we gave it to them on writing and video. They had to create artefacts with Play-Doh and gelatin but could get extra points for using more artefact species. Again, there were 5 smartphones available for the evaluation (same ones as in study 1) and each simulated attacker was given one at random. Also, those who owned one of these cell phones could use them for the attack. Thus, some smartphones were used more than others. They had one week to perform at least 120 attack attempts per artefact species. To make sure that they were performing the evaluation correctly, attackers had to record themselves on video using the fake fingers on the sensors. Moreover, they had to take pictures of all molds and artefacts and finally had to hand over a box with all of them.

Table 1: Characteristics of desktop fingerprint sensors according to sensor interaction type and sensing technology.

Fingerprint sensor	Sensor type	Technology
THER	Touch	Thermal
CAP1	Touch	Capacitive
CAP2	Touch	Capacitive
OPT	Touch	Optical

Table 2: Characteristics of the mobile device sensors used for the evaluation, according to sensor type, shape, and location on the device.

Mobile device (MD)	Sensor type	Sensor shape	Sensor location
MD1	Swipe	Rectangular	Front
MD2	Touch	Circular	Front
MD3	Touch	Rectangular	Front
MD4	Touch	Rectangular	Side
MD5	Touch	Circular	Back

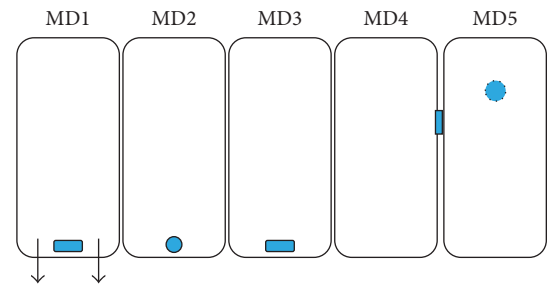


Figure 1: Fingerprint sensor placement on smartphones.

On the following subsections, the TOE (Target of Evaluation) and target application will be described for each study and, with these in mind, the penetration test will be specified.

3.1.1. Description of the TOE (Target of Evaluation). Four fingerprint sensors with different sensing technologies were used for the first study. All of them were a gray box, because the only intermediate result that could be obtained was the quality score, measured by NIST's NFIQ quality algorithm [38]. The brands cannot be disclosed, but their main characteristics can be seen in Table 1.

Both PAD evaluations on mobile devices were performed on 5 smartphones with different embedded fingerprint sensors. Their main characteristics (sensor type, shape, and location) can be seen in Table 2 and Figure 1.

As the biometric systems evaluated on these 2 studies are full systems embedded in mobile devices, they are a black box that only reveals whether the verification with the artefact passed or failed.

3.1.2. Description of Target Application. A PAD evaluation has no meaning unless the target evaluation and the conditions

under which it is performed are specified. These differ in the case of desktop sensors and smartphones.

(a) Desktop Fingerprint Sensors. In the case of desktop fingerprint sensors, there are many possible applications, for example, entering an office, spending food coupons at work, or entering critical infrastructures (factories and nuclear plants). Thus, the consequences of an attacker hacking those sensors would be having an unauthorized person enter a certain building, which could be more or less critical, or having access to privileges that do not belong to the person.

The implemented functions for the systems are enrolment and verification, being their policies in detail:

(i) Enrolment: all images of the artefacts were captured in an acquisition process. There were 2 transactions to obtain images with an NFIQ (NIST Fingerprint Image Quality) value equal or lower than 3 (good quality), and 3 attempts were made for each transaction until both images were successfully compared. If after all the attempts, no successful comparisons could be done between the two samples, the enrolment for this particular finger could not be completed.

(ii) Verification: the verification was done offline. The artefacts were captured in the same enrolment process as the real fingers, so they could be discarded by NFIQ higher than 3 (low quality). Then, in an offline process, the artefact samples were compared with the NBIS algorithm to the real fingers.

(b) Mobile Phones. The target application in the case of smartphones is to unlock the user's mobile phone. From there, all apps can be accessed (as only a few require additional security, like a PIN), including bank accounts, password managers, or personal pictures. Moreover, some apps have implemented fingerprint logging-in by using the system's fingerprint manager. For the case of smartphones, there is never any surveillance by a guard (as it is used solely by the owner). The implemented functions in this case are also enrolment and verification, with the particularity that the evaluator cannot decide on the policies.

(i) Enrolment: it is crucial for the evaluation, as it can influence the performance, and each mobile device has a different enrolment policy. As it is a black box, the evaluator cannot decide on the enrolment and verification policies, as can be seen in Table 3.

(ii) MD5's policy must be noted, as it only needs 6 samples to do the enrolment and that could influence the final performance of the fingerprint sensor.

(iii) Verification: the artefacts are used in a verification process, that is, the artefact attempts to be verified as the real finger that has been previously enrolled. The number of allowed attempts for each smartphone is detailed on Table 4.

As it can be seen, devices MD1, MD3, MD4, and MD5 accept an unlimited number of attempts to attack them, as when 5 attempts have failed, they just wait for 30 seconds and the attacker can try again, with no more restrictions (as far as the author's knowledge). MD2 asks for a PIN after 3 attempts.

Also, for all cases, if the phone is turned off and the attacker wants to turn it on and access its data, he or she will need additional information apart from the bona fide user's fingerprint, like a PIN or a password. So, if the phone is found turned off or without power and the attacker does not know the user's additional information, he or she will not be able to gain access.

Table 3: Enrolment policy for each mobile device. Policies are given by the manufacturer and cannot be changed by the evaluator.

Mobile device	Attempt policy for enrolment
MD1	Around 10 attempts needed (can be extended 10 more times if the user wants). Asks the user to move finger to get different samples and *checks it*.
MD2	Around 15 attempts needed (10 for center part of the finger and 5 for corners). Asks the user to move finger to get different samples and *checks it*.
MD3	Around 20 attempts needed. Asks the user to move finger to get different samples and *checks it*.
MD4	Around 20 attempts needed. Asks the user to move finger to get different samples and *checks it*.
MD5	Around 6 attempts needed. It just asks the user to move the finger to get different samples, but *does not check it*.

Table 4: Transaction policies for verification for each mobile device. Policies are given by the manufacturer and cannot be changed by the evaluator.

Mobile device	Allowed failed attempts	Policy
MD1	5	Waits for 30 seconds, can do attack again
MD2	3	Asks for PIN
MD3	5	Waits for 30 seconds, can do attack again
MD4	5	Waits for 30 seconds, can do attack again
MD5	5	Waits for 30 seconds, can do attack again

3.1.3. Specification of Penetration Test. Once the systems under test have been analyzed and described, it is necessary to specify the penetration test and how it will be performed. In Table 5, the characteristics of the test are shown, as well as the final amount of attempts for clarity purposes.

For the experiment of the desktop sensors, an app was developed for the capture process. As it can be seen in Figure 2, the program showed which finger to capture, the amount of attempts, and samples left and an image of the captured fingerprint. It allowed an enrolment and verification process.

For the experiment of the smartphones, a mobile app was made for iOS and Android (Figure 3). For acquiring the data, the visit screen is filled in by the evaluator (genuine user's ID, attacker's ID, device, finger ID, type of attack, and artefact species) and the app logs whether the attack succeeded or not. The enrolment was performed using the phone's native settings procedure.

Table 5: PAD evaluation characteristics for each study: desktop sensors, mobile devices (1 expert), and mobile devices (15 laymen). Details are given according to ISO/IEC 30107-3 requirements.

	Desktop sensors	Mobile devices (1 expert)	Mobile devices (15 laymen)	
Capture subjects	6	7 (3 female + 4 male)	3 (more for extra points)	53
Sources for artefacts	36 fingers, 6 capture subjects (index, middle and thumb, both hands)	42 fingers, 6 capture subjects (index, middle and thumb, both hands)	4 per capture subject (both thumbs, indexes)	212
Number of attempts	3 per transaction, 2 transactions per finger, 4 sensors (4672 in total)	10 per finger, 5 mobile devices (2669 in total)	10 per finger	5,481
Evaluator's access	Gray box (pass/fail result + NFIQ quality).	Black box (only pass/fail result). Very slight quality feedback on some devices (MD1 and MD3) ("finger too wet").		
Artefact species	7 (Play-Doh, gelatin, latex, silicone, white glue, latex with graphite, silicone with graphite)	3 (Play-Doh, gelatin, latex with graphite)	2 (Play-Doh and gelatin), more for extra credit	10 (Play-Doh, gelatin, gelatin with glycerin, silicone, TAC adhesive, wax, clay, wood glue, conductive ink, latex)
Attack type	Cooperative and non-cooperative	Cooperative		

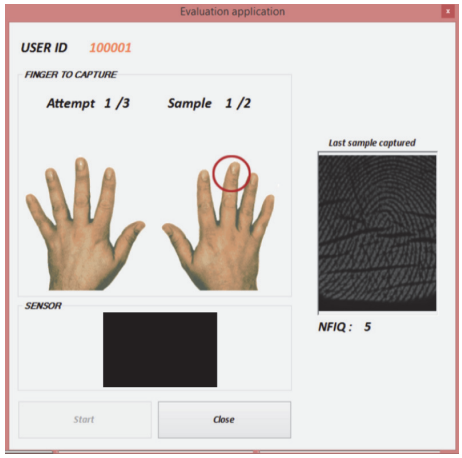


Figure 2: Example display of the desktop capture program.

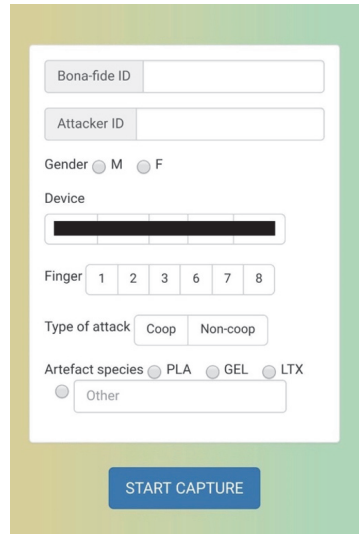


Figure 3: Smartphone app (Android and iOS) for logging the PAD evaluation. A pass/fail result is logged for each attempt.

3.2. Executing the Evaluation. After the careful planning of the evaluation, it was finally executed with the specified penetration test. The 3 steps needed for the execution according to CEM [16] are detection, capture, and processing.

3.2.1. Detection. Before the actual execution, the different artefact species were put to the test. If the sensor could detect the fingerprint, then it could be selected for the evaluation. Moreover, for the third study, simulated attackers tried more materials apart from Play-Doh and gelatin but found that not all of them were suitable.

3.2.2. Capture. As checking the quality is not always possible for the evaluator, some artefacts with different qualities (examined by the evaluator) can be used to check which ones are obtained successfully by the sensor and continue with that technique [39]. For desktop sensors, the quality of the sample could be measured with NFIQ and an image of the sample could be seen at the moment of capture. Thus, it was possible for the evaluator to improve the attack during the evaluation. In the case of mobile devices, some of them give a slight feedback on quality by telling the user that, for example, the finger is too wet. If an attacker is using a gelatin artefact and the smartphone prompts "finger too wet," the attacker will make another artefact with less proportion of water and try again.

Lastly, it must be noted that each mobile device will have different algorithms, sensor technologies, and quality and decision thresholds.

3.2.3. Processing. In the case of desktop fingerprint sensors, all images were stored for an offline verification with the database of subjects' real fingers by using the NBIS algorithm by NIST [40]. The most significant values from the analysis

were the proportion of times that an artefact was verified as a normal presentation and the proportion of images that were rejected by the system due to low quality (NFIQ > 4).

In the case of smartphones, no quality or similarity scores can be obtained, just pass/fail results. Thus, the data used for the posterior analysis was based solely on this.

4. Analysis of Attack Potential

The attack potential is a standardized measure given by Common Criteria. According to Common Criteria methodology [16], the attack potential is a measure of the *effort to be expended in attacking a TOE (Target of Evaluation) with a PAI (Presentation Attack Instrument), expressed in terms of an attacker's expertise, resources, and motivation*, which can be divided into more specific parameters. Thus, TOEs are given a rating to assess their resistance to specific attacks.

4.1. Threats and Attacks. Every threat has a corresponding possible attack [18] and they are analyzed before calculating the attack potential.

4.1.1. Desktop Fingerprint Sensors

(i) Possible Threats. Although fingerprint sensor systems can have vulnerabilities at many points, this study only focuses on the presentation attack side, that is, using an artefact generated from a user's real finger on the sensor. The intended operation of the system depends on the target application, for example, opening a door to an office or to a gym or getting privileges like food coupons.

(ii) Possible Attacks. The attack that can exploit the threat explained in the previous point is the presentation attack. The biometric characteristic can be obtained in two ways: with or without cooperation from the capture subject. In this case, both cooperative and noncooperative attacks were done. The level of expertise of the evaluator is proficient, although the materials needed for the evaluation can be found at any supermarket.

4.1.2. Mobile Phones

(i) Possible Threats. Fingerprint sensors embedded in mobile devices can have vulnerabilities at many points, too. The intended operation of the system is to unlock a smartphone, thus accessing private data. As it was said above, the only vulnerable point used for the scope of this paper is the capture process.

(ii) Possible Attacks. The attack that can exploit the threat from the first point is the presentation attack. In this case, only cooperative attacks were done. The level of expertise of the simulated attackers is low: they had no prior knowledge on how to attack fingerprint sensors.

4.2. Attack Potential Calculation. This calculation is used by the evaluator to determine whether or not the TOE is resistant to attacks assuming a specific attack potential of an attacker [16, pp. 422–432]. If the evaluator determines that a potential vulnerability is exploitable in the fingerprint sensor, they must confirm that it is exploitable by doing penetration tests (as specified on Section 3.1.3).

With this in mind, the evaluator determines the minimum attack potential required by an attacker to successfully carry on an attack and arrives at some conclusion about the TOE's resistance to attacks. This attack potential is confirmed on the penetration tests performed in this work on Section 5.2.

A score can be assigned to each of the attack potential parameters following Common Criteria's CEM (Common Evaluation Methodology) [16, p. 429]. By adding all the values from the different parameters, the attack potential of an artefact species is rated as basic, enhanced-basic, moderate, high, or beyond high.

4.2.1. Desktop Fingerprint Sensors. Following CEM specifications, a score was given to every parameter of the attack potential to calculate its total rating (each score is given according to the table on [16, p. 429]). The attack potential will be different for cooperative and noncooperative attacks, being the expertise and the elapsed time the most differentiating factors.

(a) Cooperative Attacks. See Table 6.

(b) Noncooperative Attacks. See Table 7.

4.2.2. Mobile Phones. In this case, only cooperative attacks were made, and the attack potential will be the same for both mobile device studies.

As it was calculated in Tables 6, 7, and 8, the rating for cooperative attacks on desktop fingerprint sensors is 6.5 (basic) and 10 (enhanced-basic) for the noncooperative. For the case of mobile phones, the rating is 4.5 (basic). These scores are specified on CEM. Thus, the attacks would have to be considered in penetration testing for all evaluations assuming, respectively, minimum, basic, and minimum attack potentials (or higher). If penetration tests show that the attack is successful, the TOE would fail to resist against that attack potential. Further details on how to calculate the attack potential are given on Common Criteria's CEM [16, pp. 422–432].

5. Results and Discussion

Lastly, after executing the 3 separate evaluations, the obtained results are analyzed and compared in this section.

5.1. Metrics. The standard ISO/IEC JTC1 30107-3 requires specific metrics for PAD evaluation reporting. As the access to the system differs in desktop sensors and smartphones, Figure 4 shows which metrics are suitable for each case. The possible metrics are as follows:

(i) *APNRR (attack presentation nonresponse rate)*: proportion of attack presentations using the same PAI species that cause no response at the PAD subsystem or data capture subsystem

Table 6: Attack potential calculation for cooperative attacks on desktop fingerprint sensors. Scores assigned according to the classification from Common Criteria [16, p. 429].

	Preparation phase	PAI construction + exercising phase	Attack execution phase	Total factor rating	Score
Elapsed time	<1 day (capture subject is cooperative)	<1 day or <1 week (different material difficulty)	Few seconds (perform attack)	<1 week or <2 weeks	1.5
Expertise	Layman (materials can be obtained at normal stores)	Layman (easy to create)	Layman (not much expertise needed)	Layman	0
Knowledge of TOE	Public (well known on the internet that it works)	Public (manuals can be found on the internet)	Public (no knowledge needed)	Public	0
Window of opportunity	Unnecessary (no access to TOE needed)	Easy (access to TOE for practicing)	Easy (high chance the PAI will work)	Easy	1
Equipment	Standard (no equipment needed)	Specialized (some sensors might be hard to obtain)	Standard (no equipment needed)	Specialized	4
Overall attack rating	*6.5 (Basic)*				
Attack resistance	*Minimum*				

Table 7: Attack potential calculation for noncooperative attacks on desktop fingerprint sensors. Scores assigned according to the classification from Common Criteria [16, p. 429].

	Preparation phase	PAI construction + exercising phase	Attack execution phase	Total factor rating	Score
Elapsed time	<1 week (capture subject is noncooperative)	1 week (creating PAIs)	Few seconds (perform attack)	<2 weeks	2
Expertise	Layman (materials can be obtained at normal stores)	Proficient (process needs many steps)	Layman (not much expertise needed)	Proficient	3
Knowledge of TOE	Public (well known on the internet that it works)	Public (manuals can be found on the internet)	Public (no knowledge needed)	Public	0
Window of opportunity	Unnecessary (no access to TOE needed)	Easy (access to TOE for practicing)	Easy (high chance the PAI will work)	Easy	1
Equipment	Standard (no equipment needed)	Specialized (some sensors might be hard to obtain)	Standard (no equipment needed)	Specialized	4
Overall attack rating	*10 (Enhanced-basic)*				
Attack resistance	*Basic*				

(ii) *APAR (attack presentation acquisition rate)*: proportion of attack presentations using the same PAI species from which the data capture subsystem acquires a biometric sample of sufficient quality

(iii) *IAPMR (impostor attack presentation match rate)*: proportion of impostor attack presentations using the same PAI species in which the target reference is matched. *When it is not matched*, IAPNMR *is used (Impostor Attack Presentation Nonmatch Rate)*.

5.2. Penetration Test Results. With the metrics described on the last subsection and the schema from Figure 4, we built the graphs for the error rates of desktop and smartphone fingerprint sensors.

Table 8: Attack potential calculation for cooperative attacks on smartphone fingerprint sensors. Scores assigned according to the classification from Common Criteria [16, p. 429].

	Preparation phase	PAI construction + exercising phase	Attack execution phase	Total factor rating	Score
Elapsed time	<1 day (capture subject is cooperative)	<1 day or <1 week (different material difficulty)	Few seconds (perform attack)	<1 week or <2 weeks	1.5
Expertise	Layman (materials can be obtained at normal stores)	Layman (easy to create)	Layman (not much expertise needed)	Layman	0
Knowledge of TOE	Public (well known on the internet that it works)	Public (manuals can be found on the internet)	Public (no knowledge needed)	Public	0
Window of opportunity	Unnecessary (no access to TOE needed)	Easy (access to TOE for practicing)	Easy (high chance the PAI will work)	Easy	1
Equipment	Standard (no equipment needed)	Standard (but it is necessary to buy the TOE, which can be expensive)	Standard (no equipment needed)	Standard	2
Overall attack rating			*4.5 (Basic)*		
Attack resistance			*Minimum*		

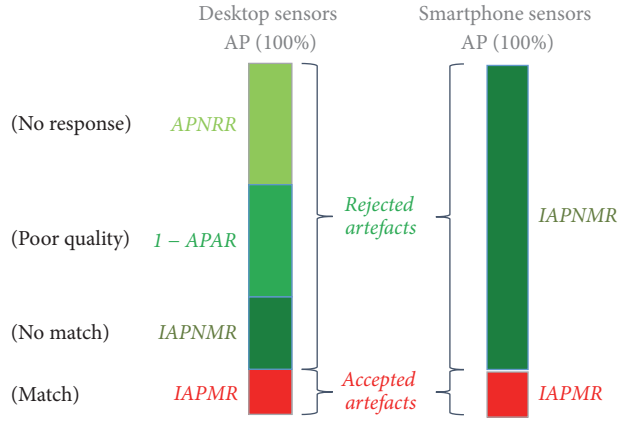


Figure 4: Metrics used for each case. Desktop fingerprint sensors give us more intermediate decisions than mobile devices.

5.2.1. Desktop Fingerprint Sensors. As these systems gave us feedback on IAPMR, APAR, and APNRR, they are represented in Figures 5 and 6.

The most meaningful value is IAPMR, as it shows the proportion of presentation attacks that defeat a comparison system. This metric exposes that the only material that can successfully attack all systems is Play-Doh, especially for the case of the optical sensor. The highest IAPMRs were obtained with silicone mixed with graphite, Play-Doh, and white glue.

For noncooperative attacks, the only vulnerable sensor was the thermal one for the case of silicone, latex, and Play-Doh, in very few times. It must be noted that very few attacks were performed in this manner, as it was decided to use the PCB molds that were created on the first try, with no room for improvement, to see the results of a first-time attacker for this case.

In addition, the greater the APNRR is, the better the system is at rejecting fake samples (by not responding when they are placed on the sensor), so in this matter, the thermal sensor responds to the highest number of artefacts, although those captured samples ended up not being successful. Systems can also reject artefacts due to their low quality, and this ability is represented by APAR. In this case, the capacitive sensors were more capable of rejecting nonconductive samples, even when breathing on them to create a conductive layer on the surface.

5.2.2. Mobile Phones (Studies 2 and 3). In this case, only IAPMR was known, so the graphs are simplified by only showing this metric and omitting IAPNMR (its contrary).

First, the overall IAPMR results are shown in Figure 7. This first graph is shown inside the corresponding smartphone shape for clarification. The rest will be shown as usual for the sake of space. The figure compares the outcome obtained in study 2 (1 expert) with the average outcome obtained in study 3 (15 laymen attackers).

As it can be seen, the experienced attacker successfully attacked the smartphones more often than the inexperienced attackers, *in average*. The results for MD3 are quite similar. Nevertheless, the IAPMR for MD4 is higher for the inexperienced attackers than for the experienced one. This outcome will be explained later in this section.

The common materials for both studies were Play-Doh and gelatin, as they were tried on every smartphone by every attacker. Thus, Figure 8 shows a comparison of IAPMR results by an experienced attacker versus the average of 15

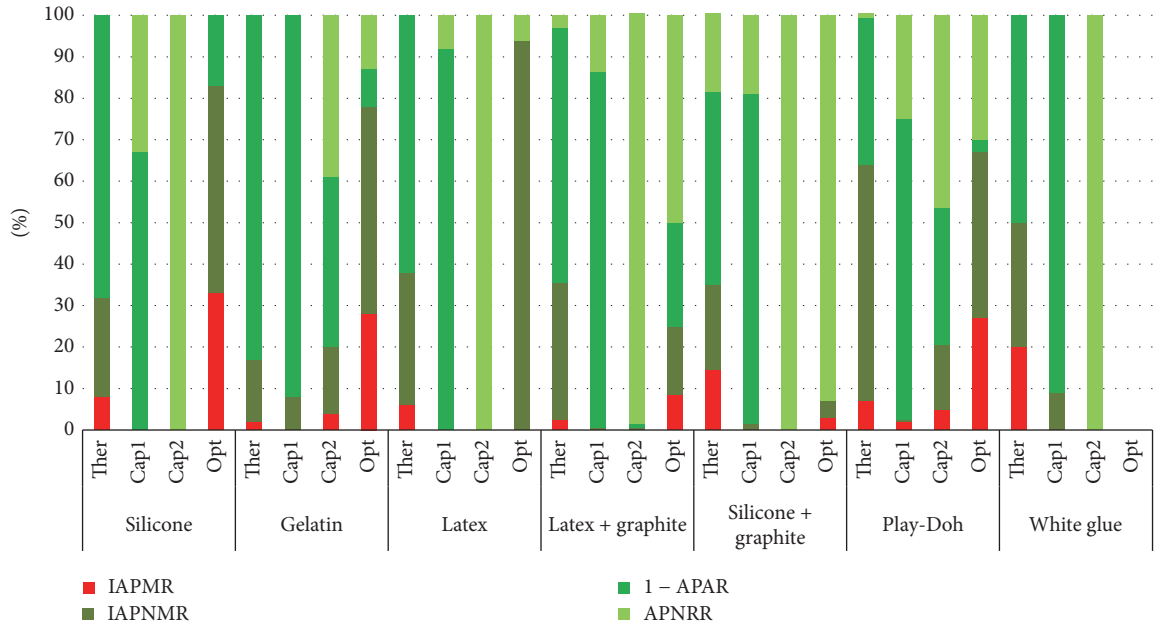


Figure 5: Cooperative attack results for desktop fingerprint sensors, separated by device and artefact species (material).

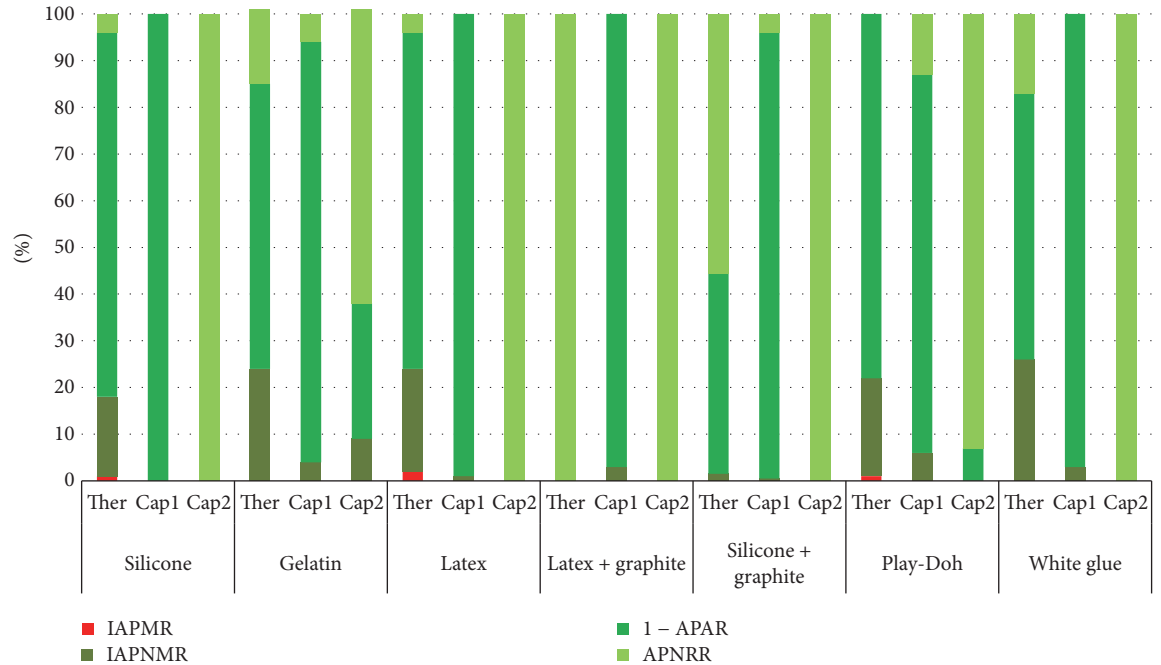


Figure 6: Noncooperative attack results for desktop fingerprint sensors, separated by device and artefact species (material).

unexperienced attackers. It can be clearly seen that the most vulnerable device is MD5, for both Play-Doh and gelatin. The experienced attacker had more successful attacks in most cases.

A break-down of the IAPMR results by attacker can be seen in Figure 9. The experienced attacker could hack into MD1 19% of the times, while the inexperienced had an almost negligible number of successful attempts. MD2 was tricky to hack for all attackers, although Att 1 did a slightly better job. It can be observed that, for the case of MD3, IAPMR varies quite a bit depending on the inexperienced attacker, although the average is similar to the experienced one. On the case of MD4, the smartphone was hard to hack for most attackers (even for Att 1), but the inexperienced attacker Att 14 could break into

Table 9: Results on the vulnerabilities of each device to different artefact species. Not all materials were used in all sensors. A tick means that the device was successfully attacked with that species at least once. A cross indicates that the device could not be attacked with that material even once. An interrogation was used when the experiment was not tried.

	Vulnerable devices								
Artefact species	Ther	Cap1	Cap2	Opt	MD1	MD2	MD3	MD4	MD5
Play-Doh	√	√	√	√	√	√	√	√	√
Gelatin	√	x	√	√	x	√	√	√	√
Clay	?	?	?	?	?	?	x	?	?
Wood glue	√	x	x	x	?	?	√	√	?
Conductive ink	?	?	?	?	?	?	x	?	?
Latex	√	x	x	x	?	?	√	?	√
Latex + graph	√	x	x	√	?	√	√	?	?
Silicone	√	x	x	√	?	?	x	?	√
Silicone + graph	√	x	x	√	?	?	?	?	?

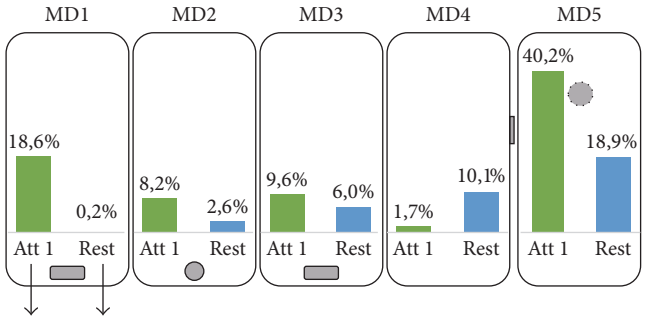


Figure 7: Overall IAPMR comparison per smartphone. Att 1 refers to the expert that performed the 2nd experiment, while Rest refers to the 15 simulated attackers.

it 19% of the times. Lastly, MD5 was the one with the most uneven results: Att 1 attacked the system successfully 40% of the times, while Att 15 did 25% of the times and Att 16, none.

The 15 simulated attackers had a chance to get additional credit for the assignment if they used additional artefact species and reported them. The results are shown in Figure 10, showing that it was discovered that white glue was the most successful material on MD3. Moreover, latex and silicone were found to be more successful than Play-Doh and gelatin on MD5.

In [21], it was said that attacking fingerprint sensors with fake fingers depends on expertise, but that luck also has a great impact. Sometimes, after trying to attack a sensor many times, the tester moves the finger slightly differently, or adds more water to the mix, or heats the artefact more and suddenly the fake finger works. This could also happen the first time a fake finger is used. Once the trick is known, the rest of the attempts will be much easier.

5.2.3. All Experiments. The only common metric for all experiments is IAPMR, which is the most significant one, as it shows the proportion of times that an artefact was verified as the real finger. It is shown for all devices across the 3 studies on Figure 11.

It can be observed that the highest IAPMR was obtained by the expert attacker on average on study 2, reaching a value of 40,2% for MD5. The lowest IAPMR on average happened on study 1 with the desktop sensors, as the policies for capture were stricter than those of the mobile devices. Our feeling when performing the evaluation on mobile devices after having done the same on desktop sensors was that it was much easier to bypass their security in comparison, which can be clearly noticed on the results.

5.3. Vulnerability Test Results. Although not all materials were tried on all devices, some vulnerabilities can be reported from the results obtained in the previous section (Table 9). A device being vulnerable means that it was hacked at least once with that artefact species. It can be observed that Play-Doh could successfully attack all devices at least once.

The consequences that may derive from these vulnerabilities are that an attacker could enter an unauthorized building, have access to privileges that do not belong to him/her, or for the case of the smartphones, unlock the phone, and have access to all the apps that do not require additional security.

5.4. Artefacts, Molds, and Captured Images. As it happens with cooking, every evaluator has different abilities to generate artefacts and it is very difficult to compare one's expertise to another. Therefore, the results of the security evaluation are dependent on who the attacker is. Thus, for a report, the best we can do is show which molds and artefacts were used in the evaluation.

5.4.1. Molds. Even inside the same evaluation, there are factors that condition the quality of the generated molds, and this influences highly the results. For example, some people tend to sweat a lot if their finger is surrounded by silicone paste for a few minutes, so the molds turn out with bubbles and with a very low quality (Figure 12).

5.4.2. Artefacts. As with the molds, it is important to include examples of the artefacts that were created of each artefact species. Figure 13 gives examples of the artefacts used for this experiment.

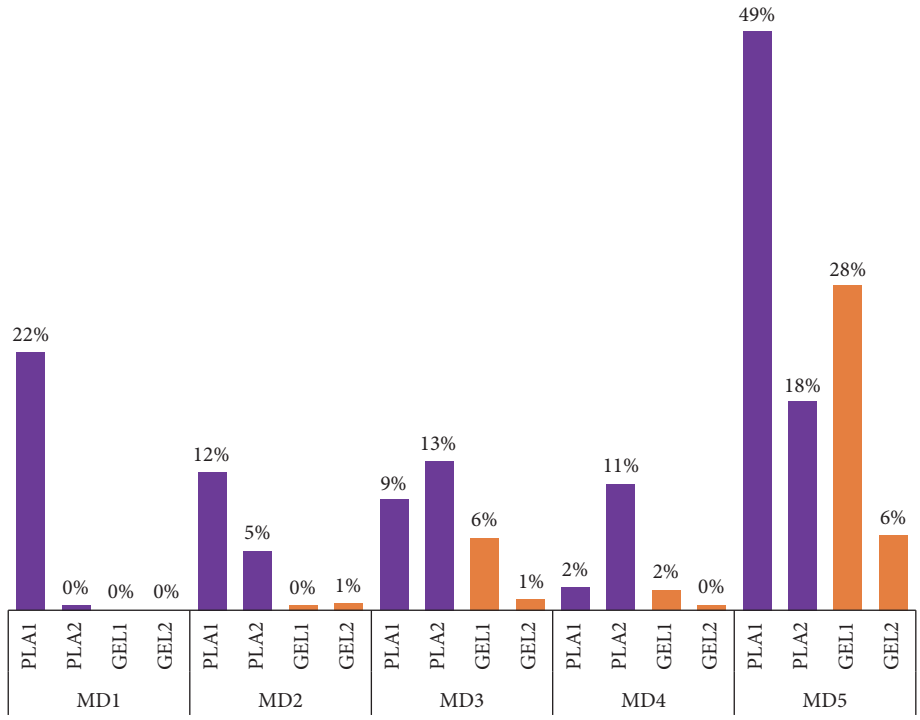


Figure 8: Comparison of IAPMR results by an experienced attacker (PLA1 for Play-Doh and GEL1 for gelatin) and 15 laymen attackers (PLA2 and GEL2).

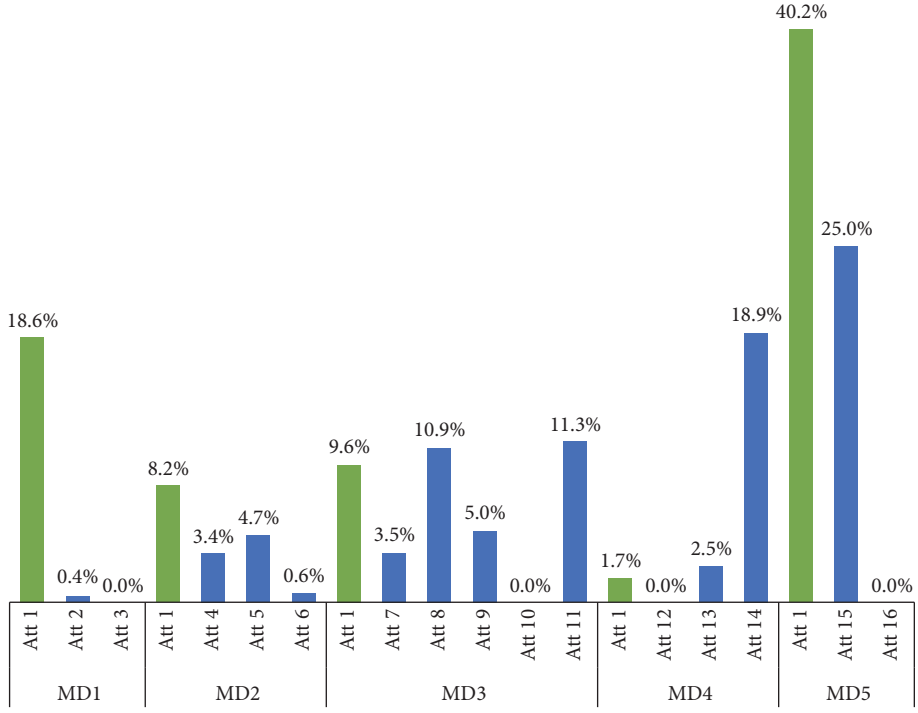


Figure 9: Break-down of IAPMR results by each of the 15 unexperienced attackers, separated by attacker and device.

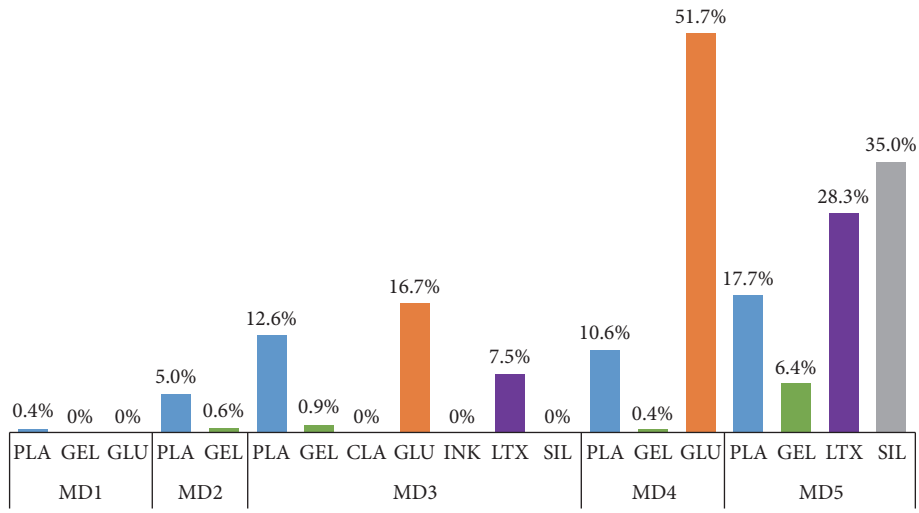


Figure 10: IAPMR results for the 15 inexperienced attackers, divided by all the artefact species used. PLA = Play-Doh; GEL = gelatin; CLA = clay; GLU = wood glue; LTX = latex; SIL = silicone.

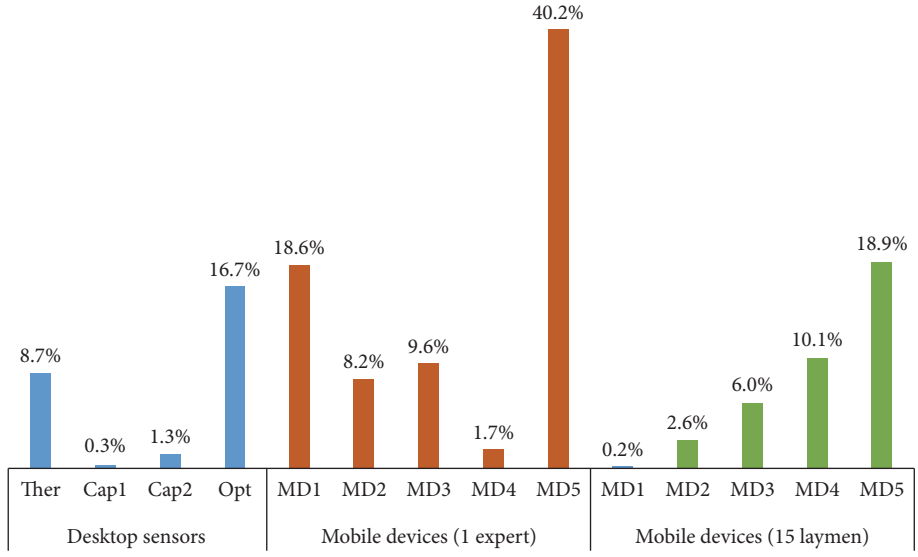


Figure 11: Overall IAPMR for all devices across 3 experiments. Desktop sensors in this case only cover cooperative attacks, for the sake of fair comparison.

5.4.3. Captured Images. The only images that we could get access to were the ones from the desktop sensors evaluation. The examples on Figures 14 and 15 are from the thermal sensor.

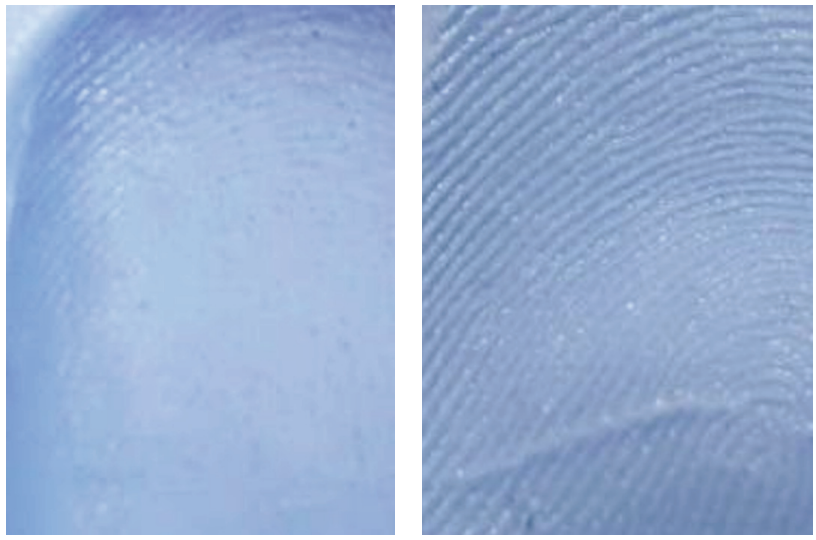
The cooperative artefact has a notably better quality than the noncooperative one, due to its acquisition process. Cooperative molds capture quite accurately the shape of the capture subject's fingerprint, while getting an accurate sample from a latent print on a glass can be trickier, as many features can be lost on the process.

It must also be noted that even, within the same evaluation, the evaluator's ability can improve. For instance, making a good quality gelatin artefact consists of having the right proportion of water and gelatin leaves, and some trials might need to be done before reaching that appropriate proportion. A clear example of this can be seen in Figure 15.

6. Conclusions

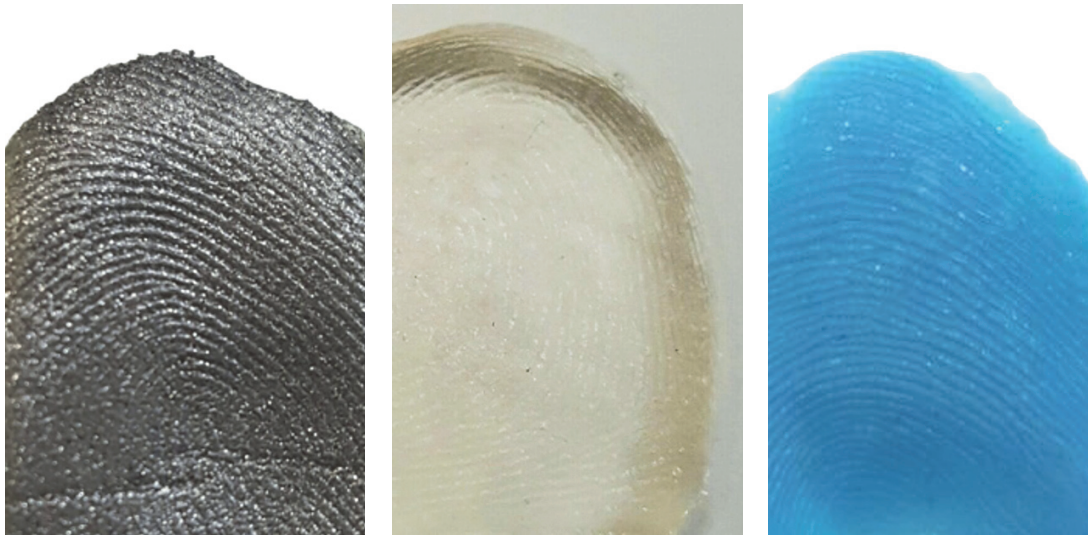
During the process of performing 3 separate PAD evaluations on fingerprint sensors, some lessons were learnt, the basic one being that standards and methodologies are necessary to compare PAD evaluations made by researchers and certification bodies. Studies are only comparable if the same metrics and procedures are used.

Enrolment and verification policies are very important for the performance and security of a system, and, with smartphones, these policies are out of hand for independent



(a) Bad quality mold (b) Good quality mold

Figure 12: Molds of different qualities. The mold on (a) has bubbles and is blurry due to the capture subject's finger characteristics.



(a) Latex with graphite artefact (b) Gelatin artefact (c) Play-Doh artefact

Figure 13: Examples of 3 artefact species: latex with graphite, gelatin, and Play-Doh.

researchers and they must adapt to the ones given by the manufacturer. For instance, for the enrolment of mobile device MD5, only 6 captures are needed (while other devices use more than 15), and this poor enrolment policy could be a reason why it is noticeably more vulnerable to attacks than the others. Also, some devices allow an unlimited amount of attempts to present a sample, giving the attacker unlimited chances. This can be fixed by asking for additional information (PIN, password, or an additional biometric modality).

Attack potential is an adequate tool to measure the effort needed to attack a system. Nevertheless, in the case of biometrics, it is difficult to calculate. Even within the same evaluation, the evaluator gets better at attacking the system in each attempt. Also, at any point of the process, anyone can get the trick to hacking a specific sensor (out of expertise and, mostly, out of luck) and the evaluation results can vary highly from that point on. To soften this variability, one solution is reporting examples of the molds, artefacts, and captured images of the evaluation.

Expertise was evaluated on this work. It was proved that all 5 smartphone sensors were hacked at some point by unexperienced attackers. The only prior knowledge they had was a short video of an expert creating fake fingers, so

(a) Cooperative (b) Noncooperative

Figure 14: Examples of captured images of cooperative and noncooperative artefacts from the desktop sensors evaluation.

(a) Bad quality artefact (b) Good quality artefact

Figure 15: Examples of a bad quality gelatin artefact versus a good quality gelatin artefact. Images from the desktop sensors evaluation.

this is especially preoccupying because similar videos can be obtained from the Internet easily. Fortunately, this was the case only for cooperative attacks, and their results were notably worse than the ones from the expert, in average. On the other hand, noncooperative attacks are more complex and need more expertise to be carried out. It was also noticed that it is easier to hack smartphone sensors than desktop sensors.

In the future, the database of inexperienced attackers will be increased, and further analyses will be made based on this. Moreover, more insights can be obtained in future evaluations: how long it takes for an inexperienced attacker to successfully attack a sensor for the first time, dependence on finger used (index, middle, and thumb), dependence on left or right hand, and so on.

Conflicts of Interest

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

Acknowledgments

This work was partially supported by the Spanish National Cybersecurity Institute (INCIBE) under the Grants Program "Excellence of Advanced Cybersecurity Research Teams."

Supplementary Materials

The supplementary video is the same file that was shared with the inexperienced attackers in order to perform their experiments. It shows the whole process from creating a mold from a real finger, to creating an artefact and using it on the smartphone sensor. *(Supplementary Materials)*

References

[1] R. Blanco-Gonzalo, R. Sanchez-Reillo, J. Liu-Jimenez, and C. Sanchez-Redondo, "How to assess user interaction effects in biometric performance," in *Proceedings of the IEEE International Conference on Identity, Security and Behavior Analysis (ISBA '17)*, February 2017.

[2] D. He and D. Wang, "Robust biometrics-based authentication scheme for multiserver environment," *IEEE Systems Journal*, vol. 9, no. 3, pp. 816–823, 2015.

[3] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2021–2040, 2003.

[4] T. van der Putte and J. Keuning, "Biometrical fingerprint recognition: do not get your fingers burned," in *Proceedings of 4th Working Conference on Smart Card Research and Advanced Applications, IFIP TC8/WG8*, vol. 31, p. 16, Bristol, UK, September 2000.

[5] S. A. C. Schuckers, "Spoofing and anti-spoofing measures," *Information Security Technical Report*, vol. 7, no. 4, pp. 56–62, 2002.

[6] A. Choiniere and T. Lubysheva, Novetta—Protecting against Fingerprint Vulnerabilities when Deploying Biometric Systems, 2015.

[7] J. Blommé, "Evaluation of biometric security systems against artificial fingers," 2003.

[8] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. A. Schuckers, "Review of the Fingerprint Liveness Detection (LivDet) competition series: 2009 to 2015," *Image and Vision Computing*, vol. 58, pp. 110–128, 2017.

[9] A. Wiehe, T. Søndrol, O. K. Olsen, and F. Skarderud, "Attacking fingerprint sensors," 2004.

[10] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial 'gummy' fingers on fingerprint systems," in *Proceedings of the Optical Security and Counterfeit Deterrence Techniques IV*, vol. 4677 of *Proceedings of SPIE*, pp. 275–289, January 2002.

[11] J. Galbally-Herrero, J. Fierrez-Aguilar, J. D. Rodriguez-Gonzalez, F. Alonso-Fernandez, J. Ortega-Garcia, and M. Tapiador, "On the vulnerability of fingerprint verification systems to fake fingerprints attacks," in *Proceedings of the 40th Annual IEEE International Carnahan Conference on Security Technology*, pp. 130–136, October 2006.

[12] S. Schuckers, "Presentations and attacks, and spoofs, oh my," *Image and Vision Computing*, vol. 55, pp. 26–30, 2016.

[13] Frank, "Chaos Computer Club breaks Apple TouchID," 2013, https://www.ccc.de/en/updates/2013/ccc-breaks-apple-touchid.

[14] A. Hern, "Hacker fakes German minister's fingerprints using photos of her hands," *The Guardian*, 2014.

[15] Common Criteria, "Common Criteria for information technology security evaluation part 1: introduction and general model September 2012 revision 4," ISO/IEC 15408 Common Criteria, Part 12012, International Organization for Standardization and International Electrotechnical Commission, 2012.

[16] Common Criteria, "Common methodology for information technology security evaluation evaluation methodology September 2012 revision 4 foreword," Tech. Rep. Ccmb-2012-09-004, 2012.

[17] ISO/IECJTC 1/SC37, *Text of FDIS 30107-3, Information Technology—Biometric Presentation Attack Detection—Part 3: Testing And Reporting*, vol. 2008, ISO-IEC Standards, 2009.

[18] B. Fernandez-Saavedra, R. Sanchez-Reillo, J. Liu-Jimenez, and J. G. Ruiz, "Best practices for the security evaluation of biometric systems," in *Proceedings of the 48th Annual IEEE International Carnahan Conference on Security Technology*, October 2014.

[19] B. Fernandez-Saavedra, R. Sanchez-Reillo, R. Alonso-Moreno, and C. Sanchez-Avila, "Evaluation methodology for fake samples detection in biometrics," in *Proceedings of the 42nd Annual 2008 IEEE International Carnahan Conference on Security Technology*, pp. 233–240, October 2008.

[20] B. Fernandez-Saavedra, R. Sanchez-Reillo, C. Sanchez-Redondo, and R. Blanco-Gonzalo, "Testing of biometric systems integrated in mobile devices," in *Proceedings of the 49th Annual IEEE International Carnahan Conference on Security Technology*, pp. 321–326, September 2015.

[21] I. Goicoechea-Telleria, J. Liu-Jimenez, R. Sanchez-Reillo, and W. Ponce-Hemandez, "Vulnerabilities of biometric systems integrated in mobile devices: an evaluation," in *Proceedings of the 50th Annual IEEE International Carnahan Conference on Security Technology*, October 2016.

[22] I. Goicoechea-Telleria, B. Fernandez-Saavedra, and R. Sanchez-Reillo, "An evaluation of presentation attack detection of fingerprint biometric systems applying ISO/IEC 30107-3," in *Proceedings of the International Biometric Performance Testing Conference*, 2016.

[23] T. Dunstone and N. Yager, *Biometric System and Data Analysis: Design, Evaluation, and Data Mining*, Springer, 2009.

[24] Y. Jo, S. Jeon, J. Im, and M. Lee, "Security analysis and improvement of fingerprint authentication for smartphones," *Mobile Information Systems*, vol. 2016, Article ID 8973828, 11 pages, 2016.

[25] E. Pagnin and A. Mitrokotsa, "Privacy-preserving biometric authentication: challenges and directions," *Security and Communication Networks*, vol. 2017, Article ID 7129505, 9 pages, 2017.

[26] C.-I. Fan and Y.-H. Lin, "Provably secure remote truly three-factor authentication scheme with privacy protection on biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 933–945, 2009.

[27] S. Yoon, J. Feng, and A. K. Jain, "Altered fingerprints: analysis and detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 451–464, 2012.

[28] A. K. Jain and S. Yoon, "Automatic detection of altered fingerprints," *Computer*, vol. 45, no. 1, pp. 79–82, 2012.

[29] S. Ghouzali, M. Lafkih, W. Abdul, M. Mikram, M. El Haziti, and D. Aboutajdine, "Trace attack against biometric mobile applications," *Mobile Information Systems*, vol. 2016, Article ID 2065948, 15 pages, 2016.

[30] L. Thalheim, J. Krissler, and P.-M. Ziegler, "Body Check: Biometrics Defeated," 2002, http://www.pcmag.com/article2/0,2817,13919,00.asp.

[31] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, LivDet 2009—Fingerprint Liveness Detection Competition 2009, 2009.

[32] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, "LivDet 2011—fingerprint liveness detection competition 2011," in *Proceedings of the 2012 5th IAPR International Conference on Biometrics (ICB '12)*, pp. 208–215, April 2012.

[33] L. Ghiani, D. Yambay, V. Mura et al., "LivDet 2013 fingerprint liveness detection competition 2013," in *Proceedings of the 6th IAPR International Conference on Biometrics (ICB '13)*, pp. 1–6, Madrid, Spain, June 2013.

[34] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers, "LivDet 2015 fingerprint liveness detection competition 2015," in *Proceedings of the 7th IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS '15)*, September 2015.

[35] K. Cao and A. K. Jain, "Hacking Mobile Phones Using 2D Printed Fingerprints," 2016.

[36] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Computing Surveys*, vol. 47, no. 2, article 28, 2014.

[37] D. Zhang, Z. Guo, and Y. Gong, *Multispectral Biometrics: Systems and Applications*, Springer, 2015.

[38] NIST, "NIST Fingerprint Image Quality," *Biometric Quality*, https://www.nist.gov/programs-projects/biometric-quality-homepage.

[39] ISO/IECJTC 1/SC37, "DRAFT INTERNATIONAL STANDARD ISO/IEC DIS 30107-3 Information technology—Biometric presentation attack detection," 2017.

[40] NIST, NIST Biometric Image Software, 2015, http://www.nist.gov/itl/iad/ig/nbis.cfm.