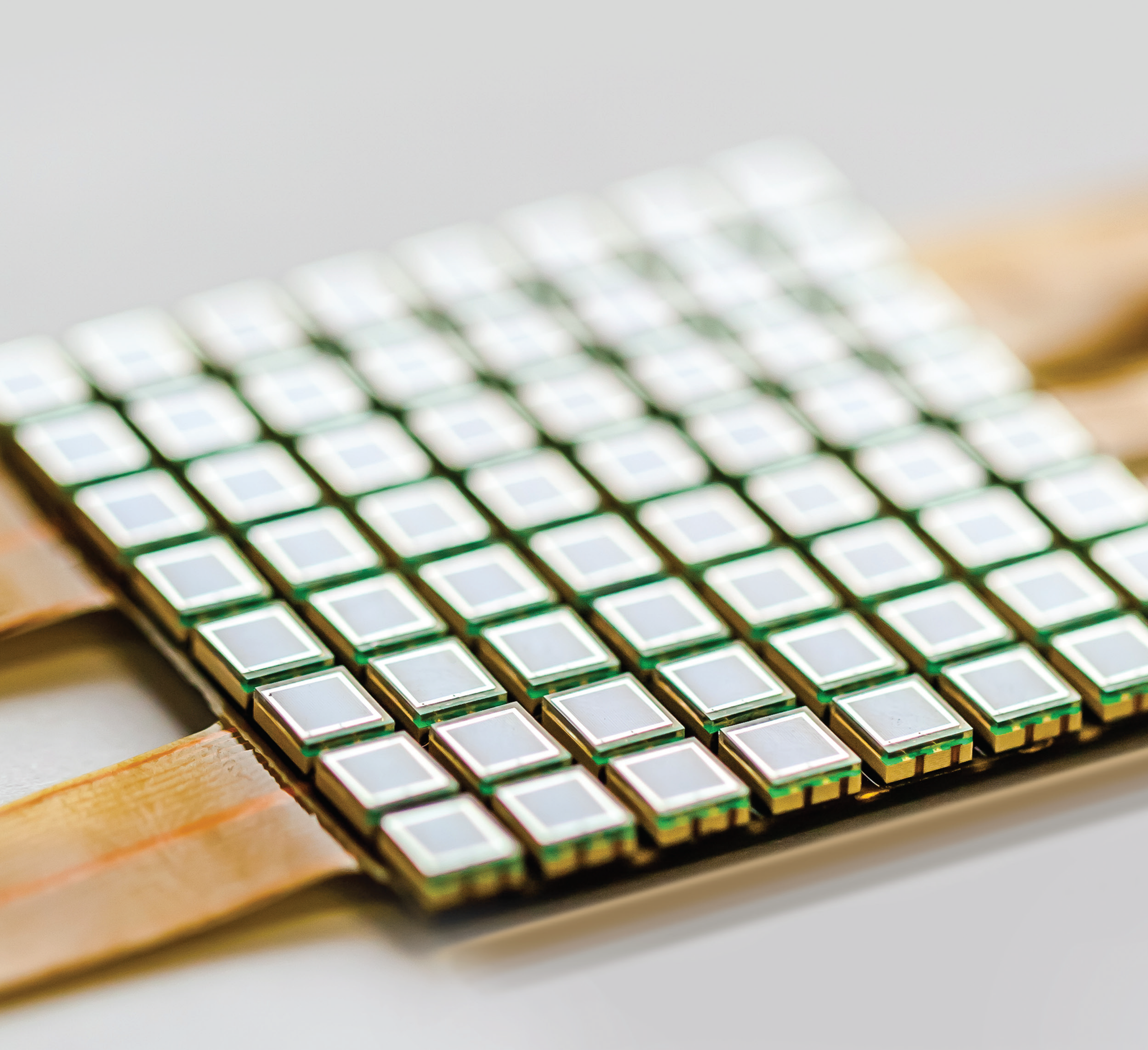
Journal of Sensors

Achieving Green Internet of Things

Lead Guest Editor: Tsu-Yang Wu
Guest Editors: Saru Kumari and Muhammad Khurram Khan



Achieving Green Internet of Things

Journal of Sensors

Achieving Green Internet of Things

Lead Guest Editor: Tsu-Yang Wu
Guest Editors: Saru Kumari and Muhammad Khurram Khan

Chief Editor

Associate Editors

Academic Editors

Ehsan Namaziandost, Iran
Heinz C. Neitzert, Italy
Sing Kiong Nguang, New Zealand
Calogero M. Oddo, Italy
Tinghui Ouyang, Japan
SANDEEP KUMAR PALANISWAMY, India
Alberto J. Palma, Spain
Davide Palumbo, Italy
Abinash Panda, India
Roberto Paolesse, Italy
Akhilesh Pathak, Thailand
Giovanni Pau, Italy
Giorgio Pennazza, Italy
Michele Penza, Italy
Sivakumar Poruran, India
Stelios Potirakis, Greece
Biswajeet Pradhan, Malaysia
Giuseppe Quero, Italy
Linesh Raja, India
Maheswar Rajagopal, India
Valerie Renaudin, France
Armando Ricciardi, Italy
Christos Riziotis, Greece
Ruthber Rodriguez Serrezuela, Colombia
Maria Luz Rodriguez-Mendez, Spain
Jerome Rossignol, France
Maheswaran S, India
Ylias Sabri, Australia
Sourabh Sahu, India
José P. Santos, Spain
Sina Sareh, United Kingdom
Isabel Sayago, Spain
Andreas Schütze, Germany
Praveen K. Sekhar, USA
Sandra Sendra, Spain
Sandeep Sharma , India
Sunil Kumar Singh Singh, India
Yadvendra Singh, USA
Afaque Manzoor Soomro, Pakistan
Vincenzo Spagnolo, Italy
Kathiravan Srinivasan, India
Sachin K. Srivastava, India
Stefano Stassi, Italy
Danfeng Sun, China
Ashok Sundramoorthy, India
Salvatore Surdo, Italy
Roshan Thotagamuge, Sri Lanka
Guiyun Tian, United Kingdom
Sri Ramulu Torati, USA
Abdellah Touhafi, Belgium
Hoang Vinh Tran, Vietnam
Aitor Urrutia, Spain
Hana Vaisocherova - Lisalova, Czech Republic
Everardo Vargas-Rodriguez, Mexico
Xavier Vilanova, Spain
Stanislav Vítek, Czech Republic
Luca Vollero, Italy
Tomasz Wandowski, Poland
Bohui Wang, China
Qihao Weng, USA
Penghai Wu, China
Qiang Wu, United Kingdom
Yuedong Xie, China
Chen Yang, China
Jiachen Yang, China
Nitesh Yelve, India
Aijun Yin, China
Chouki Zerrouki, France

Contents

Hindawi
Journal of Sensors
Volume 2022, Article ID 1309402, 16 pages
https://doi.org/10.1155/2022/1309402

Research Article

SF-LAP: Secure M2M Communication in IIoT with a Single-Factor Lightweight Authentication Protocol

Khurram Shahzad,[1] **Masoom Alam**,[1] **Nadeem Javaid**,[1] **Abdul Waheed**,[2,3] **Shehzad Ashraf Chaudhry**,[4,5] **Nafees Mansoor**,[6] **and Mahdi Zareei**[7]

[1]*Department of Computer Science, COMSATS University Islamabad, 44000, Pakistan*
[2]*Department of Computer Science, Women University Swabi, Swabi 23430, Pakistan*
[3]*School of Electrical and Computer Engineering, Seoul National University, Seoul 08826, Republic of Korea*
[4]*Department of Computer Science and Information Technology, College of Engineering, Abu Dhabi University, Abu Dhabi, UAE*
[5]*Department of Computer Engineering, Faculty of Engineering and Architecture, Nisantasi University, Istanbul 34398, Turkey*
[6]*Department of Computer Science and Engineering, University of Liberal Arts Bangladesh (ULAB), Dhaka, Bangladesh*
[7]*Tecnologico de Monterrey, School of Engineering and Sciences, Mexico*

Correspondence should be addressed to Nafees Mansoor; nafees@nafees.info

Received 20 June 2022; Revised 21 August 2022; Accepted 23 August 2022; Published 10 November 2022

Academic Editor: Akhilesh Pathak

Machine-to-machine communication allows smart devices like sensors, actuators, networks, gateways, and other controllers to communicate with one another. The industrial Internet of things (IIoT) has become a vital component. Many industrial devices are connected to perform a task automatically in machine-to-machine communication, but they are not properly secured, allowing an adversary to compromise them against a variety of attacks due to communication system vulnerabilities. Recently, a secure lightweight authentication protocol (SLAP) was proposed by Panda et al. They asserted that every known attack that could happen in the IIoT is deterred by their suggested protocol. In this study, we prove that the SLAP protocol is vulnerable to desynchronization, impersonation, replay, and eavesdropping attacks. To prevent these attacks and enhance that protocol, we need to implement a secure authentication mechanism that ensures the security of communication. This paper proposed a secure M2M Communication in IIoT with a single-factor lightweight authentication protocol (SF-LAP). Single-factor authentication is a simple and secure way to communicate. It uses less power and communication overhead while providing a secure mechanism for conversation. In the machine-to-machine (M2M) scenario, the proposed protocol uses an exclusive-OR operation and a hashing function to ensure secure communication between the sensor and the controller. The proposed mechanism uses a secure preshared key and timestamp technique to protect and safeguard this connection against desynchronization attacks and eavesdropping attacks. We used Burrows Abadi Needham (BAN) Gong, Needham, and Yahalom (GNY) logic, and the automated validation of Internet security protocols applications (AVISPA) tool for formal verification and perform a security analysis as an informal verification to make sure the suggested protocol is secure. Analysis that shows the SF-LAP consumes the least computing and communication overhead and is more secure because it prevents desynchronization and eavesdropping attacks to all of the known attacks that are modification attacks, tracing attacks, impersonation, man-in-the-middle, and replay attacks.

1. Introduction

Direct communication between devices, whether wired or wireless, is referred to as machine-to-machine and is used in industrial internet of things to increase productivity and efficiency while using less energy. Different devices such as actuators, controllers, and sensors are networked together with other devices in IIoTs. In order to increase productivity and efficiency, interconnect gathers data, analyzes it, and recommends effective measures. The main structure of our paper is depicted in Figure 1. You should be familiar with the terms machine-to-machine communication, IIoT, and lightweight authentication protocol before reading our paper.

1.1. M2M Communication. With the proliferation of wireless devices, there is a growing demand for cellular network based M2M communication. The main feature of M2M is that it allows two or more devices to communicate without the need for human interaction, as well as perform and supervise certain functions automatically. The main goal of M2M in smart devices is to ensure secure communication, automate device interaction and communication, access high-speed internet, measure the costs of different devices, minimize computation and communication costs, ensure a cheap and easy way to connect, make lightweight and efficient software for better speed, address locomotive devices and treat on location, and make lightweight and efficient software for better speed. In terms of efficiency and security, M2M is more typically linked to medium-access smart home apps, mobile M2M standard services and platforms, and LTE [1]. Figure 2 depicts many features of M2M communication.

Wireless technologies such as WLAN, RFID, near field communication, Zigbee, Bluetooth, long-term evolution (LTE), and others are widely used in M2M communication. They differ in infrastructure, ranges, throughput, sizes, and efficiency [2]. The SD-M2M framework is proposed to minimize costs and enhance end-to-end quality for renewable energy management in the house, building, industrial, grid and microgrid, and field [3]. M2M introduces LTE-A to improve the physical layer. The cellular network for M2M communication improves the short access latency and high throughput, as well as minimizing the signaling overhead [4]. M2M communication is utilized in the e-health industry to improve patient care, data protection, tamperproof electronic report delivery, dynamic doctor assignments, lower costs, and increase system performance [5]. M2M communication is used in smart home appliances to monitor and control smart home devices. These appliances are connected to smart phones and can be operated and controlled remotely, making life easier [6].

1.2. Industrial IoT. Since many smart devices are connected, one of the biggest challenges in cybersecurity. For the purpose of preventing or reducing cybersecurity events and corporate data breaches, employees must be informed about these risks and, as a result, build the resilience of their organizations to cyberattacks [7]. The blockchain system uses Merkle trees to store data; however ,this article substitutes the incremental aggregator subvector commitment for the Merkle tree to reduce the size of the proof and the quantity of communication needed. In addition to lowering the storage pressure on nodes, the proof size has been decreased from its original data scheme to 15% [8]. The results show that the encryption and verification procedures are sped up by around 70 and 90%, respectively, using this technology. It might be utilized in cloud-assisted healthcare IIoT to offer comprehensive answers to the aforementioned issues. Additionally, this approach reduces communication expenses by more than 80% during the verification stage [9].

To protect IIoT infrastructure against deadly and complex multivariant botnet attacks, the study suggests a hybrid intelligent deep learning-enabled method. The results demonstrate speed efficiency, and the suggested procedures outperform in reliably recognizing multivariant sophisticated bot assaults [10]. The IIoT's intelligent sensors are used to collect the physical attributes of objects dispersed across a large area.The heuristic technique is recommended to assist the backbone node in balancing the incoming traffic for scheduling. Aggressive scheduling media access control is the name of the proposed effort. Other nodes will actively arrange the proper time slot if the owner node does not have data to broadcast at that moment. According to calculations and comparisons with the most recent proposals, this technique outperforms time division multiple access in terms of packet loss rate, energy consumption, packet delivery rate, and time required for various numbers of sensor nodes [11].

This paper [12] examines how to use new IIoT technologies in a cloud manufacturing system to address this problem. It provides a general system architecture for a plug and play service-oriented IIoT gateway solution for a cloud manufacturing system helped by the IIoT. In order to quickly store and query data in a cloud time series database while capturing precisely the correct quantity of data for field-level manufacturing equipment, service-oriented data schemas are developed. According to research, using purposefully designed service-oriented data schemas that use plug and play IIoT technologies to gather the essential data for high-level cloud manufacturing decision-making is an efficient way to connect field-level manufacturing equipment to a cloud manufacturing platform.

Due to the widespread usage of IIoT, there is an increasing demand for sensing technology to collect data and information. Industrial wireless sensor networks with industrial information perception capabilities have developed to fill this demand. The proposed multiobjective chaotic elite adaptive ant colony optimization has obvious advantages. The population will be initialized, increasing population diversity and the algorithm's ability to deviate from the local ideal. By constantly adjusting the algorithm trend, the adaptive optimization technique significantly accelerates algorithm convergence. In various scale scenarios, the suggested algorithm's performance is assessed. The simulation results demonstrate that this technique may significantly enhance routing when compared to other cutting-edge QoS routing systems. When compared to other cutting-edge QoS routing systems, the simulation results show that this technique can successfully reduce network energy consumption, improve routing security, and fulfill multi-QoS restrictions for the end-to-end latency and reliable service [13]. Figure 3 summarizes some of the IIoT features.

1.3. Lightweight Authentication Protocols. Authentication protocols are specifically designed to transfer the data between two or more users through a proper and secure communication channel. These channels sometimes use a

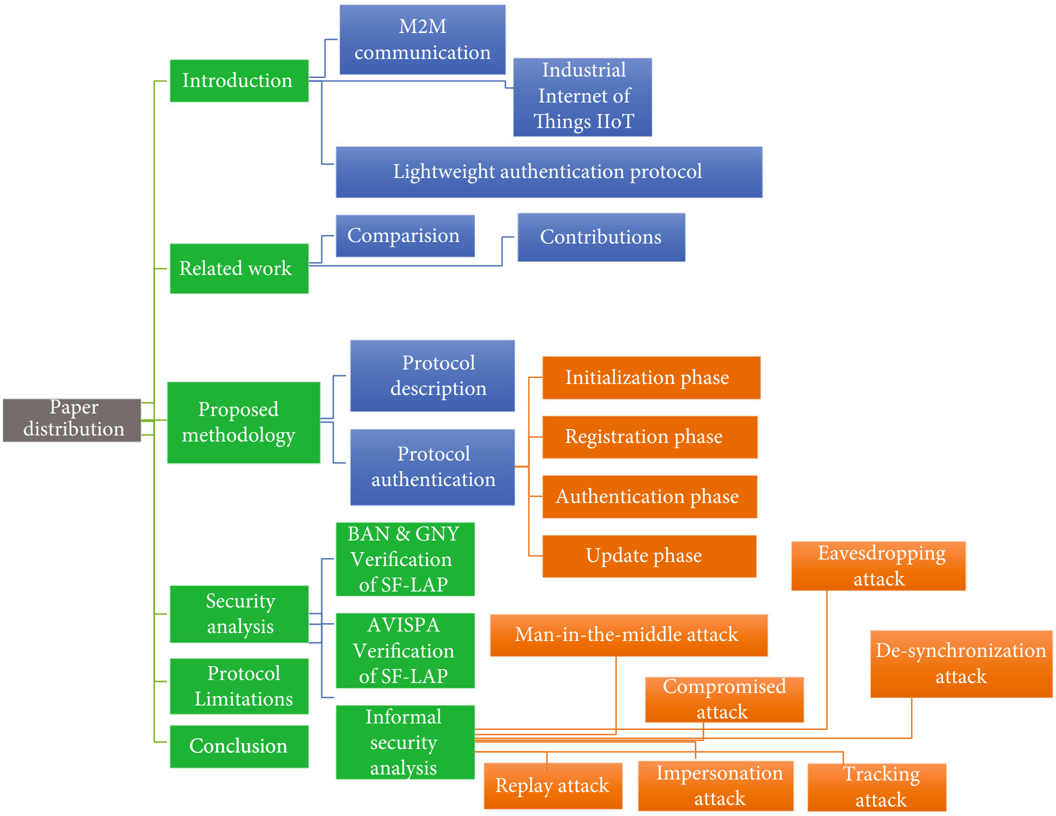


Figure 1: SF-LAP paper structure.

different cryptography algorithm for different devices and different security purposes. For instance, to achieve authority, integrity, and authenticity rely on different layers to make a system in secure manners. A bulk of authentication techniques are widely used, but we target M2M communication and propose a new SF-LAP secure protocol.

A new notion for an Internet of things micropayment framework in terms of a wearable device is presented in the paper [14]. This payment mechanism encrypts and decrypts communication messages between numerous organizations using an elliptic curve integrated encryption technology. Customers may buy things using a wearable gadget and communicate sensitive payment information via a mobile app, according to protocol. Customers, banks, and merchants are all securely connected through the program. The implementation of wearable device micropayments makes use of a secure end-to-end solution. To accomplish security features, NFC combines lightweight cryptography and elliptic curve cryptography's device pairing technology. The proposed protocol uses BAN logic to conduct mutual authentication among the many parties engaged in the protocol, enabling real-time transactions between banks and IoT devices.

The protocol's viability and safety are demonstrated using formal BAN logic. Based on the security analysis and testing findings, the proposed authentication protocol is more suitable for scaling to enormous authentication. It may reach higher security standards at a reduced cost. It also protects against desynchronization and denial-of-service assaults. Furthermore, this approach does not handle less tag storage and time. The paper [15] proposes an anonymous authentication method for resource-constrained IoT (RC-IoT) devices on page no. 1822, based on elliptic curve cryptography and signcryption. This protocol lowers the cost of communication and computation. This work uses a random oracle model to formalize theoretical security and simulated studies to show that this protocol effectively minimizes resource use.

Figure 2: Different aspects of M2M Communication.

The study [16] proposes lightweight 3FA for WSN for healthcare remote user IoT application, based on hash, and XOR includes features of biometrics, smart devices and user password, session key, and mutual authentication. Protocol

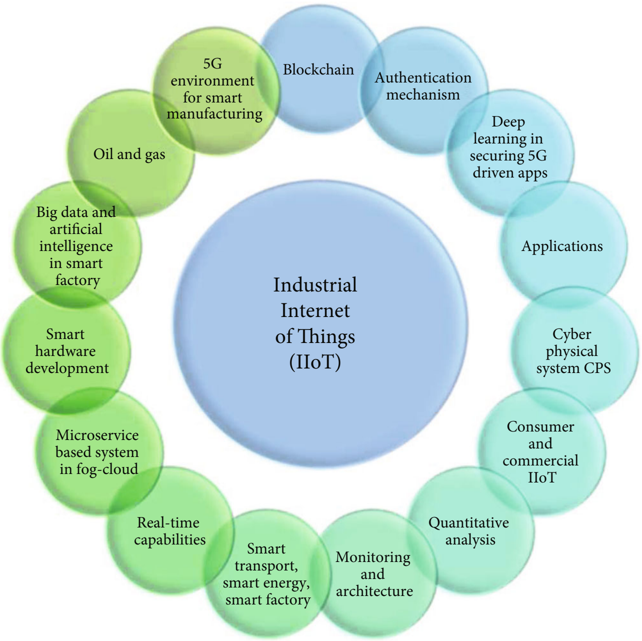

Industrial Internet of Things (IIoT)
Blockchain
Authentication mechanism
Deep learning in securing 5G driven apps
Applications
Cyber physical system CPS
Consumer and commercial IIoT
Quantitative analysis
Monitoring and architecture
Smart transport, smart energy, smart factory
Real-time capabilities
Microservice based system in fog-cloud
Smart hardware development
Big data and artificial intelligence in smart factory
Oil and gas
5G environment for smart manufacturing

Figure 3: Different features of IIoT.

is based on BAN logic and formal verified on AVISPA tool, reduces the communication and computation overhead. In [17], study proposes secure authentication mechanism provides authentication, integration, confidentiality, and access control to monitor healthcare devices and to meet complex situation, remote surgery, real time applications, mental and physical health, blood pressure ECG, etc. This scheme controls computation cost, IoT verification method, authentication mechanism in application layer, network, and perception. Also, IoT devices verification in real time application.

The fundamental challenges in the [18] M2M study are mutual authentication, authentication key agreement, and network. To overcome these challenges, several initiatives are being implemented, such as LTE and 5G for better networking speeds and protocols. The study finds that group-based secure lightweight authentication and key agreement reduce communication and computational cost. In addition, 6G is being researched in a number of industries, including e-health, military, e-education, and e-commerce, all of which depend on IoT and blockchain technologies and billions of devices [19].

ISAG is proposed in this [20] study to ensure authentication and key agreement in LTE networks to avoid MITM and DOS attacks. The lightweight authentication protocol, according to the [21] VANET study, provides a secure way for emergency vehicles to verify illegal legitimacy while also protecting them from reputation, device theft, and impersonation attacks. This study [22] proposed a secure authentication protocol for LTE networks. REPS-AKA3 improves network performance by reducing bandwidth consumption, storage, and communication overhead, as well as using the AVISPA tool for formal verification. Lightweight authentication mechanism post quantum FLAT is proposed in [23] for service and certificate providers, as well as RC-IoT devices in M2M communication devices that can withstand post quantum changes.

The rapid adoption of wireless wearable Internet of things devices in textiles where data is transmitted increases the possibility of an eavesdropping attack. The mutual authentication protocol between IoT devices and gateway is strengthened by this [24] document, which supports cutting-edge encryption methods. The study provides the necessary conclusions and also verifies the authenticity mechanism's security in comparison to other mutual authentication protocols using the AVISPA and Scyther tools, ensuring resilience against eavesdropping attacks. The blockchain technology is used in M2M communication IoT devices to avoid a centralized system. The [25] study uses blockchain technology, as well as the AES, SHA-256 algorithm, and HMAC, to improve the processing time and speed of hardware chips in smart home devices.

Light-AHAKA, a novel [26] lightweight protocol with a key agreement to encrypts sensitive data, is proposed as a secure mechanism for cloud RC-IoT systems. The study proposes [27] key agreement, text message, and registration phase for the authentication process between devices and server in order to reduce RC-IoT device, communication, and computation overhead and improve battery life. For M2M communication, [28] proposes a key exchange and mutual authentication protocol. It is also safe, light, efficient, and productive. MITM and replay attacks over the network are not possible with this protocol. AVISPA and Scyther provided formal verification.

This [29] study proposes a mutual and transitive authentication mechanism for intermediate devices and gateways. The session key is verified using BAN logic, and the resilience against critical attacks is verified using the Scyther tool. The simulation results demonstrate performance efficiency while also reducing handshake time, memory, and energy consumption. This proposed technique is small and can be used with RC-IoT devices. The study of telerobotic surgery in M2M communication devices introduces a new efficient device scheme, as well as its detection and classification against malicious nodes. This system is more efficient, secure, and capable of detecting and classifying rogue devices. This protocol is resistant to replay attacks, DDOS attacks, and eavesdropping. Its results are superior to those of the ECFU and LEACH protocols [30].

As described in the paper [31], the ultra-lightweight RFID systems security authentication protocol uses bit-crossing XOR restructuring operations to successfully deflect denial of service threats, replay, resist forgery, and desynchronization while authenticating the data center. The research [32] compares and contrasts two mutual authentication protocols: hash and Rabin public key. The reader, server, and tag are all desynchronized as part of the impersonation attack. This method secures the connection between the tag and the reader. The verification process uses the Scyther and random oracle models to ensure that this scheme is secure and efficient for RFID systems.

The remaining portion of the paper is structured as section 2 examines related research, compares the proposed protocol to all previous proposals, and highlights this paper's

Table 1: Comparison of SF-LAP with other protocols.

Resilience against attacks	[33]	[34]	[35]	[36]	[37]	[38]	SF-LAP
Replay attack resilience	✓	✓	✓	✓	✓	✗	✓
MITM attack resilience	✓	✓	✗	✓	✓	✓	✓
Impersonation attack resilience	✓	✓	✓	✓	✓	✗	✓
Modification attack resilience	✗	✗	✓	✓	✓	✓	✓
Compromise attack resilience	✗	✗	✗	✓	✗	✓	✓
Tracing attack resilience	✗	✗	✗	✓	✗	✓	✓
Eavesdropping attack resilience	✗	✓	✓	✗	✗	✗	✓
Session-key breach	✗	✓	✓	✓	✓	✓	✓
Identity confidentiality resilience	✗	✗	✓	✓	✓	✓	✓
Formal verification by AVISPA	✓	✗	✓	✓	✗	✓	✓
Desynchronization attack resilience	✗	✗	✗	✗	✗	✗	✓

contribution. Section 3 proposes methodology and provides a description and steps for authentication. Section 4 uses the BAN and GNY logic and the AVISPA tool, respectively, to conduct formal and informal security analyses. The final section 5 of our paper includes the references and the conclusion.

2. Related Work

This study proposes a three-factor mutual authentication mechanism based on elliptic curve cryptography enabling patient privacy protection in the Telecare Healthcare Information System. The communication expenses are lower using this protocol. Uses the real-or-random model, BAN logic, and the AVISPA to perform formal security evaluations. This protocol outperforms the Ryu et al. approach, which exposes itself to privileged suspicious assaults, patient anonymity, and password update errors. This protocol can protect against many security threats, such as privileged suspicious assaults, stolen mobile phones, and insider attacks. Using the ROR model and the BAN logic demonstrated, believing their approach could assure mutual session and authentication key security. Compared to similar current protocols, this protocol is more secure and has reduced communication costs [33].

SAKE is a symmetric authentication key exchange protocol is proposed in this [34] article for IIoT. It relies on hash function operations, concatenation, and XOR to ensure message integrity, key exchange, and lightweight authentication. Regarding security and performance on page no. 8, analyzes the SAKE protocol to seven modern and IoT-related authentication schemes. According to the comparison results, the SAKE protocol, page no. 9, uses the least number of computational resources and has the 3rd communication overhead of both eight AKE protocols providing twelve security measures. To show that the proposed protocol offers complete forward secrecy and secure session key agreement, it is compared to AKE-related protocols. The [35] data confidentiality, message privacy, and authenticity are among the security services provided by the suggested paradigm. According to the protocol's security verification and performance findings, the proposed system is suitable for use in real-world IoT applications. However, desynchronization assaults are not protected by this article.

The author [36] proposes an industrial RC-IoT authentication protocol. For M2M communication in the IIoT, the LAKD authentication protocol has been presented. This project aims to achieve a low computational cost that is appropriate for IIoT devices with limited resources. To do this, the proposal employs lightweight operations such as addition, subtraction, XOR, and hash function. Although the AVISPA tool and BAN logic confirmed mutual authentication and resistant to MITM and replay attacks, this suggested technique does not solve many of the previously described threats, such as forgery, session key disclosure, impersonation, and desynchronization attacks.

For storage overhead and M2M communication, the document [37] presents a based authentication system that relies on XOR operations and hashes for session key agreement, mutual authentication, device identity confidentiality, and protection against impersonation, modification, and man-in-the-middle and replay attacks. This article, however, does not provide protection against desynchronization assaults. As a result, the system we propose is resistant to desynchronization assaults. [38] SLAP is lightweight authentication and secure protocol that uses two techniques, hash and XOR operations. Modification attacks, tracing, man-in-the-middle, impersonation, and replay attacks are all safe with the proposed protocol. The proposed SLAP's correctness is checked using BAN logic. The protocol is safe, according to the AVISPA tool. SLAP protocol is suitable for various attacks, but it does not protect against desynchronization attacks. The proposed authentication framework covers forgery attacks, session key disclosure, impersonation, MITM, and replay attacks.

2.1. Comparison. It should be noted that the protocols in question are vulnerable to some well-known attacks. The proposed protocols do not provide sufficient security for session keys. As a result, an adversary can quickly obtain and exploit the session key. When we compare the security features of the SF-LAP protocol to those of other M2M communication protocols, we can see how well it defends against various attacks. [33] TMIS proposed 3FA based on

Elliptic curve cryptography, securing joint sessions and authentication keys but not against all other attacks. [34] Concatenation, XOR, and hash functions are used in the SAKE protocol. Of the eight AKE protocols, it consumes the fewest computing resources and has the third-lowest communication overhead, with twelve security features but no protection against further attacks. This [35] protocol protects against forgery, session key, impersonation, MITM, and replay attacks but not against desynchronization attacks. [36] LAKD is a RC-IoT authentication protocol for M2M communication that uses four different techniques to achieve low computational cost for IIoT devices, including addition, subtraction, XOR, and hash. Many of the previously discussed attacks, such as forgery, session key disclosure, impersonation, and desynchronization attacks, are not addressed by this proposed method. So this protocol is vulnerable.

This [37] protocol achieved session key agreement, device identity confidentiality, mutual authentication, protection against impersonation, modification, replay, and MITM attack for storage overhead and M2M communication. Still, it was unable to protect against desynchronization attacks. As a result, our proposed scheme, SF-LAP, is providing protection against all of these attacks as well as desynchronization attacks. Table 1 shows a comparison of all of these procedures to SF-LAP. [38] SLAP is lightweight authentication uses two techniques, hash and XOR operation. The author claimed that his protocol is safe and provide resilience against modification attacks, tracing, man-in-the-middle, impersonation, and replay attacks, but we have some proof that the proposed protocol is vulnerable against eavesdropping, replay, impersonation, and desynchronization attack. The proposed protocol is checked by BAN logic with the extension of GNY logic includes more rules for verification purpose.

2.1.1. Eavesdropping Attack on SLAP [38]. An attacker can identify and track the messages of a controller among the communications exchanged between a controller and a sensor. To track the victim, the attacker must eavesdrop on two of their sessions. The attacker successfully eavesdrops on the session between the controller and the sensor to get $b_i = h($ $r\|PSK), D_1 = b_i \oplus R_1, D_2 = h(R_1) \oplus \mathrm{ID}_s,$ and $D_{3'} = h(\mathrm{ID}_s \|$ $R_1\|r\|TS_0)$. Since the value of r is a simple random number that repeats and can be known by the attacker by guessing it, an adversary may simply learn the value of b_i. By eavesdropping in on the starting session, the attacker was able to successfully retrieve the value of D_1 and the controller ID is now available for use in the next authentication session.

Now that the attacker is aware of the message the controller will send the sensor as ID_r, they may watch the controller in the subsequent session. The attacker has a variety of options for completing this level. In this case, the attacker acts as a sensor and sends a message of request to the controller. With knowledge of the controller's responses to this message, the attacker obtains the messages that were sent by it and compares the values with D_1. When a value matches D_1, the attacker knows that the message is about the controller, and the controller is monitored. The adversary is aware of the value of M_1 sent from the sensor to the controller. The attacker can find and follow a specific controller with this kind of attack. The attacker employs this strategy because it is aware of the value of ID_r and has learned this from previously eavesdropping in on a session.

2.1.2. Replay and Impersonation Attack on SLAP [38]. In this attack, the adversary sends the controller the messages it got from the previous session to mimic a real sensor. The adversary eavesdrops on the session between controller and sensor. An opponent who had previously known the values of D_1 and D_2 now gets after the session as a consequence of preventing $D_{3'} == D_3, r == h(ID\|rr)$ and $D_{3'} == D_{3'}$. The attacker then pretends to be a sensor and sends the controller a request message. $D_{3'}$ is sent to the adversary by the controller. After receiving $D_{3'}$, the adversary provides M_1, which was obtained from the previous session and forwarded to the controller.

The controller receives M_1, which verifies that the opponent is the actual sensor. The opponent then receives M_2 from the controller. The opponent makes sure the controller has validated the validity of the controller after getting M_2. The adversary has therefore effectively completed its offensive task. Finally, by leveraging the stolen messages from the previous session, the attacker may have fooled the controller into thinking it was a legitimate sensor. The controller in this message decided to utilize random numbers rather than the Nonce technique because of the circumstances that led to this approach. The controller's identification is accepted once the sensor's identity has been verified.

2.1.3. Desynchronization Attack on SLAP [38]. SLAP was unable to stop the desynchronization attack because the values of b_i are stored as $b_i = h(r \| PSK)$. The b_i joins the opponent and can disrupt the synchronization between the controller and sensor when the b_i cannot be effectively received by the controller. The attacker passes messages from the controller to the sensor. In other words, the attacker uses this method to perform a man-in-the-middle attack by intercepting a request message from the sensor and sending it to the controller.

The attacker then sends the controller the M_1 message that was sent by the sensor. The controller, rather than the sensor, intercepts the communication and resends it to the adversary. The $D_{3'}$ message is compared to the D_3 and r that the adversary was already known by this controller as a result. The attacker executes the attack and is known that utilize M_2 to deliver the values of D_4, D_5, D_6, and D_7 to the sensor through $b_{i'} = h(r' \| PSK)$ makes it to get for him to do so. The attack is effective since the attacker may now compute the shared key because he knows the value of $D_7 == D_{7'}$.

2.1.4. SF-LAP VS SLAP [38] Protocol. The two key distinctions between these two are the involvement of the server in SF-LAP as compared to S and C, and the server clock timestamp function which is an added feature of SF-LAP. These modifications in SF-LAP mitigate the preceding attacks that were addressed in SLAP as giving resilience and providing resistance against eavesdropping, replay, impersonation, and desynchronization attacks. The following are some further variations between these two protocols are shown in Table 2.

2.1.5. SLAP [38] Limitations

(i) The SLAP provides a simple random number that may be reused and can cause of pseudo-random number generator attack. While the threshold time safety factor is missing when not synchronized properly

(ii) Both the devices are synchronized separately can cause of desynchronization attack. The comparison between SLAP and SF-LAP is shown in Figure 4

(iii) Attack resilience is depicted on the left side. SLAP is represented with orange bars and SF-LAP with blue bars. The absence of bars indicates that the paper lacks resistance to the attack. For instance, the fact that a desynchronization attack represents a blue bar but not an orange one indicates that SF-LAP can provide resilience against that attack while SLAP cannot

2.2. Contribution. To enhance SLAP, we add two to three new techniques called SF-LAP, compare them, and then explain why SF-LAP is more secure than SLAP. SF-LAP identifies some limitations of using the SLAP protocol as a base paper. SF-LAP makes it better by including some techniques like adding a nonce technique to a secure authentication channel and including a server and server clock between these two devices. This server clock's job is to simultaneously provide a timestamp for both devices. This timestamp has the advantage of being resistant to eavesdropping and desynchronization attacks and includes the safety chain threshold time for the authentication process. To obtain the appropriate results for SF-LAP in comparison to SLAP, we analyze it using the BAN with an extension of GNY logic [39] and run it on AVISPA and prove the SF-LAP protocol. SF-LAP provides resilience against eavesdropping, replay, impersonation, and desynchronization attack while taking into account all previous attacks. SF-LAP protocol is more secure than previous lightweight authentication protocol.

3. Proposed Methodology

The previously proposed authentication protocol was good and secure in terms of modification, tracing, man-in-the-middle, impersonation, and replay attacks, but it was unable to offer resilience against desynchronization and eavesdropping attacks. We proposed a new sensor-to-controller authentication technique SF-LAP that enables secure communication between the sensor and the controller while also offering resistance against desynchronization and eavesdropping threats while accounting for all prior assaults. It has a different process, but we must declare and highlight the similarities in SF-LAP, which we derived from SLAP rather than the timestamp technique, server involvement, resilience against desynchronization attack, and eavesdropping attack.

Moreover, to prevent a pseudo-random number generation attack from the repetition of random numbers, we used the nonce technique in this paper inspired by the paper of Rostampour et al. [39]. SF-LAP provides pseudocode for the authentication process while BAN with extension of GNY logic is used for formal verification and security analysis for informal verification. It was necessary to use the same criteria of overcomes and to provide a more secure protocol than the previous one. These logics are believes, possesses, once said, jurisdiction, fresh, session, and encryption used to prove the authentication protocol. The rules of these logics are the message-meaning rule, nonce verification rule, possession rule, freshness rule, recognizability rule, jurisdiction rule, and message interpretation rule. Table 3 shows the notation we use for this authentication mechanism.

3.1. Protocol Description. In SF-LAP protocol PSK has already been shared between S and A as shown in Figure 5. S, C, and A make up SF-LAP protocol. S sends information to C with ID_s and timestamp TS, which verifies the message and sends it back to S. By using the server clock to generate timestamps, the server ensures that the S and C communicate with a single timestamp TS. Desynchronization and eavesdropping attacks are prevented by the server timestamp for both the S and C.

3.2. Protocol Authentication Process. Simple timestamp can be use but due to obtain several timestamps by an adversary. Adversary can guess the pattern which can be reason for the several attacks like guessing and replay attack. Nonce is an arbitrary number, random number, or pseudo-random number that used for once where the values are dynamically changed. We use nonce along with the timestamp, the purpose of N along with timestamp is to secure the authentication process. The authentication mechanism is depicted in Figure 6, shows how the S and C exchange data and share the secret message.

3.2.1. Initialization Phase. Authentication server (A) believes that sensor (S), and controller (C) are already known PSK then, A always believe that session key is shared between S and C. S requests to C for communication and shares the data, then C request to A for TS. A generates TS with its clock that shares the TS_1 to both S and C, simultaneously. To make theoretically understandable, we says $\mathrm{TS}_1 = \mathrm{TS}_{S1}$ for S and TS_{C1} for C to avoid further confusion because one timestamp is shared on both sides, but practically its T S_1 with same hash on both sides. So now we can say that, A generates TS_{S1} and TS_{C1} with its clock and shares to both S and C, simultaneously.

3.2.2. Registration Phase. S generates N_s and hash the value of $a = (\mathrm{ID}_s \parallel \mathrm{PSK})$ and XOR with ID_s that is $D_1 = h(a) \oplus \mathrm{I}$ D_s and concatenate the value of TS_1 with $b = (N_a \parallel \mathrm{PSK})$ and XOR with N_s. S encrypts these values and sends message $M_1 = \{D_1, D_2, N_a, \mathrm{TS}_{S1}\}$ to C. Then, C selects the timestamp TS_{S1} and compares it with TS_{C1} that C already taken from

Table 2: SF-LAP VS. SLAP Protocol.

SF-LAP	SLAP
It provides a clear use-case in 3 steps	Does not provide the use-case diagram
Provides security against eavesdropping attacks	Does not provide resilience against eavesdropping attacks
Provides security against replay attacks	Does not resist replay attack
Provides security against impersonation attacks	Does not resist impersonation attack
Provides security against desynchronization attacks	Does not resist desynchronization attack
Timestamps both the S and C simultaneously	Does not provide this phenomenon
Communicates between S and C via server	There is no medium other than the sensor and controller
Provides the facility of timestamp through the server clock	Not provide the facility of this type of a timestamp
Provides pseudocode of our procedure	Does not provide any pseudocode
A special nonce function is used which provides the random number only for one time and then it never uses the same random number again in its communication	There is a simple random number is used in communication that is vulnerable for replay attack

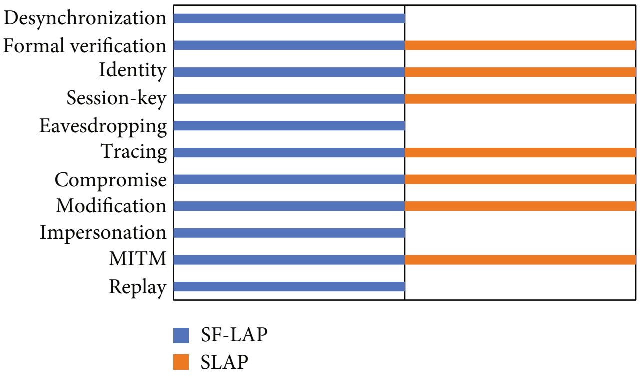


Figure 4: Comparison between SF-LAP and SLAP.

the A and verifies for the same values of both TS_{S1} and TS_{C1}. If false, then terminate the whole authentication process but if true, then computes the values of $ID_s = D_1 \oplus h(a)$ and $N_s = D_2 \oplus h(b \| TS_{S1})$. Then calculate the value of D_3 while take TS_2 from A and concatenate with N_s such as $D_3 = h(N_s \| TS_{C2}) \oplus N_c$, while C share its ID as ID_c, XORed with both the nonce values such that $D_4 = IDc \oplus h(b \| N_s \| N_c)$. Concatenate ID_c, a and value of nonce that is $D_5 = h(IDc \| a \|N_s \| N_c)$ and returns the message to S that is $M_2 = \{D_3, D_4, D_5, N_a, TS_{C2}\}$ while proceeding for authentication phase.

3.2.3. Authentication Phase. Sensor matches TS_{C2} and TS_{S2}, both the timestamp values and terminate the authentication process if both timestamp are mismatch while compute the values of $N_c = D_3 \oplus h(N_s \| TS_{S2})$ and $ID_c = D_4 \oplus h(b \| N_s \| N_c)$. Then check the match case of D_5 with $D_{5'}$, if equals, then take a timestamp TS_3 from A and calculates $D_6 = h(b \| N_s \| TS_{S3})$ and send the encrypted message to C as $M_3 = \{D_6, TS_{S3}\}$. C selects the timestamp TS_{C3} shared by A and compares with TS_{S3} and terminate the authentication process if both timestamp are not matches. In match case, C verifies the value of D_6 with $D_{6'}$. If equal, computes $D_7 = h(a \| N_s \| N_c \| TS_{C4})$, the key $\kappa = h(N_s \| N_c \| ID_s)$ and send the message $M_4 = \{D_7, TS_{C4}\}$ to the sensor S.

3.2.4. Update Phase. The update phase is an additional feature in which S sends OK after decrypts the message and make some calculation here. When S decrypts the message $M_4 = \{D_7, TS_{C4}\}$ and make some calculation here, sends OK for the confirmation of authentication phase to C that the authentication process is successfully secured. Moreover, the sensor selects TS_4 and verifies TS_{C4} with TS_{S4}, if true then verifies $D_7 = h(a \| N_s \| N_c \| TS_{C4})$ with the fresh value. If not verify, then terminates. If true, then computes $\kappa = h(N_s \| N_c \| ID_s)$. While after completing the authentication process, sensor fresh the values and sends the OK message that means to update all the values such that $ID_s = h(a' \|b' \| N_{s'} \| N_{c'})$, to avoid from the guessing attack. So, S and C updates there values to avoid from guessing attack on the previous values and now they can share secrets in a secure manner, and the process of all the authentication mechanism is completed. For more illustration, the whole authentication process taken from Figure 6 of SF-LAP protocol is given in pseudocode 1.

4. Security Analysis

To prove that our suggested SF-LAP protocol is more secure than others, we performed formal and informal security research. The three methods of verification that we used in our paper are BAN and GNY logic [39] verification of SF-LAP, AVISPA verification of SF-LAP, and informal security analysis.

4.1. BAN and GNY Logic Verification of SF-LAP. We used these [39] logics to verify the authentication mechanism, prove that SF-LAP provides acceptable authentication between sensor and controller and synchronizes both of them with the server clock. In these logics, there are a variety of notations that are employed. In our scenario, S and C are the agents instead of P and Q. M is used for the message, κ

for a secret key, K^{-1} for the private key, S for the sensor, and C for the controller. The following constructions are employed to demonstrate the use and connections of these logic.

(a) $S\equiv M$: means that the S believes M, and M are true

(b) $S\triangleleft M$: indicates that the S received/sees the message M.

(c) $S \ni M$: indicates that the S possesses the message M.

(d) $S \sim M$: indicates that S sends messages at the sending time of message M. The S believed M, or S once said M.

(e) $S \Rightarrow M$: if other principals believe S believes M, then M means the other principals believe S believes M. M is under jurisdiction of S.

(f) $P\#(M)$: indicates that M is fresh M had never been sent before the current run of the protocol. Nonces are expressions created to demonstrate freshness, and they frequently include a timestamp. Without using nonces, we can get that not-so-fresh message feeling

(g) $S \leftrightarrow^{K} C$:S and C share the key K and may use it to communicate. No principal other than S or C or a principal trusted by S or C can find K.

(h) $\longrightarrow^{K} S$:S has the session key K.

(i) $\{M\}_K$:K is the key used to encrypts M. This represents a signature of M when K is a private key. For example, K^1

To get from protocol steps to logical inferences, we use idealization. This tries to convert the sent message into the intended semantics. For example: in the given protocol $C \leftrightarrow^{K} S$ means that the key is shared between C and S. One goal of idealization is to leave out parts of the message that do not add to the recipients' beliefs. The plaintext is not used in these logics because it can be forged. The protocol's idealization is not clearly defined. It is contingent on how specific steps are interpreted.

4.1.1. BAN and GNY Logic Rules. These rules are taken from a paper [40] that provides the formal analysis required to verify the authentication protocol. S and C are used for sensor and controller, M and N are used for statement. Message meaning rules govern how messages are interpreted. Rather than write the English equivalents, we used the predefined symbols. When using a shared key set:

$$\frac{S\mid\equiv C\xrightarrow{K}S,\ S\triangleleft\{M\}K}{S\mid\equiv C\mid\sim M}. \tag{1}$$

Table 3: Paper's notations and description.

Notation	Description
PSK	Preshared key
S	Sensor
C	Controller
A	Authentication server
ID_s	Identity of S
ID_c	Identity of C
M	Message
N_s	Nonce of S
N_c	Nonce of C
N_a	Nonce of A
N_K	Nonce revoke
h	Hash function
||	Concatenation
$\oplus$	XOR operation
$\{M\}_K$	Encrypted message with session key
TS	Timestamp
K^{-1}	Private key
κ	Session key

When using session keys,

$$\frac{S\mid\equiv\ \xrightarrow{K} C,\ S\triangleleft\{M\}_K^{-1}}{S\mid\equiv C\mid\sim M}. \tag{2}$$

A principal S must believe in $C^{'}s$ beliefs according to the Jurisdiction rule:

$$\frac{S\mid\equiv C\Rightarrow M,\ S\mid\equiv C\mid\equiv M}{S\mid\equiv M}. \tag{3}$$

The rules for nonce-verification show how to ensure that communication is both new and fresh in the senders believe:

$$\frac{S\mid\equiv(\#(M)),\ S\mid\equiv C\mid\sim M}{S\mid\equiv C\mid\equiv M}. \tag{4}$$

If any component of the formula is fresh, it cannot be changed; therefore, the whole formula has to be fresh:

$$\frac{S\mid\equiv(\#(M))(M)}{S\mid\equiv(\#(M))(M,N)}. \tag{5}$$

When both communication devices have believe on each other, then Belief rule is:

$$\frac{S\mid\equiv C\mid\equiv(M,N)}{S\mid\equiv C\mid\equiv(M)}. \tag{6}$$

SF-LAP is a secured way to share a key between the S

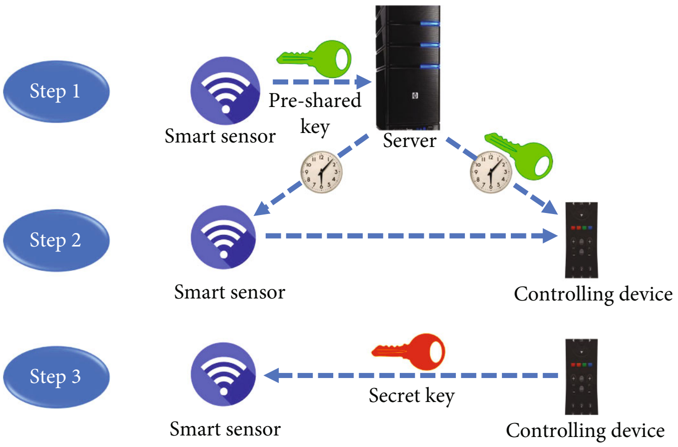


Figure 5: A use case diagram for the authentication process.

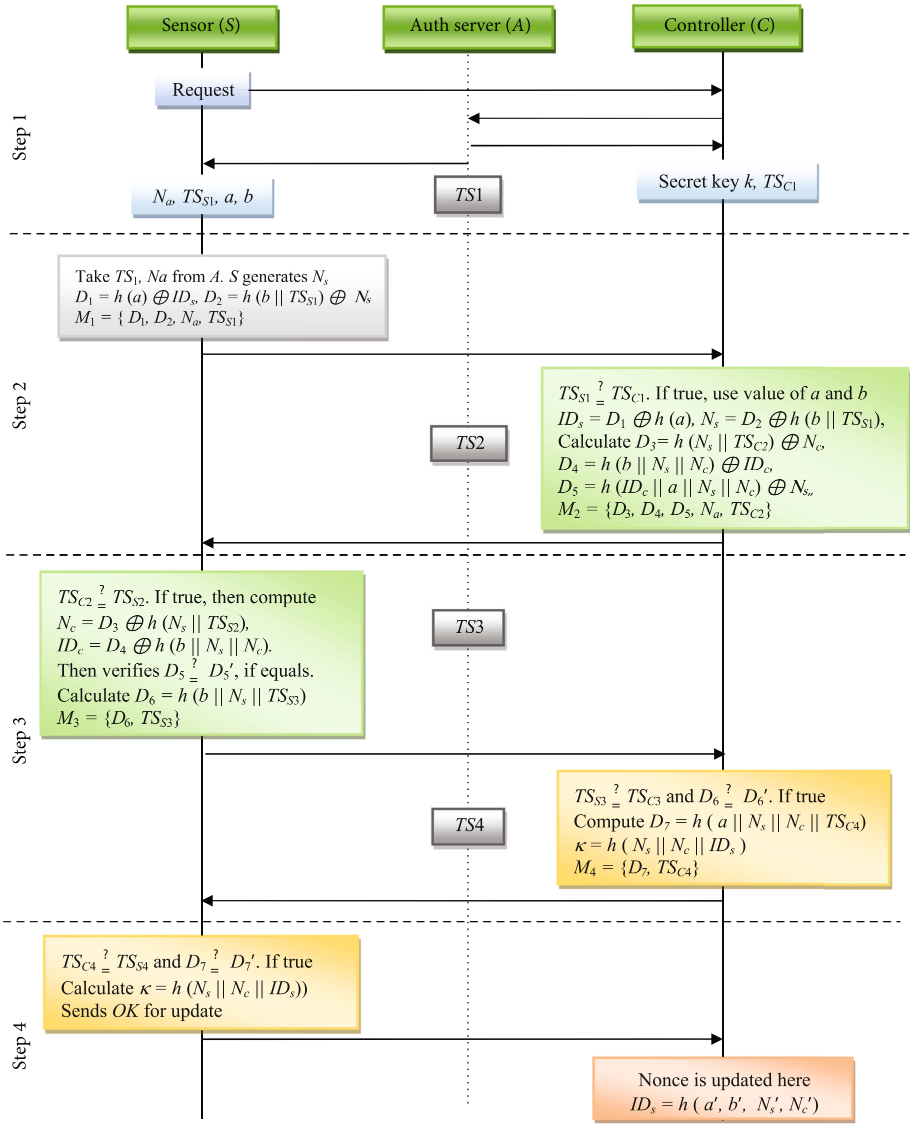

Sensor (S)
Auth server (A)
Controller (C)
Step 1
Request
N_a, TS_{S1}, a, b
$TS1$
Secret key k, TS_{C1}
Step 2
Take TS_1, Na from A. S generates N_s
$D_1 = h(a) \oplus ID_s$, $D_2 = h(b \,||\, TS_{S1}) \oplus N_s$
$M_1 = \{D_1, D_2, N_a, TS_{S1}\}$
$TS2$
$TS_{S1} \overset{?}{=} TS_{C1}$. If true, use value of a and b
$ID_s = D_1 \oplus h(a)$, $N_s = D_2 \oplus h(b \,||\, TS_{S1})$,
Calculate $D_3 = h(N_s \,||\, TS_{C2}) \oplus N_c$,
$D_4 = h(b \,||\, N_s \,||\, N_c) \oplus ID_c$,
$D_5 = h(ID_c \,||\, a \,||\, N_s \,||\, N_c) \oplus N_{s,,}$
$M_2 = \{D_3, D_4, D_5, N_a, TS_{C2}\}$
Step 3
$TS_{C2} \overset{?}{=} TS_{S2}$. If true, then compute
$N_c = D_3 \oplus h(N_s \,||\, TS_{S2})$,
$ID_c = D_4 \oplus h(b \,||\, N_s \,||\, N_c)$.
Then verifies $D_5 \overset{?}{=} D_5'$, if equals.
Calculate $D_6 = h(b \,||\, N_s \,||\, TS_{S3})$
$M_3 = \{D_6, TS_{S3}\}$
$TS3$
$TS4$
$TS_{S3} \overset{?}{=} TS_{C3}$ and $D_6 \overset{?}{=} D_6'$. If true
Compute $D_7 = h(a \,||\, N_s \,||\, N_c \,||\, TS_{C4})$
$\kappa = h(N_s \,||\, N_c \,||\, ID_s)$
$M_4 = \{D_7, TS_{C4}\}$
Step 4
$TS_{C4} \overset{?}{=} TS_{S4}$ and $D_7 \overset{?}{=} D_7'$. If true
Calculate $\kappa = h(N_s \,||\, N_c \,||\, ID_s))$
Sends OK for update
Nonce is updated here
$ID_s = h(a', b', N_s', N_c')$

Figure 6: Sensor and controller authentication mechanism via server.

```
(//) double slash are used for comments
Procedure SF-SLAP AUTHENTICATION PROTOCOL
//we assume that A already known PSK
//S takes $TS_1$ and $N_a$ from A and generates $N_s$
//S takes the value of a and x or with $ID_s$
//S takes value of b, concatenate with $TS_1$ and x or with $N_s$
$S| \sim M_1$ to C //sensor sends message $M_1$ to controller
    if ($C \lhd M_1$, and $TS_1$ match) then //controller receives message $M_1$ and match the timestamp $TS_1$
      return true;
        if ($C | \sim D_3, D_4, D_5, N_a, TS_2$ to S) then //controller sends $D_3$, $D_4$, $D_5$, $N_a$, and $TS_2$ to sensor
          return true;
            if ($S \lhd M_2$, and $TS_2$ match) then //sensor receives message $M_2$ and match the timestamp $TS_2$
              return true:
                if ($S | \sim D_6, TS_3$ to C) then //sensor sends $D_6$ with timestamp $TS_3$ to controller
                  return true;
                  if ($C \lhd M_3$, and $TS_3$ match) then //controller receives $M_3$ and matches the timestamp $TS_3$
                      return true;
                        if ($C | \sim D_7$ and $TS_4$ to S) then //controller sends $D_7$ and timestamp $TS_4$ to sensor
                          return true;
                            if ($S \lhd M_4$, and $TS_4$ match) then //sensor receives message $M_4$ and match timestamp $TS_4$
                              return true;
                              while (S sends OK to C for update the values of a', b', $N_s$', and $N_c$')
                            else
                            return false; //$TS_4$ does not match
                        else
                        return false;
                    else
                    return false; //$TS_3$ does not match
                else
                return false;
            else
            return false; //$TS_2$ does not match
        else
        return false;
    else
    return false; //$TS_1$does not match
end procedure
```

Pseudocode 1: SF-SLAP authentication protocol.

and *C*. So, the goal of our proposed technique is to firstly ensure that the shared key is secured between the *S* and *C*.

4.1.2. Goals. Goal 1: $S| \equiv (S \xrightarrow{K} C)$

Goal 2: $S| \equiv C| \equiv (S \xrightarrow{K} C)$

Goal 3: $C| \equiv (S \xrightarrow{K} C)$

Goal 4: $C| \equiv S| \equiv (S \xrightarrow{K} C)$

4.1.3. Idealized Form. SF-LAP protocol according to idealized Needham-Schroeder Shared-Key [BAN89a] is as follows:

M1: $S \rightarrow A: S, C, N_A$
M2: $A \rightarrow S: \{N_A, C, k_{SC}, \{k_{SC}, S\}k_{CA}\}k_{SA}$
M3: $S \rightarrow C: \{k_{CA}, S\}k_{CA}$
M4: $C \rightarrow S: \{N_C\}k_{SC}$
M5: $S \rightarrow C: \{N_C{-}1\}k_{SC}$

4.1.4. Assumptions. The following assumptions we have been made are based on the SF-LAP authentication protocol from Figure 6:

(i) A believes PSK is shared to S and C $S \overset{Ksc}{\leftrightarrow} C$

(ii) S and C believes that A fresh TS $(S \overset{TS1}{\leftrightarrow} C)$

(iii) A believes S, and C receive TS1

(iv) S believes that C receives M1

(v) C believes that M1 is send by S

(vi) C and S believe that A fresh TS $(S \overset{TS2}{\leftrightarrow} C)$

(vii) A believes S, and C receives TS2

(viii) C believes that S receives M2

(ix) S believes that M2 is send by C

(x) S and C believe that A fresh TS (S $\overset{TS3}{\leftrightarrow}$ C)

(xi) A believes S and C receive TS3

(xii) S believes that C receives M3

(xiii) C believes that M3 is send by S

(xiv) C and S believe that A fresh TS (S $\overset{TS4}{\leftrightarrow}$ C)

(xv) A believes S, and C receives TS4

(xvi) C believes that S receives M4

(xvii) S believes that M4 is send by C

4.1.5. Some [BAN89a] Rules for SF-LAP. R1: S |≡ (S $\overset{Ksa}{\leftrightarrow}$ A): S believes that the key is shared between S and A

R2: C |≡ (C $\overset{Kca}{\leftrightarrow}$ A): C believes that the key is shared between C and A

R3: S |≡ A ⇒S $\overset{K}{\leftrightarrow}$ C: S believes that A controls key between S and C

R4: C |≡ A ⇒S $\overset{K}{\leftrightarrow}$ C: C believes that A controls key between S and C

R5: S |≡ A ⇒#(S $\overset{K}{\leftrightarrow}$ C): S believes that A controls and fresh the value of key, shared between S and C

R6: S |≡ # (Ns): S believes that the value of nonce is fresh by S

R7: C |≡ # (Nc): C believes that the value of nonce is fresh by S

R8: S |~ {Ns, S $\overset{Ksc}{\leftrightarrow}$ C, # (Ksc), { S $\overset{Ksc}{\leftrightarrow}$ C }Kca }Ksa : S received the fresh encrypted values

R9: C |~ {S $\overset{Ksc}{\leftrightarrow}$ C}Kca from A

R10: S |~ {Nc, S $\overset{Ksc}{\leftrightarrow}$ C}Ksc from C

R11: C |~ {Nc-1, S $\overset{Ksc}{\leftrightarrow}$ C}Ksc from S

R12: C |≡ # (S $\overset{Ksc}{\leftrightarrow}$ C): C believes that value of key is fresh, shared between S and C

4.1.6. Proof from BAN and GNY Logic. By applying BAN and GNY logic [39] on SF-LAP protocol:

(i) S | ≡ A | ~ (Ns, S $\overset{Ksc}{\leftrightarrow}$ C, #(S $\overset{Ksc}{\leftrightarrow}$ C), {S $\overset{Ksc}{\leftrightarrow}$ C}Kca)

Message meaning rule using R1 and R8

(ii) S | ≡ #(Ns, S $\overset{Ksc}{\leftrightarrow}$ C, #(S $\overset{Ksc}{\leftrightarrow}$ C), {S $\overset{Ksc}{\leftrightarrow}$ C}Kca)

Freshness conjuncatenation rule using 1 and R6

(iii) S | ≡ A | ≡ (Ns, S $\overset{Ksc}{\leftrightarrow}$ C, #(S $\overset{Ksc}{\leftrightarrow}$ C), {S $\overset{Ksc}{\leftrightarrow}$ C}Kca)

Nonce verification rule using 2 and 1

(iv) S | ≡ A | ≡ (S $\overset{Ksc}{\leftrightarrow}$ C)

Belief Conjuncatenation rule using 3

(v) S | ≡ A | ≡ (# S $\overset{Ksc}{\leftrightarrow}$ C)

Belief conjuncatenation rule using 3

(vi) S | ≡ (S $\overset{Ksc}{\leftrightarrow}$ C)

Jurisdiction rule using 4 and R3

(vii) S | ≡ #(S $\overset{Ksc}{\leftrightarrow}$ C)

Jurisdiction rule using 4 and R5 (Goal 1 achieved)

(viii) C | ≡ A | ~ S $\overset{Ksc}{\leftrightarrow}$ C

Message meaning rule using R2 and R9

(ix) C | ≡ S | ≡ S $\overset{Ksc}{\leftrightarrow}$ C

Nonce verification rule using R12 and 8 (Goal 4 achieved)

(x) C | ≡ S $\overset{Ksc}{\leftrightarrow}$ C

Jurisdiction rule using R4 and 9 (Goal 3 achieved)

(xi) S | ≡ C | ~ (Nc, S $\overset{Ksc}{\leftrightarrow}$ C)

Message meaning rule using 6 and R10

(xii) S | ≡ # (Nc, S $\overset{Ksc}{\leftrightarrow}$ C)

Freshness conjuncatenation rule using 7

(xiii) S | ≡ C | ≡ (Nc, S $\overset{Ksc}{\leftrightarrow}$ C)

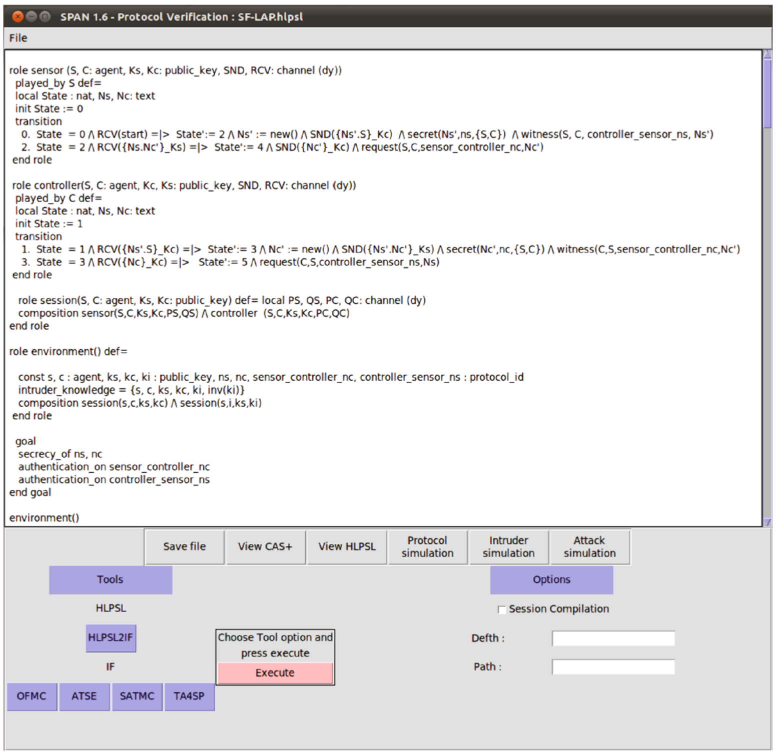
Nonce verification rule using 12 and 11

(xiv) S | ≡ C | ≡ S $\overset{Ksc}{\leftrightarrow}$ C

Belief conjuncatenation rule using 13 (Goal 2 achieved)

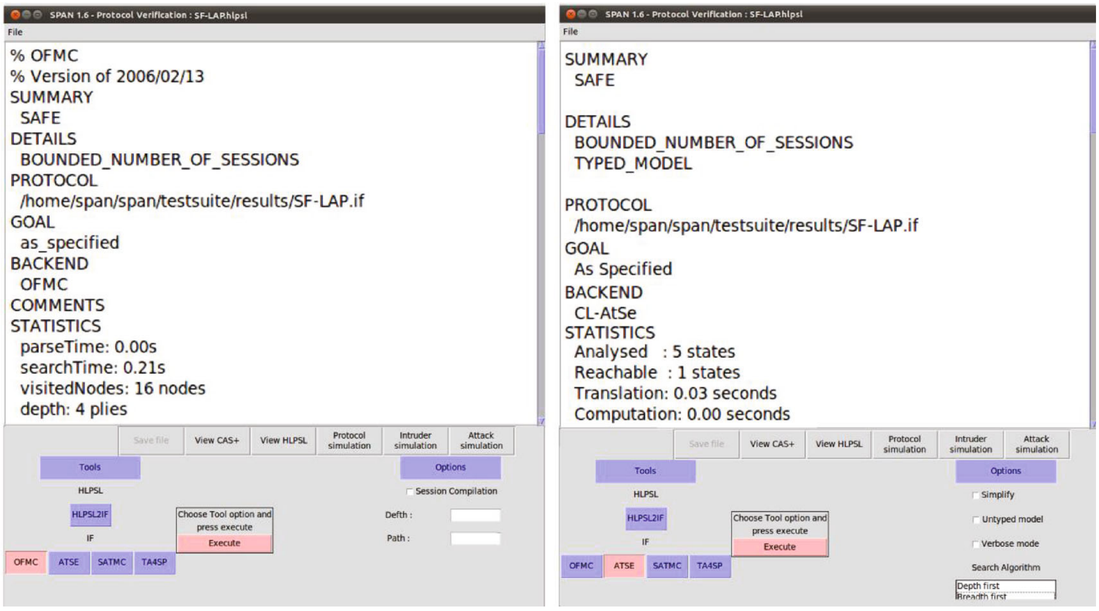
4.2. AVISPA Verification of SF-LAP. A powerful tool AVISPA determined that SF-LAP provides sufficient authentication against desynchronization attacks. In the AVISPA tool, four backends entities are responsible for a specified function at execution. HLPSL is used for a high-level language, and HLPS2IF is used for the translator.

The HLPSL representations of the proposed protocol, including the sensor and controller, are shown in Figure 7. AVISPA automatically validates security protocols and determines whether they are classified as either SAFE or UNSAFE depending on predetermined criteria. AVISPA validates that SF-LAP is SAFE protocol. Figure 8 shows the AVISPA simulation results using the on the fly model checker (OMFC) and constraint logic-based attack searcher (CT-ATSE) backends, demonstrating that SF-LAP is secure.



```
role sensor (S, C: agent, Ks, Kc: public_key, SND, RCV: channel (dy))
  played_by S def=
  local State : nat, Ns, Nc: text
  init State := 0
  transition
    0.  State  = 0 /\ RCV(start) =|>  State':= 2 /\ Ns' := new() /\ SND({Ns'.S}_Kc)  /\ secret(Ns',ns,{S,C})  /\ witness(S, C, controller_sensor_ns, Ns')
    2.  State  = 2 /\ RCV({Ns.Nc'}_Ks) =|>  State':= 4 /\ SND({Nc'}_Kc) /\ request(S,C,sensor_controller_nc,Nc')
 end role

 role controller(S, C: agent, Kc, Ks: public_key, SND, RCV: channel (dy))
  played_by C def=
  local State : nat, Ns, Nc: text
  init State := 1
  transition
    1.  State  = 1 /\ RCV({Ns'.S}_Kc) =|>  State':= 3 /\ Nc' := new() /\ SND({Ns'.Nc'}_Ks) /\ secret(Nc',nc,{S,C}) /\ witness(C,S,sensor_controller_nc,Nc')
    3.  State  = 3 /\ RCV({Nc}_Kc) =|>   State':= 5 /\ request(C,S,controller_sensor_ns,Ns)
 end role

   role session(S, C: agent, Ks, Kc: public_key) def= local PS, QS, PC, QC: channel (dy)
   composition sensor(S,C,Ks,Kc,PS,QS) /\ controller  (S,C,Ks,Kc,PC,QC)
end role

role environment() def=

   const s, c : agent, ks, kc, ki : public_key, ns, nc, sensor_controller_nc, controller_sensor_ns : protocol_id
   intruder_knowledge = {s, c, ks, kc, ki, inv(ki)}
   composition session(s,c,ks,kc) /\ session(s,i,ks,ki)
 end role

  goal
   secrecy_of ns, nc
   authentication_on sensor_controller_nc
   authentication_on controller_sensor_ns
end goal

environment()
```

Figure 7: HLPSL representation of SF-LAP.



```
% OFMC
% Version of 2006/02/13
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
  /home/span/span/testsuite/results/SF-LAP.if
GOAL
  as_specified
BACKEND
  OFMC
COMMENTS
STATISTICS
  parseTime: 0.00s
  searchTime: 0.21s
  visitedNodes: 16 nodes
  depth: 4 plies
```

```
SUMMARY
  SAFE

DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL

PROTOCOL
  /home/span/span/testsuite/results/SF-LAP.if
GOAL
  As Specified
BACKEND
  CL-AtSe
STATISTICS
  Analysed   : 5 states
  Reachable  : 1 states
  Translation: 0.03 seconds
  Computation: 0.00 seconds
```

Figure 8: AVISPA result using OMFC and ATSE back-end.

4.3. Informal Security Analysis. SLAP claims that the shared key between sensor and controller was not secured, so it is vulnerable as we proved in section 2.1 on how SLAP is compromised against eavesdropping, impersonation, replay, and desynchronization attack, and hence compromised the shared key and was unable to secure the communication. Our approach uses Nonce and timestamp to overcome this vulnerability.

4.3.1. SF-LAP Resists Replay Attacks. A replay attack repeats data and allows the attacker to delay it. For timestamp purposes, SF-LAP uses server time to synchronize the time for the S and the C. The server clock records the time of the S and sets a limited time threshold for the message; if there is a delay, the message is automatically terminated. Thus, the issue of message delay is avoided by this protocol. The authentication process's hash, all messages, and the attack are all altered if the attacker changes the timestamp. Replay attacks are therefore impossible with SF-LAP.

4.3.2. SF-LAP Resists Impersonation Attacks. By locating identity or initial handshake vulnerabilities, an adversary must launch an attack in place of the server. As we add the server between S and C, we hide the sensor IDs. S already gave A his IDs, so A has no chance to show his ID during this procedure. There is no way for an adversary to interfere with or compromise the system at the sensor location because the controller restricts new sensors here in place of verified sensors. Besides this, an adversary who is unable to discover the values of D1 and D2 is unable to overhear C and S talking. To secure this process, the request message is sent from the actual sensor to the controller using the value of Nonce from authentication and the match case value of D3. From the controller, the sensor gets D'3. The controller receives M1, which was forwarded and gleaned from the previous session after the sensor has sent D'3.

After the controller receives M1, the sensor receives M2, which verifies the validity of the message from the sensor. The sensor makes sure that the controller has confirmed the controller legitimacy after receiving M2. As a result, the adversary's offensive mission has been unsuccessful. The attacker must then prevent the controller and reliable sensor from communicating with each other using the previous session's stolen message. The controller in this message chose to use Nonce rather than a simple random number to prevent an adversary's plan. The identification of the controller is recognized once the sensor's identity has been established. Therefore, an attacker cannot guess the IDs or IDc, and the protocol will be secure from impersonation attacks.

4.3.3. SF-LAP Resists Eavesdropping Attacks. The messages of C cannot be found and followed by an attacker among the correspondence between C and S. This communication between S and C cannot be deleted or modified. An adversary cannot intercept this data. The server also knows PSK, which has previously communicated it to the C. As a result, the network is secure when $b = h$ (Na ‖ PSK), D_1 = h(a) ⊕ IDs and D_2= h (b $||TS_{S1}$), an authentication protocol successfully offers resilience against eavesdropping on the session between C and S. Since the value of Ns is a nonce technique, it is impossible for an adversary to guess because it is an arbitrary number along with the random number generated by S, which changes its value dynamically. As a result, an attacker is prevented from attacking at the beginning of the session, and the values of D_1 and IDs are protected.

In step two, the attacker can no longer read the message on the controller side, and the C will now send the ID_c to the S. At this level, the attacker has no chance of succeeding, so the attacker cannot serve as S, and hence actual S sends M_1 to the C. Response C compares the values with D_1 after receiving the M_1 and reads it that was sent from the S to the C and C does not know the value that comes from the previous session. So, SF-LAP protects against eavesdropping attacks.

4.3.4. SF-LAP Resists MITM Attacks. An adversary successfully interrupts communication between S and C resulting in a man-in-the-middle attack. IDs are not shared during the authentication process because they are declared to the C in SF-LAP ahead of time as the sensor ID. Our approach secures against MITIM attacks during communication by making sure the shared secret key is safe before the authentication procedure. When b is successfully received by C, the value of b is saved as b = h (Na || PSK), and the message from the controller to the sensor cannot be passed by the adversary or prevented from being used to carry out a MITM attack by intercepting a request message from the S and sending it to the C.

4.3.5. SF-LAP Resists Desynchronization Attacks. The sensor compared the D'3 message to the D3 and Na, which were therefore unknown to the adversary, before sending the message M1 to the controller. The attacker in this instance is ineffective and unable to compute the shared key because they are unable to carry out the attack, and are unaware that M2 is being used to deliver the sensor's values for D4, D5, D6, and D7 through b'= h (Na' || PSK), and do not know the value of D7 == D7'. On the A is where the PSK is kept, and the server clock is used to time stamp the S and C. The S and C are timestamped using the server clock, and the PSK is stored on the A.

An adversary cannot interrupt on the back-end side because Ns, Nc, and Na prevent desynchronization and the TS of the server is used simultaneously on both sides. Because special Nonce Ns on the S side and Nc on the C side, along with TS1, TS2, TS3, or TS4, are concatenated with each other and TS is fresh in each phase and does not store the old value in it, if an adversary wants to modify something, hashes will change. Additionally, because Nonce is tied to the server clock, it changes dynamically every second. Therefore, SF-LAP resists desynchronization attacks since there is no chance that IDs and IDc will be changed.

4.3.6. SF-LAP Resists Modification Attacks. The values are one-way hashes in step 2 of the SF-LAP authentication protocol; specifically, D1 hashes the value of a, D2 hashes the value of b, D3 hashes Ns concatenated with TSC2, D4 hashes the concatenated values of b, Ns, and Nc, and D5 hashes the concatenated value of IDc, a, Ns, and Nc. A modification attack happens when an adversary modifies one of the messages M1, M2, M3, or M4. S and C exchange information and timestamp each other using the server clock. The values of b, Ns, and D7 are hashed in step 3, along with the values of a, Ns, and Nc. The server timestamp and the value of each of these Ds, such as TSS1 with D2, TSC2 with D3, TSS3 with D6, and TSC4 with D7, make the hash values unique at each stage. The hash of all D's values will change if an adversary attempts to modify any of them because they are all concatenated and XORed together, making it impossible for the adversary to modify any step. SF-LAP resists modification attacks as a result.

4.3.7. SF-LAP Resists Tracing Attacks. When the IDs or IDc remain the same throughout the process, a tracing attack takes place. As the authentication mechanism is synchronized with the server clock, which an adversary cannot alter, the nonce dynamically changes the values here. The timestamp produced by the server clock ensures that an adversary cannot break it. Additionally, we use a nonce technique in conjunction with a timestamp to make sure that every number differs from the one before it in a random manner, preventing the adversary from determining the true key. Additionally, SF-LAP added an update phase in which S validates TSS4 and D7 before updating all the values, including a', b', Ns', and Nc', which are sent to C and updated upon the OK request. The adversary will be unable to guess or track these values once all the previous values have been deleted from this authentication system.

4.3.8. SF-LAP Resists Compromised Attacks. If the attacker has access to the key and can compute the message M1, M2, M3, or M4, it constitutes a compromised attack. Before this communication, we have a secret key that is shared over a secure channel and contains a nonce technique. A compromised attack cannot succeed against this protocol because an adversary cannot guess it.

5. Conclusion

A single-factor lightweight authentication protocol for machine-to-machine communication in the industrial Internet of things is proposed in this paper. The proposed protocol uses an exclusive-OR operation and a hashing function to provide a secure mechanism for conversation, ensuring secure communication between the sensor and the controller. It protects and safeguards this connection against desynchronization and eavesdropping attacks using a secure preshared key and timestamp technique. To ensure that SF-LAP is secure, we used BAN and GNY logic and the formal verification by using AVISPA tool and security analysis, verify that SF-LAP is secure. Finally, we conclude that SF-LAP is more secure because it prevents desynchronization, eavesdropping, and all other types of attacks such as modification attacks, tracing attacks, man-in-the-middle, impersonation, and replay attacks. The secure method of authentication provided by the SF-LAP protocol is available when a sensor is already connected to a controller. What steps will be taken if we want to connect a new sensor to the same controller? We will explain this authentication mechanism in more detail later on, along with how a new sensor can connect to the same controller.

Data Availability

All the relevant data are available in the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This project has been sponsored by the University of Liberal Arts Bangladesh (ULAB), Dhaka Bangladesh.

References

[1] O. A. Amodu and M. Othman, “Machine-to-machine communication: an overview of opportunities,” *Computer Networks*, vol. 145, pp. 255–276, 2018.

[2] M. Weyrich, J.-P. Schmidt, and C. Ebert, “Machine-to-machine communication,” *IEEE Software*, vol. 31, no. 4, pp. 19–23, 2014.

[3] Z. Zhou, J. Gong, Y. He, and Y. Zhang, “Software defined machine-to-machine communication for smart energy management,” *IEEE Communications Magazine*, vol. 55, no. 10, pp. 52–60, 2017.

[4] M. Hasan, E. Hossain, and D. Niyato, “Random access for machine-to-machine communication in lte-advanced networks: issues and approaches,” *IEEE Communications Magazine*, vol. 51, no. 6, pp. 86–93, 2013.

[5] K. Saleem, A. Derhab, J. Al-Muhtadi, and B. Shahzad, “Human-oriented design of secure machine-to-machine communication system for e-healthcare society,” *Computers in Human Behavior*, vol. 51, pp. 977–985, 2015.

[6] R. Daş and G. Tuna, “Machine-to-machine communications for smart homes,” *International Journal of Computer Networks and Applications*, vol. 2, no. 4, pp. 196–202, 2015.

[7] A. Corallo, M. Lazoi, M. Lezzi, and A. Luperto, “Cybersecurity awareness in the context of the industrial Internet of things: a systematic literature review,” *Computers in Industry*, vol. 137, article 103614, 2022.

[8] J. Wang, J. Chen, Y. Ren, P. K. Sharma, O. Alfarraj, and A. Tolba, “Data security storage mechanism based on blockchain industrial Internet of things,” *Computers & Industrial Engineering*, vol. 164, article 107903, 2022.

[9] M. Ali, M.-R. Sadeghi, X. Liu, Y. Miao, and A. V. Vasilakos, “Verifiable online/offline multi-keyword search for cloud-assisted industrial Internet of things,” *Journal of Information Security and Applications*, vol. 65, article 103101, 2022.

[10] T. Hasan, J. Malik, I. Bibi et al., “Securing industrial internet of things against botnet attacks using hybrid deep learning approach,” *IEEE Transactions on Network Science and Engineering*, 2022.

[11] D. K. Sah, T. N. Nguyen, K. Cengiz, B. Dumba, and V. Kumar, “Load-balance scheduling for intelligent sensors deployment in industrial internet of things,” *Cluster Computing*, vol. 25, no. 3, pp. 1715–1727, 2022.

[12] C. Liu, S. Ziwei, X. Xun, and L. Yuqian, “Service-oriented industrial internet of things gateway for cloud manufacturing,” *Robotics and Computer-Integrated Manufacturing*, vol. 73, article 102217, 2022.

[13] C. Li, Y. Liu, J. Xiao, and J. Zhou, “Mceaaco-qsrp: a novel qos secure routing protocol for industrial internet of things,” *IEEE Internet of Things Journal*, 2022.

[14] B. Sriramulu, P. V. Rao, D. R. Vemula, B. R. Reddy, and T. J. Lakshmi, “A secure iot-based micro-payment protocol for wearable devices,” *Peer-to-Peer Networking and Applications*, vol. 15, no. 2, pp. 1163–1188, 2022.

[15] X. Ding, X. Wang, Y. Xie, and F. Li, “A lightweight anonymous authentication protocol for resource-constrained devices in

internet of things," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1818–1829, 2022.

[16] B. H. Taher, H. Liu, F. Abedi, H. Lu, A. A. Yassin, and A. J. Mohammed, "A secure and lightweight three-factor remote user authentication protocol for future IoT applications," *Journal of Sensors*, Article ID 8871204, 2018 pages, 2021.

[17] K. Dewangan, M. Mishra, and N. K. Dewangan, "A review: a new authentication protocol for real-time healthcare monitoring system," *Irish Journal of Medical Science*, vol. 190, no. 3, pp. 927–932, 2021.

[18] M. M. Modiri, J. Mohajeri, and M. Salmasizadeh, "A novel group-based secure lightweight authentication and key agreement protocol formachine-type communication," *Scientia Iranica*, 2021.

[19] K. Shahzad, A. O. Aseeri, and M. A. Shah, "A blockchain-based authentication solution for 6g communication security in tactile networks," *Electronics*, vol. 11, no. 9, p. 1374, 2022.

[20] S. Gupta, B. L. Parne, and N. S. Chaudhari, "ISAG: IoT-enabled and cecrecy aware group-based handover scheme for e-health services in M2M communication network," *Future Generation Computer Systems*, vol. 125, pp. 168–187, 2021.

[21] C. Wang, R. Huang, J. Shen, J. Liu, P. Vijayakumar, and N. Kumar, "A novel lightweight authentication protocol for emergency vehicle avoidance in VANETs," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14248–14257, 2021.

[22] M. A. Mobarhan and M. Salamah, "REPS-AKA3: a secure authentication and re- authentication protocol for LTE networks," *Journal of Network and Computer Applications*, vol. 201, article 103345, 2022.

[23] E. Karacan, S. Akleylek, and A. Karakaya, "Pq-flat: a new quantum-resistant and lightweight authentication approach for m2m devices," in *2021 9th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1–5, Elazig, Turkey, 2021.

[24] H. Dalkilic and M. H. Ozcanhan, "A strong mutual authentication protocol for securing wearable smart textile applications," *Advances in Electrical and Computer Engineering*, vol. 22, no. 1, pp. 31–38, 2022.

[25] V. Joanne Marie, J. E. Santos, V. Pascua, and N. M. C. Tiglao, "Hardware-accelerated blockchain-based authentication for the Internet of things," in *Cognitive Radio Oriented Wireless Networks and Wireless Internet. CROWNCOM WiCON 2021 2021*, H. Jin, C. Liu, A. S. K. Pathan, Z. M. Fadlullah, and S. Choudhury, Eds., vol. 427 of Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, pp. 283–295, Springer, Cham, 2022.

[26] A. H. Aly, A. Ghalwash, M. Nasr, and A. A. El-Hafez, "A new lightweight authenticated key agreement protocol for Iot in cloud computing," *Journal of Engineering Science and Technology*, vol. 16, no. 5, pp. 3987–4005, 2021.

[27] M. M. Samy, W. R. Anis, A. A. Abdel-Hafez, and H. D. Eldemerdash, "An optimized protocol of m2m authentication for internet of things (Iot)," *International Journal of Computer Network & Information Security*, vol. 13, no. 2, pp. 29–38, 2021.

[28] C. Thammarat and C. Techapanupreeda, "A secure authentication and key exchange protocol for m2m communication," in *2021 9th International Electrical Engineering Congress (iEECON)*, pp. 456–459, Pattaya, Thailand, 2021.

[29] R. Krishnasrija, A. K. Mandal, and A. Cortesi, *A lightweight mutual and transitive authentication mechanism for iot network*.

[30] M. P. Lokhande, D. D. Patil, L. V. Patil, and M. Shabaz, "Machine-to-machine communication for device identification and classification in secure telerobotics surgery," *Security and Communication Networks*, vol. 2021, Article ID 5287514, 16 pages, 2021.

[31] X. Wang, K. Fan, K. Yang et al., "A new RFID ultra-lightweight authentication protocol for medical privacy protection in smart living," *Computer Communications*, vol. 186, pp. 121–132, 2022.

[32] M. Hosseinzadeh, O. H. Ahmed, S. H. Ahmed et al., "An enhanced authentication protocol for RFID systems," *IEEE Access*, vol. 8, pp. 126977–126987, 2020.

[33] J. Ryu, O. Jihyeon, D. Kwon et al., "Secure ecc-based three-factor mutual authentication protocol for telecare medical information system," *IEEE Access*, vol. 10, pp. 11511–11526, 2022.

[34] Q. Fan, J. Chen, M. Shojafar, S. Kumari, and D. He, "SAKE∗: a symmetric authenticated key exchange protocol with perfect forward secrecy for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 6424–6434, 2022.

[35] A. Adeel, M. Ali, A. N. Khan et al., "A multi-attack resilient lightweight IoT authentication scheme," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 3, article e3676, 2022.

[36] E. Lara, L. Aguilar, M. A. Sanchez, and J. A. Garcia, "Lightweight authentication protocol for m2m communications of resource-constrained devices in industrial internet of things," *Sensors*, vol. 20, no. 2, p. 501, 2020.

[37] A. Esfahani, G. Mantas, R. Matischek et al., "A lightweight authentication mechanism for m2m communications in industrial iot environment," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 288–296, 2019.

[38] S. Panda, S. Mondal, and N. Kumar, "SLAP: a secure and lightweight authentication protocol for machine-to-machine communication in industry 4.0," *Computers & Electrical Engineering*, vol. 98, article 107669, 2022.

[39] P. Syverson and I. Cervesato, "The logic of authentication protocols," in *Foundations of Security Analysis and Design. FOSAD 2000*, R. Focardi and R. Gorrieri, Eds., vol. 2171 of Lecture Notes in Computer Science, pp. 63–137, Springer, Berlin, Heidelberg, 2000.

[40] S. Rostampour, N. Bagheri, Y. Bendavid, M. Safkhani, S. Kumari, and J. J. P. C. Rodrigues, "An authentication protocol for next generation of constrained Iot systems," *IEEE Internet of Things Journal*, 2022.

Hindawi
Journal of Sensors
Volume 2022, Article ID 4583512, 11 pages
https://doi.org/10.1155/2022/4583512

Review Article

A Comparative Analysis of Fraudulent Recruitment Advertisement Detection Methods in the IoT Environment

Ruiqi Wang,[1] **Ning Cao**,[1] **Yajie Guo**,[1] **Shujuan Ji**,[1] **and Sachin Kumar**[2]

[1]*Key Laboratory for Wisdom Mine Information Technology of Shandong Province, Shandong University of Science and Technology, Qingdao, China*
[2]*Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, India*

Correspondence should be addressed to Shujuan Ji; jane_ji2003@aliyun.com

Received 7 August 2022; Revised 16 September 2022; Accepted 7 October 2022; Published 8 November 2022

Academic Editor: Kathiravan Srinivasan

The growth of the Internet of Things has changed the way of job hunting. Online recruitment has gradually replaced the traditional offline recruitment mode. Some unscrupulous people use online recruitment platforms to publish fraudulent recruitment advertisements, which not only bring financial and reputational losses to job seekers but also harm the sustainable development of society. However, previous studies have not used unified evaluation metrics and datasets, and detecting fraudulent recruitment advertisements lacks systematic research. To resolve this problem, this paper selects four representative traditional learning methods (i.e., random forest, support vector machine (SVM), logistic regression, and Naïve Bayes) and three deep learning methods (i.e., TextCNN, gate recurrent unit (GRU), and bidirectional long-short-term memory (Bi-LSTM)), which perform good in natural language processing (NLP) and use the same evaluation metrics and datasets conducting comparative experiments on balanced and unbalanced datasets, respectively. The experimental results show that the TextCNN method achieves the best detection performance with relatively low energy consumption on the balanced dataset. All the metrics values are more significant than 0.93. On unbalanced datasets, the TextCNN method still performs best with increasing imbalanced proportion.

1. Introduction

With the development of the Internet of Things (IoT), people can quickly get information from electronic devices. At the same time, the primary recruitment method in the labor market has rapidly shifted from offline to online, and getting recruitment information from the Internet has become a mainstream way. IoT has changed the inefficient way of finding a job. Online recruitment has the advantages of effectiveness, easiness, and efficiency [1]. However, some unscrupulous people use the network platforms' weakness to post fraudulent recruitment advertisements on the Internet to deceive money and exploit labor in the name of recruitment. Some fake recruitment has evolved from fraud to violent robbery, threats, restriction of personal freedom, and other serious violations [2].

Fraudulent recruitment advertisements have become a nationwide social nuisance. According to the *2017 China Internet Users' Consumer Protection Rights Report* (https://wenku.baidu.com/view/f50682067ed5360cba1aa8114431b90d6d85894e.html), among all the fraud cases reported by protection rights, fraudulent part-time work was the most frequently reported type of fraud (accounting for 22.1%), and most of the fraudulent recruitment occurred in well-known recruitment platforms. Data from *China's Justice Big Data Service Platform* (http://data.court.gov.cn/pages/research.html) shows that fraudulent recruitment cases increased yearly. Data from *Ai Media Consulting* through the *2019 China Internet Recruitment Industry Market Research* (https://www.iimedia.cn/c400/63879.html) shows that among the various bad experiences on the online platforms, the most situation that job seekers minded was that the enterprise information was not true (accounting for 34.8%). The second was personal information leakage (accounting for 31.8%). It is thus clear that the detection of fraudulent recruitment advertisements is a critical problem to be solved urgently. Detecting fraudulent recruitment advertisements based on data generated by the IoT system not only helps safeguard the rights and interests of job seekers but also promotes economic growth and creates a green IoT environment. Figure 1 shows the detection process for fraudulent recruitment advertisements in an IoT environment. However, this area is still a relatively untapped field, not receiving much attention from the academic community. In addition, detecting fraudulent recruitment advertisements among legitimate ones is a technically challenging problem [3]. Most research on fraudulent recruitment advertisements is carried out from the theoretical aspect. For example, Rubin [4] analyzed the causes and countermeasures of commercial fraud through advertising from the perspective of information economics, stating that deception was manipulating information to gain advantages. From a legal perspective, Jiang [5] puts forward the slogan "*Taking the law as a guarantee, strengthening advertising supervision functions, increasing rectification efforts, and cracking down on false and fraudulent advertisements.*"

The technology for detecting fraudulent recruitment advertisements is limited and immature. According to the methods adopted, the limited existing studies on the detection of fraudulent recruitment advertisements generally can be divided into three parts: traditional learning-based detection methods, traditional learning + feature extraction-based detection methods, and deep learning-based detection methods. Traditional learning-based detection methods mainly use traditional learning algorithms to detect fake job advertisements. Traditional learning + feature extraction-based detection methods mainly use feature extraction methods to improve the performance of traditional learning algorithms. Deep learning-based detection methods use various deep learning algorithms to detect employment frauds without feature engineering. The similarity of these detection methods is to identify the implicit fraud patterns in data. In particular, existing research involved different detection methods, different feature extraction methods, different evaluation metrics, and different datasets. Therefore, it is necessary to study the detection of fake recruitment advertisements systematically. To conduct a systematic study on detecting fake recruitment advertisements, this paper selects seven algorithms with good performance in NLP. It conducts systematic and comprehensive experiments using the same evaluation metrics and datasets. The main contributions are summarized as follows.

(1) The existing detecting methods of fraudulent recruitment advertisements are described and analyzed in detail

(2) A comparative analysis of the existing work is carried out experimentally using the same dataset and evaluation metrics. Seven algorithms are used to conduct comparative experiments on the public employment fraud detection dataset, and the experimental results are analyzed in detail

(3) The experimental results show that the TextCNN of the deep learning methods outperforms all other compared methods in the accuracy, precision, recall, and $F1$-score. In terms of time performance, though the training time of TextCNN is much higher than the traditional learning methods and the traditional learning + feature extraction methods, its testing time is similar to the ones of SVM and SVM+TF-IDF, which is acceptable for IoT devices. Therefore, comprehensively considering the accuracy, precision, recall, $F1$-score, and testing time, the TextCNN method performs best among all these compared methods

(4) This paper has a specific reference value for further research of higher performance and lower energy consumption detection methods to achieve the goal of green IoT

The organization of this paper is as follows. Section 2 reviews the relevant research on fraudulent recruitment detection. Section 3 introduces the representative methods. Section 4 presents the setting of experiments and the analysis of experimental results. Section 5 is the summary of the work of this paper and the prospect for the future.

2. Related Work

This section reviews the related work from three categories (i.e., the traditional learning-based detection methods, the traditional learning + feature extraction-based detection methods, and the deep learning-based detection methods) that we classify in Section 1 in detail.

2.1. Traditional Learning-Based Detection Methods. Some researchers used traditional learning-based methods to learn some rules. For example, Vidros et al. [6] analyzed employment fraud for the first time. They explained the work process of online recruitment and the role of the applicant tracking system in this process in detail. A set of rules of thumb was summarized from analyzing feasible data in the real world. Meanwhile, spam [7], insurance fraud, and phishing are highly similar to recruitment fraud, and they also discussed these areas in detail.

Besides, a more comprehensive and extensive study [8] was conducted based on the previous work [6]. They contributed and evaluated a real dataset of 17880 recruitment advertisements—Employment Scam Aegean Dataset (EMSCAD) (https://www.kaggle.com/datasets/shivamb/real-or-fake-fake-jobposting-prediction). Based on a subset of this dataset, they conducted word bag modeling and trained six WEKA (http://www.cs.waikato.ac.nz/ml/weka) classifiers (logistic regression, J48 decision trees, Naïve Bayes, random forest, etc.). As a result, an extension of the empirical ruleset was derived.

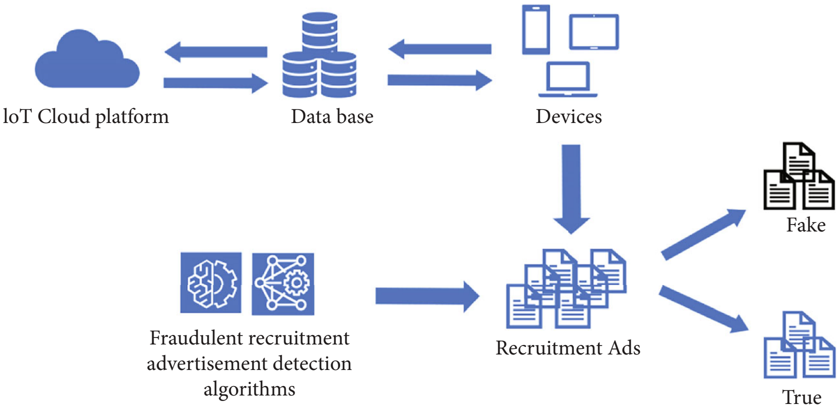


Figure 1: Detection process for fake job ads in an IoT environment.

Based on supervised learning methods, Dutta and Bandyopadhyay [9] proposed an automatic classification tool. They used single and ensemble classifiers to detect fraudulent recruitment advertisements, respectively. The single classifiers applied Naïve Bayes, multi-layer perceptron (MLP), K-nearest neighbor (KNN), and decision tree. And in the ensemble classifiers, random forest, and AdaBoost were applied. In addition, they compared the performance of these classifiers on the original highly unbalanced dataset.

Recruitment advertisements should be drawn from various sources to collect data in an all-around way. To solve this problem, Nindyati and Nugraha [10] extracted the Indonesian Employment Scam Dataset (IESD) from Indonesian recruitment data and manually labeled it based on empirical analysis and public reports of employment scam complaints. They considered the platforms where fraudulent recruitment advertisements were posted. In addition, they added behavioral features to previous studies [8] and used behavioral activities as contextual features for fraud detection. Naïve Bayes, logistic regression, KNN, decision tree, and SVM were applied as classifiers. The result indicated that adding behavioral characteristics can improve the detection effect of fake recruitment advertisements.

2.2. Traditional Learning + Feature Extraction-Based Detection Methods. According to the classification method proposed by Vidros et al. [8], the features were divided into three categories in the feature extraction stage, i.e., language-based, context-based, and metadata-based features [3]. They selected J48, logistic regression, and random forest methods as three baselines. Then, they combined voting techniques, including average vote, majority vote, and maximum vote, with these three baselines when constructing detection models. Moreover, they evaluated the performance of the models on unbalanced datasets, which made the experiment more comprehensive.

The lack of sufficient background information on recruitment websites makes the detection of fraudulent recruitment advertisements even more challenging. To address this problem, Mahbub and Pardede [11] focused on a novel feature space design that further extracted information about recruitment companies. They extracted contextual features manually and considered not only textual and structural features but also contextual features. The experiments result on the Naïve Bayes, J48, and JRip classifiers suggested that adding contextual features improved the detection performance and further enriched the ruleset.

Alghamdi and Alharby [12] used ensemble method based on random forest classifier to detect fraudulent recruitment advertisements. The SVM algorithm extracted the main features, including company profile, company logo, and required experience. Moreover, the detection performance was improved based on the reliable model obtained in the preprocessing and feature selection stages.

Mehboob and Malik [13] focused on the influential features of the EMSCAD to detect fake recruitment advertisements. They used information gain to select significant features. Their experimental results showed that the company profile, salary range, organization type, and required education were the most influential. Therefore, they considered combining well-performing features and adding valuable features to help improve the model's performance. In addition, they established a robust detection model using gradient boosting techniques.

2.3. Deep Learning-Based Methods. Fraudsters may know the ruleset in advance, which makes detecting fraudulent recruitment advertisements increasingly tricky. Kim et al. [14] believed there was an internal correlation between frauds, so they proposed a deep neural network algorithm based on hierarchical clustering to detect implicit fraud in data. They took the anomaly characteristics as the initial weights of the deep neural network and trained them with an autoencoder. This method discarded the feature information, took the global and local data structure as the entry point, and used clustering and deep neural networks to reveal the internal relationships between frauds.

2.4. Summary. Existing studies did an excellent job of detecting fraudulent recruitment advertisements. However, using the ruleset to detect fake job advertisements has poor expansion and is challenging to apply to new datasets. At the same time, existing studies used different detection methods, different evaluation metrics, and different datasets. They lacked comparative analysis, which makes the detection of fake job advertisements lack systematic research.

To solve this problem, this paper selects four traditional algorithms, random forest, logistic regression, SVM, and Naïve Bayes, which are frequently used in the above literature and are proven to be good on various datasets under multifarious evaluation metrics. Moreover, three popular deep learning algorithms, including GRU, Bi-LSTM, and TextCNN, are adopted to detect fake recruitment advertisements, all performing well in the field of NLP. For machine learning methods, the Bag of Words (BoW) algorithm and the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm [15] are adopted to realize feature extraction. At the same time, in deep learning methods, we also try to use pretraining models, including Word to Vector (Word2-Vec) [16] and Global Vectors (GloVe) [17] for word embedding. Moreover, we use the same evaluation metrics and datasets to carry out systematic and comprehensive comparative experiments.

3. Comparison and Analysis of Fraudulent Recruitment Advertisement Detection Methods

Since this paper is aimed at comparing and analyzing the detection methods used in the existing work experimentally, this section will analyze the four typical traditional learning methods frequently adopted in recruitment advertising detection literature [3, 8–10, 12, 13] and the three currently popular deep learning methods. In detail, the four traditional learning methods include random forest, logistic regression, SVM, and Naïve Bayes. The three deep learning methods include GRU, Bi-LSTM, and TextCNN.

3.1. Random Forest Method. As a classical ensemble learning method, random forest was first proposed by Breiman [18], who combined the Bagging ensemble learning theory [19] with the random subspace method [20]. Random forest is a classifier based on the decision tree, which can solve the performance bottleneck of the decision tree by outputting the result by voting. In addition, random forest has better tolerance to noise and outliers, has stronger scalability for high-dimensional data classification, and has a stronger generalization ability of the model. Moreover, random forest is a data-driven nonparametric classification method that only needs to learn classification rules from a given sample without prior knowledge.

3.2. SVM Method. Based on statistical learning theory [21], SVM is a data mining method that can successfully handle regression and pattern recognition problems. The target of SVM is to search for the optimal hyperplane in space to satisfy the classification requirements. SVM is one of the most commonly used and most effective classifiers, and it has good generalization performance based on the principle of structural risk minimization. Moreover, SVM has a solid theoretical basis and specific mathematical model and has been widely concerned by researchers since it was proposed.

3.3. Naïve Bayes Method. Based on the probability model, Naïve Bayes was proposed by Maron and Kuhns [22]. The "naive" of Naïve Bayes refers to the two primary hypotheses: conditional independence and positional independence. In detail, the conditional independence hypothesis assumes that the property values are independent of each other, namely, there is no dependency between terms. The positional independence hypothesis means that the term's location in the document does not affect the probability calculation.

3.4. Logistic Regression Method. The mechanism of logistic regression [23] uses a group of data to fit a logistic regression model and form multiple regression relationships to predict the occurrence probability of an event in any area. The advantage of logistic regression lies in that the independent variables in statistical analysis can be continuous or discrete, which does not need to meet the normal distribution.

Logistic regression is good at solving binary classification problems, and detecting fraudulent recruitment advertisements is a common textual binary classification problem. The logistic regression classifier is simple and easy to understand, and the model is highly interpretable. It does not need to assume data distribution in advance and directly models the possibility of classification, avoiding the problem of inaccurate hypothesis distribution. In addition, only the eigenvalues of each dimension are stored, and the memory resource consumption is small.

3.5. TextCNN Method. Convolutional neural network (CNN) [24] is a feedforward neural network that recognizes two-dimensional images with amplification, shrinkage, and displacement invariance. In recent years, CNN has been mostly used for image processing or classification recognition. Kim [25] first adopted CNN for text classification and proposed the TextCNN model. Figure 2 is the structure of TextCNN. TextCNN and CNN are very similar in design. The difference is that CNN uses convolution kernels of the same width and height when processing images. Still, TextCNN's convolution kernel width is consistent with the word vector dimension. When CNN processes images, it carries out a two-dimensional convolution operation at the convolution layer, while TextCNN carries out a one-dimensional convolution operation when it processes text.

3.6. LSTM Method. Long-short-term memory (LSTM) [26] is a special kind of recurrent neural network (RNN) [27]. Its network structure is the same as general RNN. The difference is that the memory module replaces the summation unit in the hidden layer. Figure 3 is the structure of the LSTM model. The information of cell state can be enhanced or weakened by the design of the "gate" of LSTM, so that long-term dependent information can be learned, effectively overcoming the defect of traditional RNN.

GRU [28] and Bi-LSTM [29] are the two most classical variants of LSTM. GRU was proposed to solve long-term memory problems and gradient in backpropagation. And Bi-LSTM overcomes the shortcoming that the LSTM model can only get one-way information from front to back but cannot from back to front. The forward and backward LSTM networks obtain the context information, and the model's performance is effectively improved.

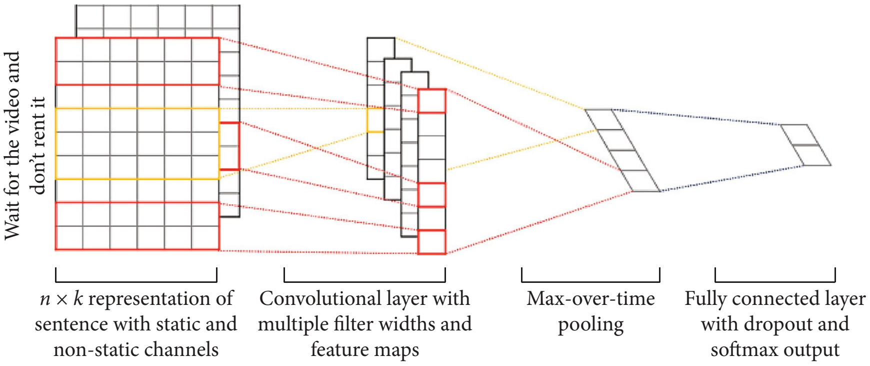

Wait for the video and don't rent it
$n \times k$ representation of sentence with static and non-static channels
Convolutional layer with multiple filter widths and feature maps
Max-over-time pooling
Fully connected layer with dropout and softmax output

Figure 2: Structure of TextCNN model.

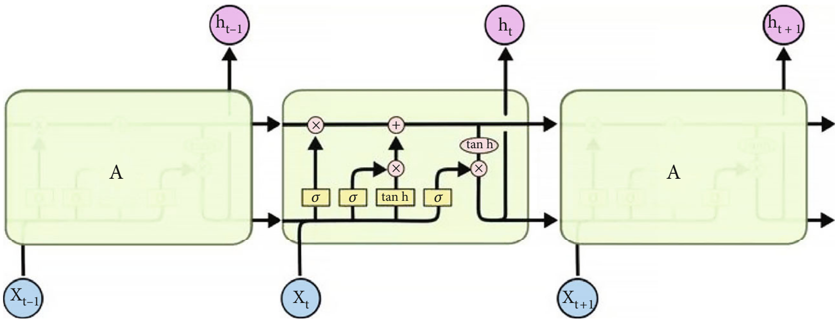


Figure 3: Structure of LSTM model.

Table 1: Dataset description.

Attribute number	Attribute name	Attribute type
1	Title	String
2	Location	String
3	Department	String
4	Salary_range	String
5	Company_profile	HTML fragment
6	Description	HTML fragment
7	Requirements	HTML fragment
8	Benefits	HTML fragment
9	Telecommuting	Binary
10	Has_company_logo	Binary
11	Has_questions	Binary
12	Employment_type	Nominal
13	Required_experience	Nominal
14	Required_education	Nominal
15	Industry	Nominal
16	Function	Nominal

4. Experiments and Results

In this paper, we have designed two sets of experiments. The first set of experiments is to verify the performance of each algorithm on the balanced dataset. The second set of experiments is conducted to verify the influence of the dataset imbalance ratio on the experimental performances. The following subsections illustrate the data and experimental setting details, the evaluation metrics, and the experimental results.

4.1. Data and Experimental Settings. In the experimental data section, the EMSCAD is used in this paper. The dataset contains 17880 real-life recruitment advertisements from 2012 to 2014, classified as legitimate and fraudulent, with 17014 legitimate and 866 fraudulent. The dataset description is shown in Table 1. We construct five datasets based on the original dataset using the downsampling method, including one balanced dataset and four unbalanced datasets. Table 2 is the detailed information on the datasets, and the last column is the datasets for each set of experiments we used. As many values are missed, we select company profile, description, requirements, and benefits, primarily selected in existing literature as text features. Before starting the experiment, we "clean" the experimental data, including removing punctuations, stopping words, and processing missing values.

In experiments, seven algorithms are selected for comparison, that is, the random forest method (abbr. RF), the logistic regression method (abbr. LR), the SVM method, and the Naïve Bayes method (abbr. NB), which are the four typical traditional learning algorithms adopted in recruitment advertisement detection literature [3, 8–12] that are described in the related work section. The TextCNN method, the GRU method, and the Bi-LSTM method are three promising deep learning algorithms. For feature

Table 2: Datasets statistics established based on the EMSCAD.

Datasets	True	Fraudulent	Fraudulent%	Total	Used experiments
Balanced dataset	866	866	50%	1732	1st
Unbalanced dataset-1	1300	866	40%	2166	2nd
Unbalanced dataset-2	2598	866	25%	2364	2nd
Unbalanced dataset-3	4908	866	15%	5774	2nd
Unbalanced dataset-4	17017	866	4.8%	17880	2nd

Table 3: Confusion matrix for binary classification.

Confusion matrix		Predict	
		Positive	Negative
True	Positive	TP	FN
	Negative	FP	TN

extraction, the BoW algorithm is used in the traditional learning methods, and the classical TF-IDF algorithm is adopted in traditional learning + feature extraction methods. Besides, deep learning methods use Word2Vec and GloVe for word embedding. All the experiments are validated by five-fold cross-validation.

4.2. Evaluation Metrics. Accuracy, precision, recall, and $F1$ -score are adopted as evaluation metrics. A confusion matrix is introduced first to introduce the above four evaluation metrics. Table 3 is a binary confusion matrix. In this table, TP represents the quantity of truly positive samples and classified as positive, FP represents the quantity of actually negative samples but classified as positive, FN represents the quantity of actually positive samples but classified as negative, and TN represents the quantity of actually negative samples and classified as negative.

The details of accuracy, precision, recall, and $F1$-score are shown as follows.

(1) *Accuracy* is the proportion of the number of correctly classified samples to the total samples

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

(2) *Precision* is the proportion of truly positive samples in a group of predicted positive samples

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

(3) *Recall* is the percentage of all truly positive samples that are successfully predicted

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

(4) *F1 -score* refers to the weighted summed average of precision and recall

$$F1\text{-score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

4.3. Experimental Results and Analysis

4.3.1. Results and Analysis of the First Set of Experiments. Table 4 lists the results of experiments on the balanced dataset. Table 4 shows that all the traditional learning methods achieve good results (A, P, R, and F are all greater than 0.88), and random forest performs best on the balanced dataset. After using the TF-IDF algorithm for feature extraction, the results of all the traditional learning + feature extraction methods are improved (A, P, R, and F are all greater than 0.9), except for the Naive Bayes. In particular, the SVM+TF-IDF methods achieve the best performance. TF-IDF algorithm measures the importance of a word in terms of frequency. These results show that using TF-IDF for feature extraction plays a significant part in enhancing the effectiveness of methods. In the deep learning methods, TextCNN performs best (A, P, R, and F are all greater than 0.93). The reason is that TextCNN uses three different sizes of convolution kernels for word embedding. Compared with random forest and SVM +TF-IDF, the performance of TextCNN improves (A: 3.1%, P: 3.2%, R: 3%, F: 3%) and (A: 0.2%, P: 0.1%, R: 0.3%, F: 0.2%), respectively. After using Word2Vec and GloVe for word embedding, the performance of GRU, Bi-LSTM, and TextCNN are worse than before. In my opinion, these two pretraining models are trained on a specific dataset, which is different from the writing style of the EMSCAD dataset, so the experimental results are not ideal.

4.3.2. Results and Analysis of the Second Set of Experiments. Table 5 lists the results of experiments on the four unbalanced datasets. From Table 5, when the datasets are slightly unbalanced, for example, on the unbalanced dataset-1 and unbalanced dataset-2, the results are roughly the same as those on the balanced dataset. In detail, on the unbalanced dataset-1 and unbalanced dataset-2, we can see that random forest performs best in traditional learning methods, while SVM+TF-IDF performs best in traditional learning + feature extraction methods. Similarly, TextCNN in the deep learning category

Table 4: Experimental results on the balanced dataset.

Category	Model	Balanced dataset			
		A	*P*	*R*	*F*
Traditional learning	RF	0.899	0.899	0.900	0.900
	SVM	0.884	0.885	0.885	0.884
	LR	0.897	0.898	0.899	0.897
	NB	0.893	0.894	0.894	0.893
Traditional learning + feature extraction	RF+TF-IDF	0.923	0.923	0.923	0.923
	SVM+TF-IDF	0.928	0.930	0.927	0.928
	LR+TF-IDF	0.908	0.908	0.909	0.908
	NB+TF-IDF	0.876	0.879	0.879	0.876
Deep learning	GRU	0.835	0.836	0.835	0.834
	Bi-LSTM	0.884	0.886	0.884	0.884
	TextCNN	**0.930**	**0.931**	**0.930**	**0.930**
Deep learning + pretraining	GRU+Word2Vec	0.748	0.750	0.745	0.747
	GRU+GloVe	0.748	0.748	0.748	0.748
	Bi-LSTM + Word2Vec	0.875	0.875	0.876	0.875
	Bi-LSTM + GloVe	0.831	0.832	0.830	0.830
	TextCNN + Word2Vec	0.923	0.923	0.923	0.923
	TextCNN + GloVe	0.926	0.927	0.927	0.927

performs best compared to the above methods. On the unbalanced dataset-1, compared with random forest and SVM+TF-IDF, the performance of TextCNN improves (*A*: 3.1%, *P*: 2.3%, *R*: 3.8%, *F*: 3.3%) and (*A*: 3%, *P*: 2.7%, *R*: 3.3%, *F*: 3.2%), respectively. On the unbalanced dataset-2, compared with random forest and SVM+TF-IDF, the performance of TextCNN improves (*A*: 3.2%, *P*: 1.2%, *R*: 6.8%, *F*: 4.8%) and (*A*: 2.4%, *P*: 0.8%, *R*: 5%, *F*: 3.5%), respectively. In the deep learning + pretraining methods, after using the pretraining model Word2Vec and GloVe for word embedding, the performance of GRU and Bi-LSTM decreased significantly. And the result of the TextCNN methods decreases slightly. We guess these results also may lead by the fact that these two pretraining models are trained on a specific dataset, which is different from the writing style of the EMSCAD dataset.

On the unbalanced dataset-3 and unbalanced dataset-4, though the unbalanced ratio of these two datasets increases, the TextCNN method maintains its advantage. Still, it has the best detection effect compared with other methods. In particular, on the unbalanced dataset-3 and unbalanced dataset-4, TextCNN performs best compared to the other methods. As the datasets become more unbalanced, the influence of TF-IDF on accuracy, precision, recall, and $F1$ -score gradually decreases, and the recall and $F1$-score of the traditional learning methods and traditional learning + feature extraction methods significantly reduce, which shows that the TextCNN method has good robustness even when the dataset is very unbalanced. We think the TextCNN method uses multiple convolution kernels of different sizes to embed documents that enrich the semantic representation. In conclusion, TextCNN has advantages in dealing with balanced and unbalanced datasets; therefore, it is more suitable for dealing with the data in real life.

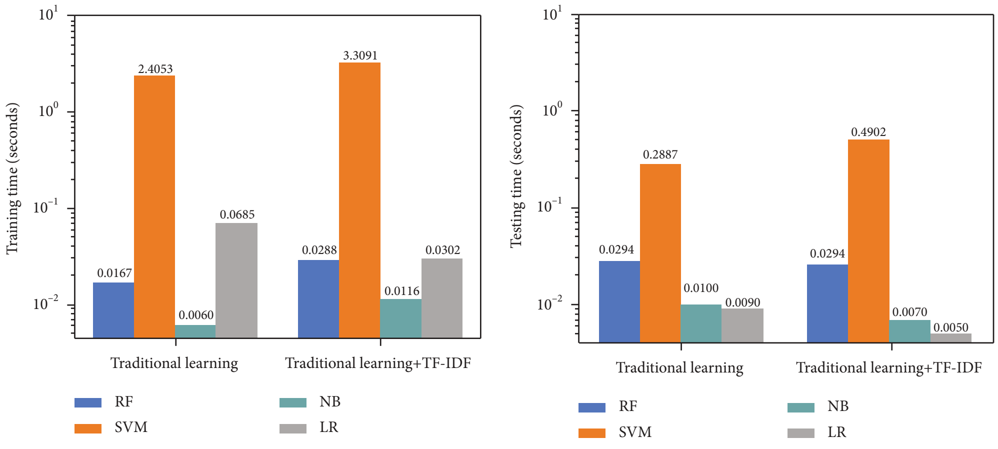
4.4. Analysis of Energy Consumption. Since the goal of the green IoT is to achieve better results in an environmentally friendly manner (i.e., less computing consumption), for the computing consumption, two aspects should be considered, i.e., training time and testing time. Testing time is even more critical for IoT devices than training time for getting a good model. This is because IoT devices with less testing time are more sensitive and can contribute to better human-computer interaction. Therefore, to further compare the performance of the compared methods, this section analyzes the experimental results from the perspective of training time and testing time and takes the testing time as the primary consideration. Figures 4 and 5 are the results of training time and testing time on the balanced dataset. Two figures are drawn because the training time of traditional and deep learning methods is of different orders of magnitude.

As can be seen from Figure 4(a), among the four traditional learning methods (see left columnar plexus of Figure 4(a)), Naïve Bayes has the shortest training time and SVM has the longest one. After feature extraction (see right columnar plexus of Figure 4(a)), expecting for logistic regression, the training time of all the other methods, such as RF, SVM, and NB, increases. Compared with traditional learning methods, the training time of the detection methods based on deep learning is significantly increased to a larger order of magnitude, shown in Figure 5(a). That is because deep learning methods need much longer to train the deep neural network. In addition to the GRU method, the training time of Bi-LSTM and TextCNN increases after using pretraining models for word embedding (see right columnar plexus of Figure 5(a)), and the training time of the TextCNN method increases most significantly.

Table 5: Experimental results on the unbalanced datasets.

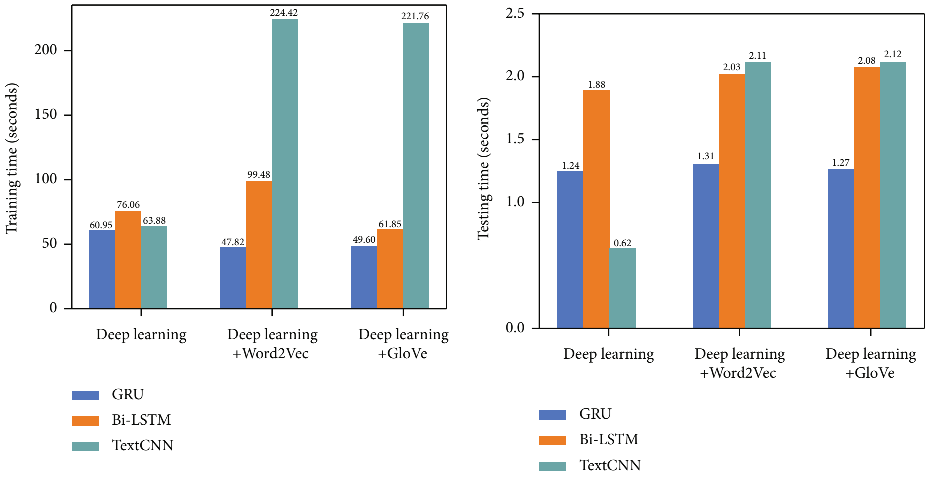
Category	Model	Unbalanced dataset-1				Unbalanced dataset-2			
		A	*P*	*R*	*F*	*A*	*P*	*R*	*F*
Traditional learning	RF	0.910	0.919	0.896	0.905	0.928	0.946	0.864	0.896
	SVM	0.896	0.898	0.885	0.890	0.924	0.932	0.865	0.892
	LR	0.877	0.872	0.874	0.873	0.906	0.874	0.880	0.877
	NB	0.830	0.824	0.830	0.826	0.832	0.781	0.826	0.796
Traditional learning + feature extraction	RF+TF-IDF	0.904	0.912	0.890	0.898	0.918	0.940	0.844	0.879
	SVM+TF-IDF	0.911	0.915	0.901	0.906	0.936	0.950	0.882	0.909
	LR+TF-IDF	0.882	0.885	0.870	0.877	0.904	0.912	0.830	0.860
	NB+TF-IDF	0.836	0.834	0.824	0.828	0.861	0.855	0.761	0.791
Deep learning	GRU	0.865	0.861	0.857	0.857	0.912	0.893	0.868	0.879
	Bi-LSTM	0.886	0.881	0.881	0.880	0.915	0.894	0.875	0.883
	TextCNN	**0.941**	**0.942**	**0.934**	**0.938**	**0.960**	0.958	**0.932**	**0.944**
Deep learning + pretraining	GRU+Word2Vec	0.802	0.796	0.808	0.798	0.855	0.820	0.783	0.800
	GRU+GloVe	0.808	0.804	0.797	0.799	0.859	0.814	0.787	0.799
	Bi-LSTM + Word2Vec	0.889	0.884	0.884	0.884	0.918	0.896	0.885	0.889
	Bi-LSTM + GloVe	0.854	0.848	0.855	0.851	0.901	0.895	0.825	0.853
	TextCNN + Word2Vec	0.938	0.939	0.932	0.936	0.955	**0.960**	0.920	0.938
	TextCNN + GloVe	0.932	0.937	0.922	0.928	0.949	0.958	0.906	0.928
Category	Model	Unbalanced dataset-3				Unbalanced dataset-4			
		A	*P*	*R*	*F*	*A*	*P*	*R*	*F*
Traditional learning	RF	0.944	**0.963**	0.824	0.875	0.977	0.985	0.768	0.842
	SVM	0.940	0.941	0.821	0.867	0.974	0.984	0.733	0.811
	LR	0.937	0.879	0.881	0.880	0.969	0.830	0.832	0.831
	NB	0.853	0.738	0.840	0.769	0.868	0.613	0.827	0.646
Traditional learning + feature extraction	RF+TF-IDF	0.939	0.962	0.806	0.861	0.976	**0.987**	0.755	0.831
	SVM+TF-IDF	0.954	0.960	0.861	0.901	0.980	0.984	0.799	0.867
	LR+TF-IDF	0.923	0.920	0.771	0.822	0.965	0.938	0.656	0.724
	NB+TF-IDF	0.890	0.891	0.657	0.705	0.955	0.961	0.541	0.564
Deep learning	GRU	0.947	0.918	0.861	0.885	0.973	0.875	0.833	0.850
	Bi-LSTM	0.946	0.920	0.862	0.888	0.979	0.922	0.837	0.874
	TextCNN	**0.969**	0.956	**0.923**	**0.938**	**0.986**	0.954	**0.896**	**0.922**
Deep learning + pretraining	GRU+Word2Vec	0.916	0.855	0.781	0.809	0.967	0.895	0.733	0.791
	GRU+GloVe	0.923	0.867	0.791	0.791	0.969	0.885	0.775	0.820
	Bi-LSTM+Word2Vec	0.947	0.908	0.869	0.887	0.973	0.905	0.802	0.845
	Bi-LSTM+GloVe	0.932	0.902	0.807	0.845	0.972	0.920	0.758	0.817
	TextCNN+Word2Vec	0.968	0.967	0.900	0.929	0.985	0.981	0.860	0.911
	TextCNN+GloVe	0.949	0.958	0.906	0.928	0.962	0.963	0.879	0.915

For the testing time, according to Figure 4(b), we can see that almost all the methods need less than 0.05 s in response to each test except the SVM and SVM+TF-IDF methods. I think SVM and SVM+TF-IDF spend more time searching for the optimal hyperplane. In particular, the logistic regression method has the shortest testing time, which means it is the most sensitive method for IoT devices. After using TF-IDF for feature extraction (see right columnar plexus of Figure 4(b)), the testing time of all the other two methods, such as NB and LR, is decreased except for the RF and SVM methods. From Figure 5(b), we can see that in the deep learning methods, the testing time of TextCNN is shorter than GRU and Bi-LSTM. With the adoption of Word2Vec and GloVe, the testing time of GRU, Bi-LSTM, and TextCNN all increased. It is worth noting that the response time of TextCNN is about 0.6 s per test, which is slightly higher than that of traditional learning methods. Compared with the results, we get by using the traditional learning methods and the traditional learning + feature extraction methods in Table 4 and Table 5, we can see that TextCNN achieves the best detection performance on accuracy, precision, recall, and $F1$-score. Moreover, comparing the results



(a) (b)

Figure 4: (a) Training time and (b) testing time on the four traditional learning methods.



(a) (b)

Figure 5: (a) Training time and (b) testing time on the two deep learning methods.

in Figures 4 and 5, we can see that TextCNN achieves acceptable energy consumption during testing time. Therefore, TextCNN is a worthwhile method for detecting fraudulent recruitment advertisements in the IoT Environment.

5. Conclusion

This paper analyzes and compares seventeen methods, four traditional learning methods (i.e., random forest (RF), SVM, logistic regression (LR), and Naïve Bayes (NB)). These four traditional learning methods combined with feature extracting method (i.e., RF+TF-IDF, SVM+ TF-IDF, LR +TF-IDF, and NB+ TF-IDF) and three deep learning methods (i.e., GRU, Bi-LSTM, and TextCNN); these three deep learning methods combined with pretraining model for word embedding (i.e., GRU+Word2Vec, GRU+GloVe, Bi-LSTM+Word2Vec, Bi-LSTM+GloVe, TextCNN+Word2Vec, and TextCNN+GloVe). To further analyze the performance of each method, comprehensive experiments are carried out based on the EMSCAD dataset.

The experimental results indicate that the deep learning methods are generally better than the traditional learning methods, the traditional learning + feature extraction methods, and even the deep learning + pretraining-based methods, regardless of the balanced or the unbalanced datasets. In particular, TextCNN outperforms other deep learning methods. In terms of time performance, though TextCNN needs considerably longer offline training time, the testing time (i.e., response time) is slightly higher than the traditional learning-based methods. Those results indicate that the TextCNN method can detect real-life fraudulent recruitment advertisements in the IoT environment.

In summary, using unified evaluation metrics and datasets and considering the impact of the imbalance rates make the comparison and analysis of fraudulent recruitment advertisement detection methods more systematic and comprehensive.

Therefore, this paper can help researchers systematically understand the detecting methods of fraudulent recruitment advertisements and provide directions for selecting and exploring suitable methods. The experimental results have a specific reference value for further research of higher performance recruitment advertising detection methods.

Based on this paper, in the future, we aim to collect our employment fraud detecting dataset, and we will study a higher performance and lower energy consumption fraudulent recruitment advertising detection method to help achieve the goal of green IoT. It is an exciting direction to ensemble high-performance methods such as TextCNN, LSTM, or other popular deep learning methods and technologies (e.g., attention mechanism, mask mechanism). Moreover, deeper pretraining models for word embeddings are also an interesting direction.

Data Availability

The data we used is available and can be obtained from the author (202082060057@sdust.edu.cn).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper is supported by the Natural Science Foundation of China (71772107), the Natural Science Foundation of Shandong Province of China (ZR2020MF044, ZR202102230289, and ZR2019MF003), Shandong Education Quality Improvement Plan for Postgraduate (2021), and Open Research Fund of Anhui Province Engineering Laboratory for Big Data Analysis and Early Warning Technology of Coal Mine Safety (no. CSBD2022-ZD01).

References

[1] P. Kaur, “E-recruitment: a conceptual study,” *International Journal of Applied Research*, vol. 1, no. 8, pp. 78–82, 2015.

[2] C. Wang, “False information analysis of online recruitment,” *Modern Marketing (Business Edition)*, vol. No. 335, no. 11, pp. 140-141, 2020.

[3] S. Lal, R. Jiaswal, N. Sardana, A. Verma, A. Kaur, and R. Mourya, “ORFDetector: ensemble learning based online recruitment fraud detection,” in *2019 Twelfth International Conference on Contemporary Computing (IC3)*, pp. 1–5, Noida, India, 2019, August.

[4] P. Rubin, “Regulation of information and advertising,” *CPI Journal*, vol. 4, 2008.

[5] Y. Jiang, “On the construction of advertising credit supervision system,” *Industrial and Commercial Administration*, vol. 4, 2004.

[6] S. Vidros, C. Kolias, and G. Kambourakis, “Online recruitment services: another playground for fraudsters,” *Computer Fraud & Security*, vol. 2016, no. 3, pp. 8–13, 2016.

[7] I. Androutsopoulos, J. Koutsias, K. V. Chandrinos, and C. D. Spyropoulos, “An experimental comparison of naive Bayesian and keyword-based anti-spam filtering with personal e-mail messages,” in *Proceedings of the 23rd annual international ACM SIGIR conference on research and development in information retrieval*, pp. 160–167, Athens, Greece, 2000, July.

[8] S. Vidros, C. Kolias, G. Kambourakis, and L. Akoglu, “Automatic detection of online recruitment frauds: characteristics, methods, and a public dataset,” *Future Internet*, vol. 9, no. 1, p. 6, 2017.

[9] S. Dutta and S. K. Bandyopadhyay, “Fake job recruitment detection using machine learning approach,” *International Journal of Engineering Trends and Technology*, vol. 68, no. 4, pp. 48–53, 2020.

[10] O. Nindyati and I. G. B. B. Nugraha, “Detecting scam in online job vacancy using behavioral features extraction,” in *2019 International Conference on ICT for Smart Society (ICISS)*, vol. 7, pp. 1–4, Bandung, Indonesia, 2019, November.

[11] S. Mahbub and E. Pardede, “Using contextual features for online recruitment fraud detection,” in *the 27th International Conference on Information Systems Development*, Lund, Sweden, 2018.

[12] B. Alghamdi and F. Alharby, “An intelligent model for online recruitment fraud detection,” *Journal of Information Security*, vol. 10, no. 3, pp. 155–176, 2019.

[13] A. Mehboob and M. S. I. Malik, “Smart fraud detection framework for job recruitments,” *Arabian Journal for Science and Engineering*, vol. 46, no. 4, pp. 3067–3078, 2021.

[14] J. Kim, H. J. Kim, and H. Kim, “Fraud detection for job placement using hierarchical clusters-based deep neural networks,” *Applied Intelligence*, vol. 49, no. 8, pp. 2842–2861, 2019.

[15] G. Salton and C. T. Yu, “On the construction of effective vocabularies for information retrieval,” *ACM SIGPLAN Notices*, vol. 10, no. 1, pp. 48–60, 1975.

[16] T. Mikolov, G. Corrado, C. Kai, and J. Dean, “Efficient estimation of word representations in vector space,” in *Proceedings of the international conference on learning representations (ICLR 2013)*, Scottsdale, AZ, USA, 2013.

[17] J. Pennington, R. Socher, and C. Manning, “Glove: global vectors for word representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014.

[18] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[19] S. W. Kwok and C. Carter, “Multiple decision trees,” in *Machine Intelligence and Pattern Recognition*, vol. 9, pp. 327–335, North-Holland, 1990.

[20] T. K. Ho, “The random subspace method for constructing decision forests,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[21] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, 2013.

[22] M. E. Maron and J. L. Kuhns, “On relevance, probabilistic indexing and information retrieval,” *Journal of the ACM (JACM)*, vol. 7, no. 3, pp. 216–244, 1960.

[23] B. Efron, “The efficiency of logistic regression compared to normal discriminant analysis,” *Journal of the American Statistical Association*, vol. 70, no. 352, pp. 892–898, 1975.

[24] Y. LeCun, B. Boser, J. S. Denker et al., “Backpropagation applied to handwritten zip code recognition,” *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[25] Y. Kim, “Convolutional neural networks for sentence classification,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014.

[26] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[27] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[28] K. Cho, B. Van Merriënboer, C. Gulcehre et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014.

[29] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.

Hindawi
Journal of Sensors
Volume 2022, Article ID 2022586, 11 pages
https://doi.org/10.1155/2022/2022586

Research Article

Adaptive VR Video Data Transmission Method Using Mobile Edge Computing Based on AIoT Cloud VR

Min Wang,[1] **Fuquan Zhang**,[2,3] **Linjuan Ma**,[4] **and Ye Tian**[5]

[1]*Dean's office, Fujian Chuanzheng Communications College, Fuzhou, China 350007*
[2]*College of Computer and Control Engineering, Minjiang University, Fuzhou, China 350108*
[3]*Digital Media Art, Key Laboratory of Sichuan Province, Sichuan Conservatory of Music, Chengdu, China 610021*
[4]*School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China 100081*
[5]*Experimantal Art School, Sichuan Conservatory of Music, Chengdu, China 610041*

Correspondence should be addressed to Fuquan Zhang; zfq@mju.edu.cn and Linjuan Ma; malinjuan@bit.edu.cn

Received 24 May 2022; Accepted 29 July 2022; Published 7 September 2022

Academic Editor: Saru Kumari

Aiming at the high requirements of cloud service-based virtual reality in AIoT for data transmission rate and delay sensitivity, a cloud VR system scheme based on MEC (Mobile Edge Computing) is proposed, which mainly incorporates viewpoint-based VR video data processing and hybrid digital-to-analog (HDA) transmission optimization and can be served for AIoT transmission filed. Firstly, a learning-driven multiaccess MEC offloading strategy is designed, in which the VR terminal automatically selects the optimal MEC server for task offloading, thereby effectively improving network efficiency and reducing service delay. Secondly, the progressive transmission of the VR data is realized through viewpoint-aware dynamic streaming based on RoI (region of interest) and the priorities of different objects. The transmission priority of each object in the scene is determined through the ROI layering, which effectively solves the contradiction between the large data volume in the VR scenes and the network bandwidth limitation when applied in AIoT domain, and further improves the real-time performance of the system. Then, the HDA (hybrid digital-analog) technique is introduced to optimize the transmission. Finally, the base station protocol stack is modified on the basis of the LTE (Long-Term Evolution) system, and the MEC technology is integrated to realize a complete cloud VR system in AIoT. The experimental results show that compared with other advanced schemes, the proposed scheme can achieve more robust and efficient data transmission performance and provide better VR user experience.

1. Introduction

Recently, it is well known that VR (virtual reality) systems combined with AIoT, incorporating multiple technologies such as digital image processing, computer graphics, multimedia, computer simulation, sensors, and computer networks, have a wide range of applications in many fields such as entertainment, simulation training, aerospace, scientific and computing visualization, art, [1]. VR has three most prominent features: immersion, interactivity, and imagination, which are also known as the 3I characteristics of VR [2] and can realize human-computer interaction based on these characteristics. Cloud VR is a kind of real-time VR technology where can be used for amount data transmission in AIoT based on cloud computing, in which cloud servers are used to replace users' local computing devices, which promotes the popularization of VR applications [3]. The cloud VR architecture consists of four parts: the content layer, the platform layer, the network layer, and the terminal layer. Among them, the content provided by the platform layer to the content layer includes cloud VR video services and cloud VR strong interaction services and is responsible for video import, transcoding, storage, broadcast control and distribution processing, logical calculation, and real-time rendering for strong interaction services. The network layer consists of four parts: backbone network, MAN (metropolitan area network), access network, and home network, which meet cloud VR's requirements for large bandwidth and low latency. Finally, the terminal layer is connected to the platform layer by accessing 5G/Wi-Fi to realize functions such as VR content presentation and user authentication [4]. Cloud VR Solution Architecture is shown in Figure 1.

However, due to the huge amount of VR video data, in addition to cloud computing and rendering, network transmission bandwidth and delay limitations have become new bottlenecks for the entire system [5]. For basic 4 K resolution cloud VR services, the network bandwidth needs to reach at least 40 Mbit/s, and the RTT (round trip delay) of transmission should be controlled within 40 ms to provide users with a good viewing experience [6]. In the current mobile network architecture, the distance between the user and the server is at least at the metro distance level. Regardless of device forwarding and image transmission, the RTT of fiber transmission only is as high as 30-40 ms, which is difficult to meet the requirements of cloud VR [7].

With the development of 5G technology and the need of AIoT, the bandwidth of mobile networks has greatly increased. By sinking computing nodes to the vicinity of the application scenario (gateways), MEC puts the data collection and analysis operations close to the user side and realizes the closed loop of data processing on the edge, which can alleviate the network transmission pressure and shorten the data processing time [8]. Under the MEC architecture, deploying VR applications on MEC nodes will bring the following advantages: (1) the MEC computing nodes are directly deployed close to the mobile gateways, which can reduce the number of hops for network transmission between VR application data and end users, and reduce network processing delay [9]; (2) the VR application runs on the MEC node, which can realize local data processing. For the users under the same UPF (User Plane Function), the data does not need to enter the Internet, which can reduce the transmission pressure on the Internet [10]; (3) edge processing of VR content, for example, in VR live broadcast scenarios, the process of VR video splicing, encoding and transcoding, and distribution can be performed directly on the MEC nodes, data offloading can be realized for local users nearby, and OTT (Over the Top) users or users covered by other MEC nodes can be quickly distributed through CDN (content delivery network) [11]; and (4) the distributed networking of MEC can realize continuous VR experience in mobile scenarios, such as watching VR live broadcasts or participating in video conferences on high-speed mobile carriers (such as cars and high-speed trains) [12].

In addition, MEC can work with low-latency application-layer protocols such as QUIC (Quick User Datagram Protocol Internet Connection) and RTP (Real-Time Transport Protocol) to make cloud VR possible [13]. On the other hand, in the MEC scenario, the source ends (edge servers) are more closely connected with the channel ends (base stations), and the bandwidths are sufficient to support the transmission of baseband data between the servers and the base stations, which greatly improves the feasibility of using pseudoanalog, HAD (hybrid digital-analog) and other source-channel joint coding techniques [14].

When a user watches a VR video, due to the limitation of the FoV (field of view) of the display device, the user often can only watch a certain part of the whole video at a moment. If the server transmits the entire video, most of the bandwidth resources will inevitably be wasted. Therefore, the adaptive block transmission method based on DASH (Dynamic Adaptive Streaming over HTTP) is the most widely used method for VR video streaming [15]. In the block transmission mode, a complete VR video is divided into many video blocks, and each video block is encoded into different quality levels. The server adaptively selects an optimal quality level for each video block according to factors such as network bandwidth and transmits it to the user. There are two basic schemes in the adaptive block transmission mode: view adaptation and rate adaptation [16]. The former is central for predicting changes in users' viewpoints, and the latter is central to resource allocation by the servers.

Based on the mobile edge computing technology in AIoT, this paper modified the LTE (Long-Term Evolution) base station protocol stack to build a mobile edge computing platform to serve for AIoT that expands to support HDA (hybrid digital-analog) transmission, in order to realize an efficient and reliable cloud VR system. The main contributions of this paper are as follows:

(1) A learning-driven MEC server unloading strategy is adopted, so that users can automatically select the optimal MEC server

(2) Realize a complete cloud VR system through viewpoint-aware dynamic streaming based on ROI and object priorities

(3) Based on HDA technique, the system transmission efficiency is optimized to provide high quality VR videos under limited bandwidth

2. Related Research

The main content in the VR system is virtual scenes, and there will be a large number of 3D scenes in the virtual scenes. In the current network environment, the system is faced with the problem of how to solve the contradiction between the data transmission of 3D scenes and the limited network bandwidth in practice in AIoT. This aspect involves the processing, sending, and receiving of the 3D scenes, and at the same time, it is also necessary to ensure that the user-side scenes can be generated with a good visual experience [17]. Whether these problems can be solved are fundamentally related to the successful implementation of the VR system.

The researchers have proposed FOV transmission schemes for differential transmission of panoramic video information based on viewpoint areas [18]. These include the pyramid projection transmission scheme proposed by Facebook and the Tile Wise transmission scheme promoted by Huawei.

In the pyramid projection transmission scheme, a full-view nonuniform quality code stream is prepared for each view, and high-quality coding is used in the user's viewpoint region, while low-quality coding is used in other regions [19]. This method greatly reduces the bandwidth

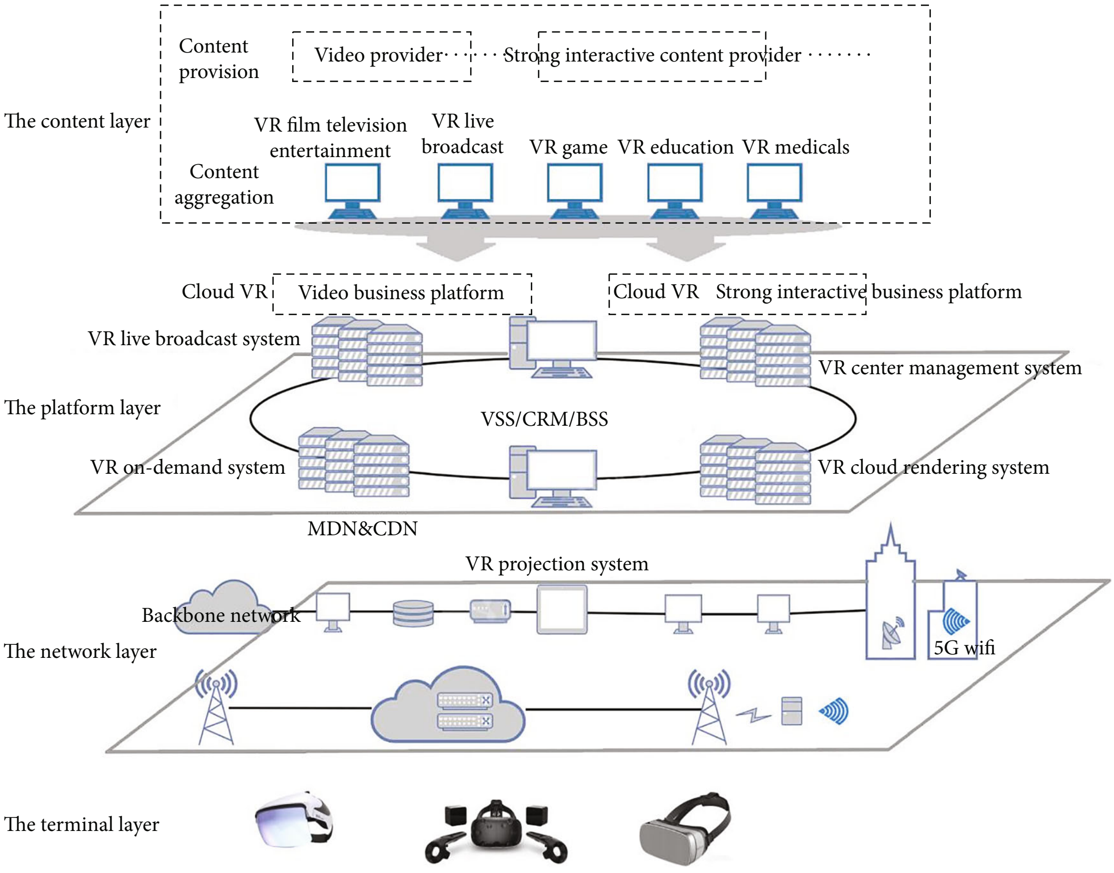


Figure 1: Cloud VR Solution Architecture.

requirements of system users when watching panoramic videos and improves the effective utilization of network bandwidth, but the sum of all perspective video files in the system server is more than 6 times that of the original files. The Tile Wise transmission scheme combines low-quality full-views and high-quality viewpoint regions. The server side does not need to prepare for each viewing angle area but divides the panoramic video image into multiple tiles at the same time, each area corresponds to a stream that can be decoded independently, and the server will prepare a low-quality panoramic full-view video stream. The client obtains a full-view stream and a high-quality tile selected according to the viewpoint information [20]. For the construction of panoramic images with nonuniform qualities, Hosseini et al. [21] proposed a viewpoint-aware adaptive VR transmission framework based on extended MPEG-DASH SRD (moving picture experts group-dynamic adaptive streaming over HTTP spatial relation description). Similarly, Kim et al. [22] proposed a SSAS (spatial segmented adaptive streaming) scheme based on the HLS (HTTP live streaming) protocol to realize real-time adaptive streaming based on user viewpoints. These solutions draw on existing HTTP adaptive transmission protocols such as DASH and further expand time-based tiling to space to achieve dynamic adaptive streaming.

According to the situational information such as the computing capability of the users' mobile terminals, appropriate MEC servers are selected for efficient task offloading, so as to ensure the network delay performance and reduce energy consumption. Guo et al. [23] proposed a MEC task offloading strategy based on nonorthogonal multiple access, which takes into account the constraints of different access technologies. By considering different business quality of service (QoS) constraints, Henri et al. [24] proposed an offloading strategy that can guarantee a strong delay boundary based on game theory. Based on the Stackelberg game theory Hosseini et al. [25] proposed a price-based distributed MEC task offloading algorithm, which enables users to make autonomous decisions. In addition, Liu and Liu [26] proposed an energy-efficient MEC task offloading algorithm for ultradense wireless network scenarios in which energy overheads are minimized by optimizing offloading decision variables and power bandwidth allocation. In the existing research on MEC task offloading, it is assumed that the computing power and storage capacity of the MEC are known, and based on the research scenario of a single MEC server, the offloading decision of computing tasks is made with the goal of optimal delay or optimal energy. However, with the densification of base station deployment in 5G networks, a large number of MEC servers will be deployed on base stations or access points (APs) that are closer to user mobile terminals. The computing and storage capabilities of different MEC servers are different. Therefore, the mobile terminals on the user side need to independently decide and select the optimal MEC server access strategy according to

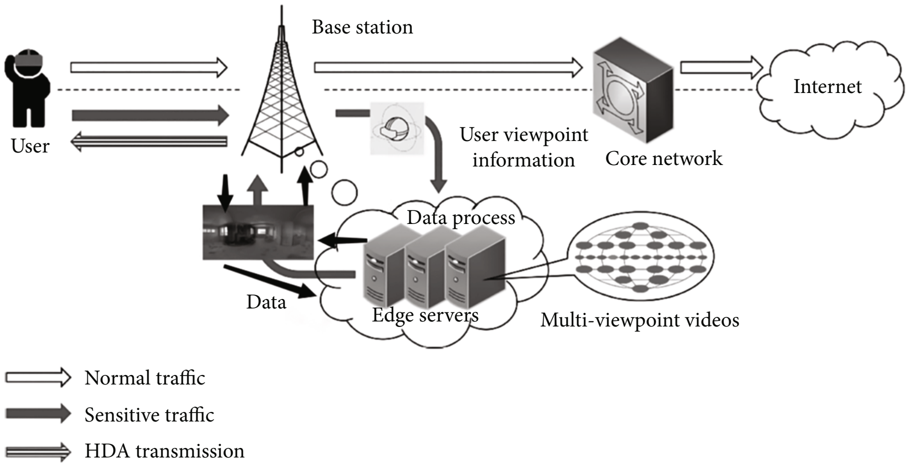


Figure 2: System structure.

the situational information such as service characteristics and network environment, so as to minimize network delay and network energy consumption at the same time, thereby realizing an energy-efficient MEC server task offloading strategy.

Liu et al. [27] proposed an efficient VR transmission mechanism based on source-channel joint coding. After tiling the VR videos with reference to the users' FOV information, different levels of error correction strategies are used to maximize the viewing quality within users' FOV. Feng et al. [28] defined a new QoE (Quality of Experience) metric to measure the user's viewing experience and presented an efficient modulation control algorithm to maximize the QoE value under different channel conditions. Zhang and Ma [32] proposed multiobject crowd real-time tracking in the dynamic environment based on a novel neural network, which can be used in AIoT cloud VR.

3. VR Adaptive Transmission Scheme Based on MEC and Viewpoint Awareness

3.1. System Architecture. The proposed solution integrates the MEC technology on the basis of the LTE system and expands the base station to realize the HDA transmission mode, so as to meet the high requirements of the interactive VR services for latency and network quality and provide high-performance network support for the cloud VR service especially in AIoT data transmission [32]. The system structure is shown in Figure 2.

By modifying the protocol stack of the base station, while introducing the MEC function, the compatibility of the system to standard LTE terminals is maintained. The tunneling protocols are used to redirect traffic at the network layer to filter and offload sensitive traffic from the edge services. The base station maintains a sensitive traffic table of the edge services, which records the IP addresses, protocols, and port numbers of data packets that need to be forwarded to the edge servers. Each passing data packet is matched. If the data packet matches the entry in the sensitive table, the destination IP in the tunnel packet header of the GPRS Tunneling Protocol (GTP) is reconstructed, and the original core network IP is replaced with the edge server IP; that is, the data packet is forwarded to the edge server. For the returned downlink data, the edge server masquerades the source IP address as the real public network address of the application server.

The proposed architecture implements cloud VR based on the MEC system. The computing tasks and services are moved down to the edge of the base stations, so as to minimize the transmission delays from both the network structures and the physical distances, and the RTT will be controlled within 10 ms. It not only improves the response speed of the server but also ensures the stability of the network service quality and greatly improves the users' viewing experience. The introduction of HDA technique provides a more flexible transmission mode for edge servers, enabling them to make full use of bandwidth resources and alleviating the saturation effect of existing digital transmission in AIoT.

3.2. Learning-Driven MEC Server Adaptive Offloading Strategy. In the MAB (multiarmed bandit) model [24], there are N gambling arms and one player for multiple rounds of selection. Each time the player selects one of the gambling arms and receives the corresponding reward, the player can only obtain the reward value of the selected arm after selection. The reward value of each gambling arm follows some unknown specific distribution and is independent of each other. The player learns the reward distribution of different gambling arms through exploration and utilization. After J rounds of games, the optimization goal of the player is to maximize the expected value of the reward.

It is assumed that there are U users and M base stations in the 5G wireless network scenario, and each base station contains an MEC server (to simplify the description, the base station and the MEC server are collectively represented by M). Let the total system bandwidth be B, and there are K subcarriers in the system bandwidth. Assuming that a user can only access one base station at time t, and at most, one user can access a subcarrier, then we have

$$\sum_{k\in K, m\in M} a^{k,j,m}(t) = 1, \forall i = U, \tag{1}$$

where $k \in K$ denotes a resource block. The SINR (signal to interference plus noise ratio) of user terminal i and base station m on resource block k is

$$\Gamma_i^{k,m} = \frac{P_{im}^{km} g^{k,i,m}}{N_0}, \tag{2}$$

where P_{im}^{km} represents the transmission power from base station m to user i in resource block k, the channel gain between base station m and user i is $g^{k,j,m}$, and N_0 is the noise power that follows $N(0, \delta)$ distribution. The transmission rate from the user to the base station is

$$R_i^m = \sum_{k_m=1}^{K_m} a_{i_m}^{k_m} \log\left(1 + \Gamma_i^{k,m}\right), k_m \in K_M^B. \tag{3}$$

For delay-sensitive services, it is assumed that the arrival rate of the data packets conforms to the Poisson distribution with the arrival rate of λ_{ds}, and the fixed length of the data packets is L_{ds}. In order to meet the QoS constraints of the delay-sensitive services, based on the effective bandwidth theory, the effective bandwidth with the transmission delay bound is defined as

$$W(\theta_v) = \lim \frac{1}{t\theta_v} \log E\left(e^{\theta_v Z(t)}\right), \tag{4}$$

where $W(\theta_v)$ is the effective bandwidth, θ_v is the QoS value of the user terminal, $Z(t)$ is the number of packets reached within the period $(0, t)$, and $E(.)$ represents the mathematical expectation.

Assuming that the maximum computing frequency of each MEC server (base station) is $f_m^{\max}(\forall m \in M)$, $f(t) = [f_i, m(t)]$, $m \in M$, and $i \in U$, the amount of computation that the MEC server m can provide for the user's mobile terminals can be expressed as

$$A_{i,m} = \frac{f_{i,m} T}{b_i}, \tag{5}$$

where $f_{i,m}$ represents the computing frequency of each server, and b_i represents the computing load of the user-side tasks, which can be measured in an offline fashion. The MEC network architecture is shown in Figure 3.

In the proposed learning-driven MEC-MAB autonomous offloading algorithm, the user's mobile terminal i is the player, and the MEC server m is the gambling arm. If the user i chooses to access the MEC server m, the corresponding random reward value $Q_{i,m}$ will be obtained. The reward value of each MEC server obeys a specific distribution with mean value as $\pi = [\pi_1, \pi_2, \cdots, \pi_m]$ and is independent of each other, where π_m is the real reward of MEC server m. Since the user cannot always choose the server with the highest real reward, the regret value R_j is defined as the difference between the actual reward value obtained after j selections and the expected maximum reward value:

$$R_j = \pi^* j - \sum_{m=1}^{M} E\left[N_j(m)\right] \pi_m, \tag{6}$$

where $\pi^* = \max_{1\le s\le M} \pi_s$, and $N_j(m)$ is the number of times the MEC server m has been selected in the previous j rounds. Since in the MAB model, the real reward value of the gambling arm is the reward value generated after the action is performed, it is necessary to estimate the reward value for the selection behavior of the gambling arm as follows:

$$W_{j+1}(m) = W_j(m) = \frac{1}{N_j(m)} \left[r_{N_j(m)} - W_j(m)\right]. \tag{7}$$

Using the Thompson-Sampling algorithm [29], the probability of each selection of the reward value of the MEC server in the MAB model is regarded as a Beta (α, β) distribution, and then the reward value distribution probability function of the MEC server selection behavior can be mathematically expressed as

$$f(m, \alpha, \beta) = \frac{1}{B(\alpha, \beta)} m^{a-1}(1-m)^{\beta-1}. \tag{8}$$

The parameter update rule for Beta distribution is given as

$$(\alpha_1, \beta_1) = \begin{cases} (\alpha_1, \beta_1) f_i, & f_i \neq l, \\ (\alpha_1, \beta_1 + (r_1, 1 - r_t) f_i, & f_i = l. \end{cases} \tag{9}$$

Initially, the user mobile terminal observes situational information such as QoS of its computing task and sets $t = 0$ and $\gamma = 0$. at time t, $t \le T$, the reward estimation of the user's mobile terminal for the MEC server selection behavior satisfies $W(m) \sim \mathrm{Beta}(\alpha_m, \beta_m)$. The user selects the MEC server with the largest reward value $\arg\max_m W(m) \longrightarrow \mathrm{MEC}_t$. The network applies the selected access behavior and measures the corresponding reward value r_t, and the parameters are updated as $(\alpha_1, \beta_1) + (r_t, 1 - r_t) \longrightarrow (\alpha_1, \beta_1)$.

In the proposed MEC-MAB algorithm, as the number of observations from the MEC server selections increases, the confidence interval of the beta distribution becomes narrower, enabling the users to automatically select the optimal MEC server with maximum reward.

3.3. Viewpoint-Aware Progressive Dynamic Streaming. This paper proposes a viewpoint-aware dynamic streaming based on ROI and object priority. The strategy first divides the data into the current scenes, potential scenes, and future scenes based on ROI judgment. And the concept of ROI is extended, and the priorities of vertical and horizontal objects are determined through ROI layering, so as to determine the transmission priority of each object in the scene.

The visibility judgment and elimination of object space can reduce the transmission of unnecessary scenes in order

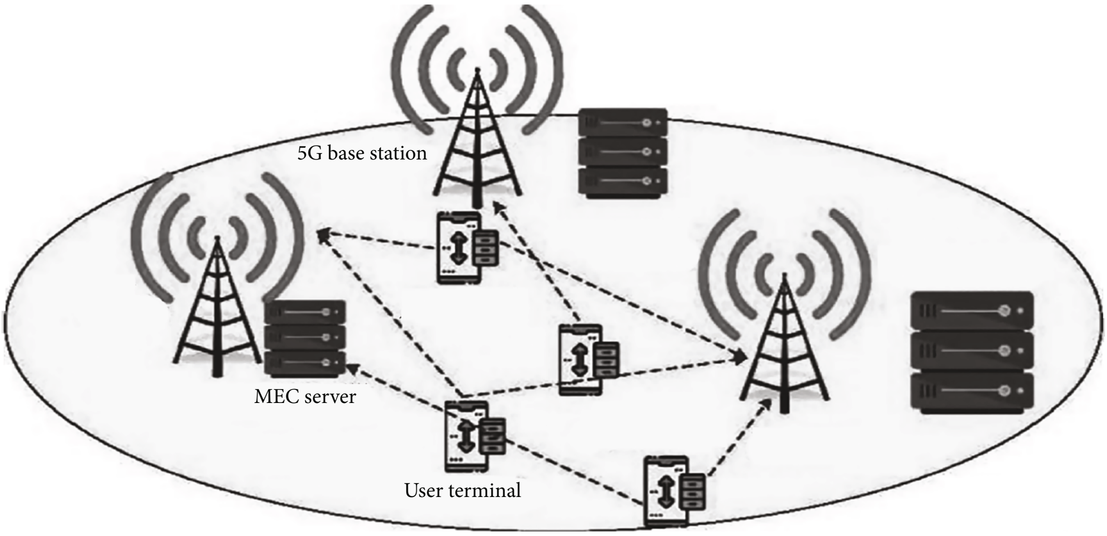


Figure 3: MEC network architect.

to achieve the minimum transmission volume, thereby improving the real-time performance of interaction. To achieve this goal, the whole scene must first be analyzed to determine the visible scenes according to the position and angle of the user, and the visibility should be calculated according to the relevant algorithm to remove those unnecessary or unimportant scenes. Then, the current visible scenes are firstly transmitted, and as the viewpoint moves, the incremental part of the scenes is gradually transmitted, which is also the original intention of progressive transmission.

As shown in Figure 4, the viewpoint range is divided into three areas: CPVS (Current Potential Visible Scenes), IPVS (Incremental Potential Visible Scenes), and FPVS (Future Potential Visible Scenes), based on the visual habit of looking directly in front and then looking around, and observing the near area first and then the far area.

CPVS Zone is the immediate and nearest currently visible scene area. All object models in this area are visible to the user and are relatively close to the user; so, objects in this area should have the highest priority in the process of transmission and interaction. IPVS Zone is as follows: this zone consists of two parts. To the user, the objects in this area are not immediately visible and are relatively far from the user. If the user's viewpoint moves (walks forward or turns), then the objects in this area are immediately visible. Therefore, the object model in this area should have a relatively high priority. After the object model in the CPVS Zone is accessed, the scenes in this area are downloaded first, so that these models can be displayed in time when the user roams. FPVS Zone is as follows: all object models in this area are not currently visible to the user and are relatively far away from the user. Object models in this area should have lower priority, and these objects can be prefetched to the client for use in scene roaming only when the network is idle or if additional network bandwidth is available. Thus, the visual areas are divided into CPVS, IPVS, and FPVS. The object models in this three zones correspond to three queues, queue 1, queue 2, and queue 3, respectively, and the priority of the three queues is queue 1 > queue 2 > queue 3.

When there are many objects in the scene, in order to better distinguish different object models and realize the progressive transmission of the scene, it is necessary to further determine the access sequence of a certain scene with multiple objects in it; that is, it is necessary to determine the visual importance of different objects in a certain area. This paper extends the concept of ROI and determines the order of a specific object in the access queue by layering a certain ROI area and considering the horizontal and vertical importance of different objects.

Level of ROI is from the servers' perspective, the ROI area of the current user is further subdivided into several levels, and the transmission order of the objects is determined accordingly. As shown in Figure 5, if there are multiple objects in the currently visible ROI, the distance and viewing angle can be used to determine the order in which the objects are accessed: objects in level A have the highest priority, objects in level B have the next priority, and objects in level C have the lowest priority.

In this way, the system mainly uses limited bandwidth resources to transmit videos within the user's visible range, compresses redundant content as much as possible, and provides the best viewing effect. At the same time, each frame of video transmitted contains full-view information, which can achieve "device-cloud asynchronous" rendering. When the user's posture changes, the local display device does not need to wait for the server to send back data and completes the rendering locally in real time, updating the scene with the shortest delay to ensure the complete and smooth transition.

The essence of progressive streaming is "download while browsing" to achieve the optimal real-time effect. According to the principle of human vision, the closer the object is to the viewpoint, the smaller the angle that the object deviates from the viewpoint, and the higher the resolution of the object that the viewpoint can observe. To save bandwidth, it is not necessary to download the full model increments,

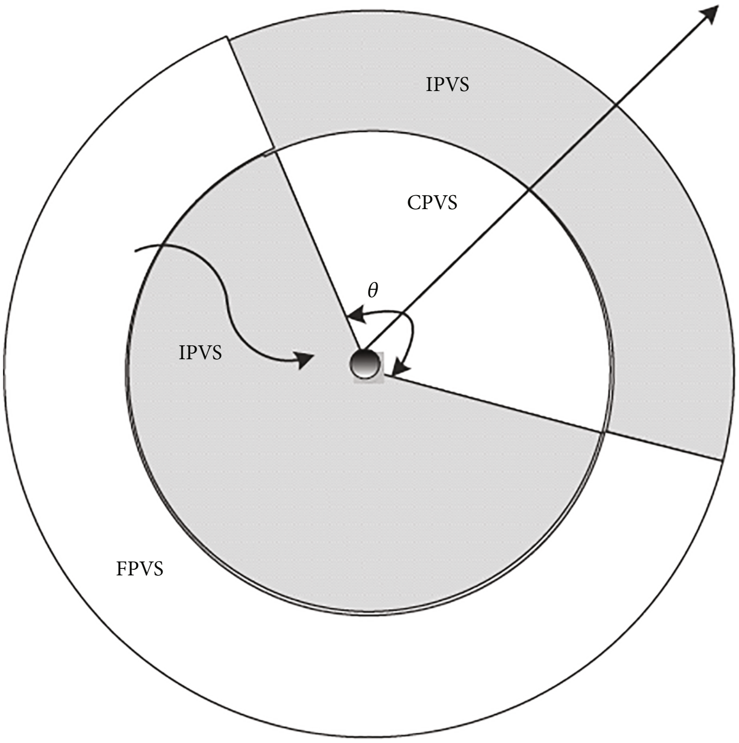


Figure 4: ROI region of the user.

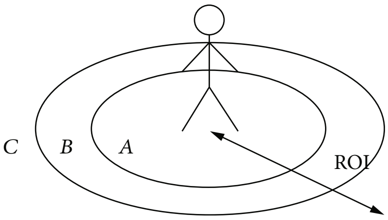


Figure 5: ROI layering.

instead, we can just download the ORM of the scene. The ORM of an object can be determined based on its visual importance to the viewpoint:

$$W(O_i) = \left(1 - \frac{D_i}{R}\right) \cos \frac{\theta_i}{2}, \quad D_i < R, \tag{10}$$

where $W\,(O_i)$ represents the visual importance of the object O_i, R is the ROI radius of the avatar, D_i is the distance between O_i and the avatar's viewpoint, and θi represents the deviation angle between the object i and the avatar's viewpoint ($0 \leq \theta \leq 180°$).

3.4. Optimized HDA Transmission Scheme. In the existing wireless communication system, the channel end encodes the video data into a bit stream for transmission. If there is a bit error in the transmission of the video code stream, the decoding of the video data will cause serious visual distortion or even a decoding failure. Although the current wireless video soft transmission scheme can realize seamless adaptation of video transmission quality to channel conditions, its transmission efficiency is not satisfactory. Combining the high efficiency of traditional digital transmission and the robustness of video soft transmission, the HDA transmission technique has the potential to provide stable, reliable, and high-efficiency VR video transmission in AIoT.

In the proposed HDA transmission system, a time-division multiplexing HDA video soft transmission scheme is designed. The video in the user window is decomposed into two layers: the first layer is the base layer signal, which is generated from the video source compressed by the HEVC encoder. The second layer is the enhancement layer signal, which is the residual value after subtracting the original video signal and the reconstructed signal of the first layer. The two layers of video signals are transmitted in a time-division multiplexing manner. On the one hand, in order to achieve reliable transmission of the digital part, the target bit rate is controlled by the quantization parameter, and the channel coding rate and modulation order are determined by the SNR. On the other hand, overall video quality is directly dependent on the MSE (mean squared error) of the analog signal, which can be expressed as a function of the data variance of the analog part, the power and bandwidth allocated to the analog part, and the noise power of the channel.

In terms of power allocation, it is first necessary to ensure that the base layer can be successfully decoded. Therefore, the overall video quality of HDA video transmission is determined by the data variance of the enhancement layer (analog source), the resources allocated to the analog part, and the channel noise power. Based on the SNR, the bitstream signals at the base layer are turbo-coded with a selected channel coding rate, and the coded signals are subjected to quadrature amplitude modulation. Considering that HEVC (high efficiency video coding) has basically removed the interframe correlation of video sequences, the residual between the original video and the reconstructed video basically does not contain interframe redundancy. The residual part is further decorrelated by 3D-DCT transform, and the power-scaled DCT coefficients are used to modulate the signal amplitude.

Since the signals in the second layer are the enhancement signal of the video, under the condition of limited bandwidth and power, restoring as many of the enhancement layer signals as possible helps to improve the quality of the reconstructed video. After the decorrelation operation of the enhancement layer signals, the energy distribution of the analog coefficients is relatively concentrated, which is manifested in some large coefficients that are concentrated in the upper left corner. In time-division multiplexing coding, appropriate parameters should be selected for the code rate and channel coding modulation mode of the first layer to ensure the correct decoding. Since the first layer is designed to be decoded correctly at a given channel noise power, the overall system distortion is determined by the reconstruction distortion of the second layer. In order to reduce the interference of the large coefficients of the analog part to the digital signal, we try to transmit the large coefficients by time-division multiplexing. Due to the bandwidth limitation, the small coefficients will be discarded. Although discarding small coefficients saves bandwidth, the high-frequency component information carried by these small coefficients cannot be recovered at the receiving end, which will bring additional performance loss.

Without loss of generality, we use MSE as the distortion measure. Let D_a and D_d be analog distortion and digital distortion, respectively. In order to successfully decode the digital base layer, the SNR of the digital part must be greater than the signal-to-noise ratio threshold $\mathrm{SNR}_{\mathrm{th}}$. The spectral efficiency corresponding to $\mathrm{SNR}_{\mathrm{th}}$ is e_f, which depends on the MCS (modulation and coding scheme), and it must satisfy

$$\mathrm{SNR}_{\mathrm{th}} \leq \frac{P_d}{\sigma_n^2}, \tag{11}$$

where P_d is the average power distribution coefficient of the base layer digital signal, and σ_n^2 is the channel noise power with Gaussian white noise added. Since the sum of digital power and analog power is limited by the total power P_T, the power constraint relationship can be expressed as

$$P_T = P_a B_a + P_d B_d, \tag{12}$$

where P_a is the average power allocated to the enhancement layer analog signals, B_a is the bandwidth occupied by the transmitted analog signal, and B_d is the bandwidth occupied by the transmitted digital signal. After the enhancement layer is transformed by 3D-DCT, each group of video frames is further divided into N blocks, and the variance of the i -th block is defined as λ_i. The analog signals transmitted through the channel are interfered by the channel noise, and the distortions can be expressed as [30]

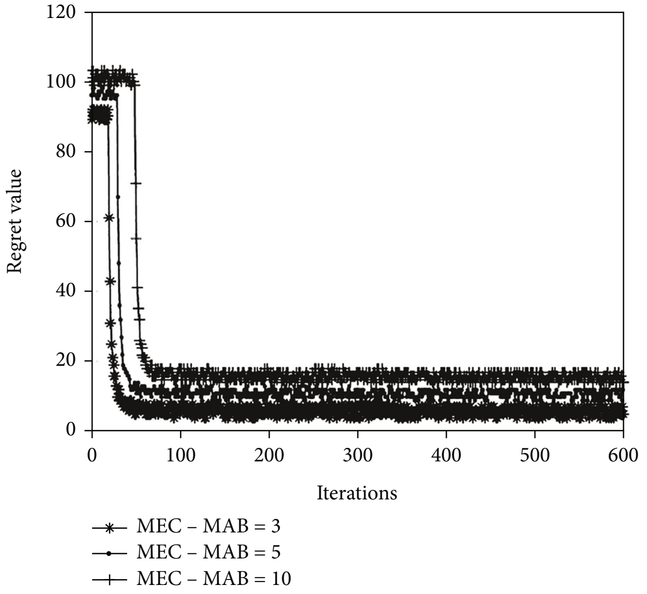


Figure 6: Validation of the proposed MEC offloading strategy.

$$D_{a1} = \sigma_n^2 \left[\left(\sum_{i=1}^{B_a} \sqrt{\frac{\lambda_i}{B_a P_a}} \right)^2 \right]. \tag{13}$$

On the other hand, the discarded coefficients cannot be recovered at the receiver, which also introduces additional distortion D_{a2}:

$$D_{a2} = \sum_{i=B_a+1}^{N} \lambda_i. \tag{14}$$

Therefore, the optimal power allocation problem can be defined as

$$\begin{cases} \min\limits_{P_a \cdot QP} D_a \\ \mathrm{ST} B_a P_a + B_d P_d = P_T, \\ B_a + B_d = N. \end{cases} \tag{15}$$

The variance of the i -th block can be expressed as [31]

$$\lambda_i = e^{k_i \mathrm{QP} + w_i}, \tag{16}$$

where QP is the quantization parameter, and k_i and w_i are two parameters in the i -th block that represent the exponential relationship between λ_i and QP. After the video is digitally compressed and encoded, the number of quantized bits per pixel can also be further fitted with an exponential function, and the fitting parameters are a and b, respectively.

Table 1: PSNR results (dB) of HEVC scheme and proposed HDA scheme under different QPs.

Methods	$QP = 20$			QP = 25			
	$\beta = 0.8$	$\beta = 0.7$	$\beta = 0.5$	$\beta = 0.8$	$\beta = 0.7$	$\beta = 0.5$	$\beta = 0.25$
HEVC	46	46	46	42	42	42	42
Proposed scheme	53	52	51	50	48	46	41

The relationship between the quantization parameter QP and the number of bits produced per pixel R is

$$R(\mathrm{QP}) = ae^{b\mathrm{QP}}. \tag{17}$$

When a group of video frames has M pixels, the total number of bits obtained after digital compression can be calculated as

$$\mathrm{Bit}(\mathrm{QP}) = R(\mathrm{QP})M = ae^{b\mathrm{QP}}M. \tag{18}$$

4. Experiment and Results

Based on the software radio platform from the laboratory and the modified LTE protocol stack of the MEC architecture, a complete MEC platform is build, and it is used as the bearer network to develop the cloud VR system which is suitable for AIoT. The system can use standard commercial terminals or professional VR headsets as client display devices. A complete performance evaluation of the proposed system is performed on this platform.

4.1. MEC Offloading Algorithm Verification. Firstly, the proposed learning-driven MEC task offloading strategy is verified by simulation. Assuming that the number of users is 10, the computing tasks of the users' mobile terminals follow the Poisson distribution, and the path loss exponent is set to 2.

The variation of the simulation regret value with respect to the number of iterations for the number of MEC server nodes (base stations) of 3, 5, and 10 is shown in Figure 6. It can be observed that the network regret value converges in a short time for different numbers of MEC servers. As the number of MEC servers increases, the convergence speed of the algorithm becomes slower, but the overall convergence speed is still reasonable, which shows that the proposed MEC-MAB offloading strategy has good convergence performance.

4.2. VR Transmission Scheme Validation. MATLAB experimental simulation of the proposed HDA transmission scheme is carried out. The HDA transmission system proposed in this paper consists of a data channel and a control channel. The data channel executes the function blocks of the transmit ends and receive ends, respectively, and is based on the power distribution calculation to solve P_a and QP with minimized distortions. The data channel combines digital transmission and pseudoanalog transmission. The digital transmission scheme uses HEVC for source coding and the LTE-based adaptive modulation and coding scheme for transmission. Different combinations of channel coding rates and modulation modes can be selected.

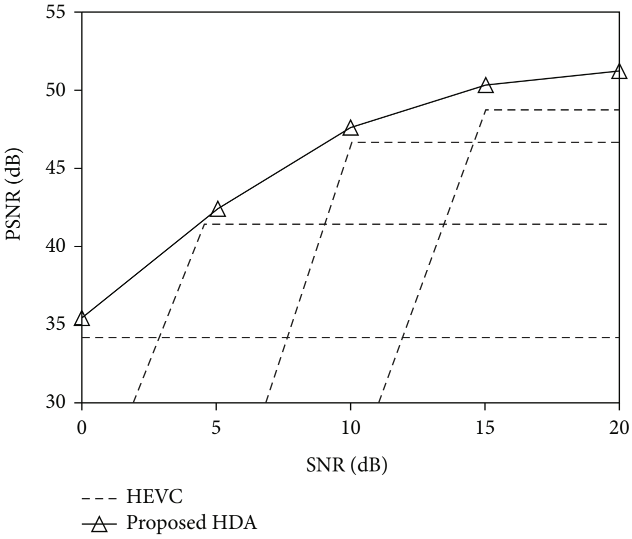


Figure 7: PSNR performance comparison.

Experimental simulations were performed using standard HD sequences with a resolution of 1664×1664 pixels. For analog transmission, each frame in the video sequence is divided into 64 blocks. In the experiment, each picture group is set to consist of 16 frames of images; so, the analog symbols of each picture group are divided into 1024 coefficient blocks. When the video frame rate is 30 fps, the source bandwidth N_s is 41.5 MHz. The bandwidth used for data transmission is defined as N_c, and a specific implementation process is designed to make the number of symbols used for source coding and channel coding in the digital part less than or equal to N_c.

The performance of the HDA transmission scheme proposed in this paper is compared with the existing digital video transmission scheme HEVC. PSNR is measured at the receiving ends to evaluate the quality of video transmission, and same bandwidth and power are used for the two schemes. According to the LTE adaptive modulation and coding scheme, the channel coding adopts LTE turbo coding, and the code rate is $R = 1/3$. The modulation scheme supports QPSK, 16QAM, and 64QAM. Taking the channel SNR = 5.5 dB as an example, the spectral efficiency is about 1.47. Table 1 gives the results of the HEVC scheme and the proposed HDA scheme on the test sequence under different QPs when the target channel SNR = 10 dB and the spectral efficiency of the digital part is 1.47. The ratio of the available video channel bandwidth to the source bandwidth is set to β, i.e., $\beta = N_c$.

It can be seen from Table 1 that when the digital part adopts a certain QP value, the video quality of the receiving end under the HEVC scheme does not change with the change of the available channel bandwidth. When QP = 20, if the available channel bandwidth is only 1/4 of the source bandwidth, the data length of the digital part will exceed the available bandwidth, causing the digital part to not be decoded correctly. In contrast, with the proposed method, under the condition that the available bandwidth resources are severely limited, increasing the QP can realize the encoding and transmission of the digital part of the data.

Next, further performance comparisons between the proposed HDA scheme and HEVC scheme are carried out. Under different channel conditions, the classic HEVC scheme is inevitably affected by the cliff effect. As the SNR increases, so does the spectral efficiency, at which point HEVC will have the opportunity to choose a lower QP. Consider an SNR of 0 to 20 dB, $\beta = 0.5$, and 10% of the bandwidth is reserved for hybrid automatic retransmission of the digital part. As shown in Figure 7, the average PSNR of the proposed HDA scheme is 0.41 dB higher than that of the HEVC scheme. By adding analog signals to the existing digital transmission scheme and dividing part of the bandwidth to the analog signals, the saturation effect of the video quality at the receiving ends can be improved. When the SNR of the target channel is high, analog signal transmission can further achieve greater performance gains and benefit for AIoT.

5. Conclusions

With the rapid development of multimedia services such as VR, there are great requirements for the data transmission in AIoT; also, the resolution of video data is increasing day by day, along with the growing challenge for network bandwidth. Therefore, new business models, such as cloud VR, attracted more attention on the transmission network latency. In this paper, we design and build a cloud VR system based on AIoT using MEC technology to perform dynamic streaming relying on user viewpoint information. Through the learning-driven MEC autonomous offloading strategy, in the absence of prior information such as MEC server computing and storage capabilities and channel status, the optimal MEC server is autonomously selected for task offloading, and the energy consumption is minimized while satisfying user delay constraints. Combined with the viewpoint-aware progressive streaming based on ROI and object priorities used on the server side, the bandwidth requirements are reduced, and a good VR viewing experience is achieved. Meanwhile, based on the relationship between edge servers and base stations in the MEC architecture, HDA transmission technology is merged to further optimize the transmission bandwidth and efficiency in AIoT cloud VR systems. It is known that the use of the AIoT data transmission can enrich the data information and enhance the human-computer interaction in AIoT cloud VR systems. However, it needs more rapidly data processing and more faster network support. In subsequent studies, the availability of better compression schemes for panoramic video and AIoT data transmission in AIoT cloud VR will be further explored, and the channel fading will be further considered with a view to achieving more performance gains.

Data Availability

The data we used is available, and the performance optimization scheme proposed in this paper can be used in adaptive data transmission of VR videos. And part of them are available from the corresponding authors upon request (malinjuan@bit.edu.cn; zfq@mju.edu.cn).

Conflicts of Interest

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

Acknowledgments

Our work is supported by the Digital Media Art, Key Laboratory of Sichuan Province, Sichuan Conservatory of Music (Project No. 21DMAKL01), First Batch of Industry-University Cooperation Collaborative Education Project funded by the Ministry of Education of the People's Republic of China, 2021 (Project No. 202101071001), Minjiang University School-Level Scientific Research Project Funding (Project Nos. MYK17021, MYK18033, and MYK21011), and Minjiang University Introduced Talents Scientific Research Start-Up Fund (Project No. MJY21030).

References

[1] P. Halarnkar, S. Shah, H. Shah, H. Shah, and A. Shah, "A review on virtual reality," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, no. 6, p. 325, 2012.

[2] Z. Lv, "Virtual reality in the context of Internet of Things," *Neural Computing and Applications*, vol. 32, no. 13, pp. 9593–9602, 2020.

[3] T. Kämäräinen, M. Siekkinen, J. Eerikäinen, and A. Ylä-Jääski, "CloudVR: cloud accelerated interactive mobile virtual reality," in *Proceedings of the 26th ACM international conference on Multimedia*, pp. 1181–1189, 2018.

[4] Z. Guo, P. Zhang, and J. Xia, "Design of virtual reality education platform based on 5G MEC," in *2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS)*, pp. 572–578, 2021.

[5] A. Mehrabi, M. Siekkinen, T. Kämäräinen, and A. yl¨-J¨¨ski, "Multi-tier cloudVR," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 17, no. 2, pp. 1–24, 2021.

[6] X. Tang, C. Cao, Y. Wang et al., "Computing power network: the architecture of convergence of computing and networking towards 6G requirement," *China Communications*, vol. 18, no. 2, pp. 175–185, 2021.

[7] Q. Cheng, H. Shan, W. Zhuang, L. Yu, Z. Zhang, and T. Q. Quek, "Design and analysis of MEC- and proactive caching-based 360° mobile VR video streaming," *IEEE Transactions on Multimedia*, vol. 24, pp. 1529–1544, 2022.

[8] Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "A survey on mobile augmented reality with 5G mobile edge computing: architectures, applications, and

technical aspects," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1160–1192, 2021.

[9] K. Bilal and A. Erbad, "Edge computing for interactive media and video streaming," in *Second International Conference on Fog and Mobile Edge Computing (FMEC)*, pp. 68–73, 2017.

[10] R. K. Srinivasa, N. K. S. Naidu, S. Maheshwari, C. Bharathi, and A. H. Kumar, "Minimizing latency for 5G multimedia and V2X applications using mobile edge computing," in *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, pp. 213–217, 2019.

[11] J. Gedeon, F. Brandherm, R. Egert, T. Grube, and M. Muhlhauser, "What the fog? Edge computing revisited: promises, applications and future challenges," *IEEE Access*, vol. 7, pp. 152847–152878, 2019.

[12] L. Wang, L. Jiao, T. He, J. Li, and M. Mühlhäuser, "Service entity placement for social virtual reality applications in edge computing," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 468–476, 2018.

[13] A. Bujari, O. Gaggi, M. Luglio et al., "Addressing the bandwidth demand of immersive applications through NFV in a 5G network," *Mobile Networks and Applications*, vol. 25, no. 3, pp. 1114–1121, 2020.

[14] D. He, C. Luo, F. Wu, and W. Zeng, "Swift: a hybrid digital-analog scheme for low-delay transmission of mobile stereo video," in *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 327–336, 2015.

[15] D. Podborski, E. Thomas, M. M. Hannuksela, S. Oh, T. Stockhammer, and S. Pham, "Virtual reality and DASH," in *International broadcasting convention, Ibc*, 2017.

[16] D. He, C. Westphal, and J. J. Garcia-Luna-Aceves, "Joint rate and fov adaptation in immersive video streaming," in *Proceedings of the 2018 Morning Workshop on Virtual Reality and Augmented Reality Network*, pp. 27–32, 2018.

[17] I. Jo, Y. Park, H. Kim, and J. Bae, "Evaluation of a wearable hand kinesthetic feedback system for virtual reality: psychophysical and user experience evaluation," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 5, pp. 430–439, 2019.

[18] R. Ghaznavi-Youvalari, A. Zare, H. Fang et al., "Comparison of HEVC coding schemes for tile-based viewport-adaptive streaming of omnidirectional video," in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1–6, 2017.

[19] T. El-Ganainy and M. Hefeeda, "Streaming virtual reality content," 2016, http://arxiv.org/abs/1612.08350.

[20] Z. Tu, T. Zong, X. Xi et al., "Content adaptive tiling method based on user access preference for streaming panoramic video," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–4, 2018.

[21] M. Hosseini and V. Swaminathan, "Adaptive 360 VR video streaming based on MPEG-DASH SRD," in *2016 IEEE International Symposium on Multimedia (ISM)*, pp. 407-408, 2016.

[22] H. S. Kim, S. B. Nam, S. G. Choi, C. H. Kim, T. T. K. Sung, and C. B. Sohn, "HLS-based 360 VR using spatial segmented adaptive streaming," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–4, 2018.

[23] F. Guo, H. Zhang, H. Ji, X. Li, and V. C. M. Leung, "An efficient computation offloading management scheme in the densely deployed small cell networks with mobile edge computing," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2651–2664, 2018.

[24] S. Henri, C. Vlachou, and P. Thiran, "Multi-armed bandit in action: optimizing performance in dynamic hybrid networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 4, pp. 1879–1892, 2018.

[25] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, 2019.

[26] M. Liu and Y. Liu, "Price-based distributed offloading for mobile-edge computing with computation capacity constraints," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 420–423, 2018.

[27] Z. Liu, S. Ishihara, Y. Cui, Y. Ji, and Y. Tanaka, "JET: joint source and channel coding for error resilient virtual reality video wireless transmission," *Signal Processing*, vol. 147, pp. 154–162, 2018.

[28] J. Feng, Y. Wu, G. Zhai, N. Liu, and W. Zhang, "An algorithm for transmitting VR video based on adaptive modulation," in *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 443–448, 2019.

[29] J. Chen, B. Li, and R. Srikant, "Thompson-sampling-based wireless transmission for panoramic video streaming," in *2020 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, pp. 1–3, 2020.

[30] K. H. Lee and D. Petersen, "Optimal linear coding for vector Channels," *IEEE Transactions on Communications*, vol. 24, no. 12, pp. 1283–1290, 1976.

[31] B. Tan, H. Cui, J. Wu, and C. W. Chen, "An optimal resource allocation for superposition coding-based hybrid digital–analog system," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 945–956, 2017.

[32] F. Q. Zhang and L. J. Ma, "Multi-object crowd real-time tracking in dynamic environment based on neural network," *Journal of Network Intelligence*, vol. 7, no. 2, pp. 386–394, 2022.

Hindawi
Journal of Sensors
Volume 2022, Article ID 1628344, 18 pages
https://doi.org/10.1155/2022/1628344

Research Article

An Efficient Revocable Identity-Based Encryption with Equality Test Scheme for the Wireless Body Area Network

Tung-Tso Tsai, Han-Yu Lin, and Hsiao-Chieh Chang

Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung 202, Taiwan

Correspondence should be addressed to Tung-Tso Tsai; tttsai@mail.ntou.edu.tw

Received 16 June 2022; Revised 18 July 2022; Accepted 1 August 2022; Published 13 August 2022

Academic Editor: Tsu-Yang Wu

With the rapid development and popularization of cloud computing, people are willing to upload their own data to the cloud to enjoy the services. However, some personal and private data are not suitable for uploading directly to the cloud. Therefore, these data must be encrypted before uploading to the cloud to ensure the confidentiality. To achieve the confidentiality of data and enjoy cloud services, a notion of identity-based encryption with equality test (IBEET) was proposed. Using IBEET, two ciphertexts encrypted under different public keys can be tested to confirm whether they contain the same plaintext. The equality test can be applied to the wireless body area network system in which the cloud can utilize ciphertexts from patients and medical institutions to perform equality tests to determine whether which patient's status is abnormal. Indeed, revoking illegal or expired users on any cryptosystem is an important issue. To the best of our knowledge, there is little research on the design mechanism of user revocation in the IBEET. In this paper, we propose a novel notion of revocable identity-based encryption with an equality test, called RIBEET. Based on the notion, we present the first RIBEET scheme. Meanwhile, the proposed scheme will be proven to be secure under the bilinear Diffie-Hellman (BDH) assumption.

1. Introduction

With the rapid development and popularization of cloud computing, people are willing to upload their own data to the cloud to enjoy the services. However, some personal and private data are not suitable for uploading directly to the cloud. To ensure the confidentiality of data, several encryption mechanisms [1–4] have been applied to cloud computing. Identity-based encryption (IBE) [5] is one of the encryption mechanisms of public key systems. The system of an IBE contains two roles: the private key generator (PKG) and users (including senders and receivers). Each user utilizes his own identity (e.g., e-mail address, name, or social security number) to register with the PKG to obtain a private key. Senders can regard the identity of the receiver as a public key to encrypt private data. After receiving the encrypted message (ciphertext), the receiver can decrypt it with her/his own private key.

To achieve the confidentiality of data and enjoy cloud services, the first identity-based encryption with equality test (IBEET) was proposed by Ma [6]. Using IBEET, two ciphertexts encrypted under different public keys can be tested to confirm whether they contain the same plaintext. Ma [6] also gave an application of IBEET used to classify encrypted e-mails. Each encrypted e-mail can be attached with a tag for classification, while the tag can be encrypted under different public keys in the IBEET system. An e-mail server in the cloud can test the equality of any two encrypted tags to classify encrypted e-mails. Subsequently, many studies on IBEET have been published in the literature [7–11].

The equality test can be applied to the wireless body area network (WBAN) system [12–17] in which the cloud can utilize ciphertexts from patients and medical institutions to perform equality tests to determine whether the patient's status is abnormal. Figure 1 shows the architecture of WBANs. A patient is equipped with wearable sensors to collect her/his health record data from sensors of electroencephalogram (EEG), electrocardiogram (ECG), blood pressure, pulse oximeter, insulin pump, electromyogram (EMG), and motion. These health record data are encrypted through the mobile device and uploaded to the cloud server. On the other hand, the medical institution also uploads the

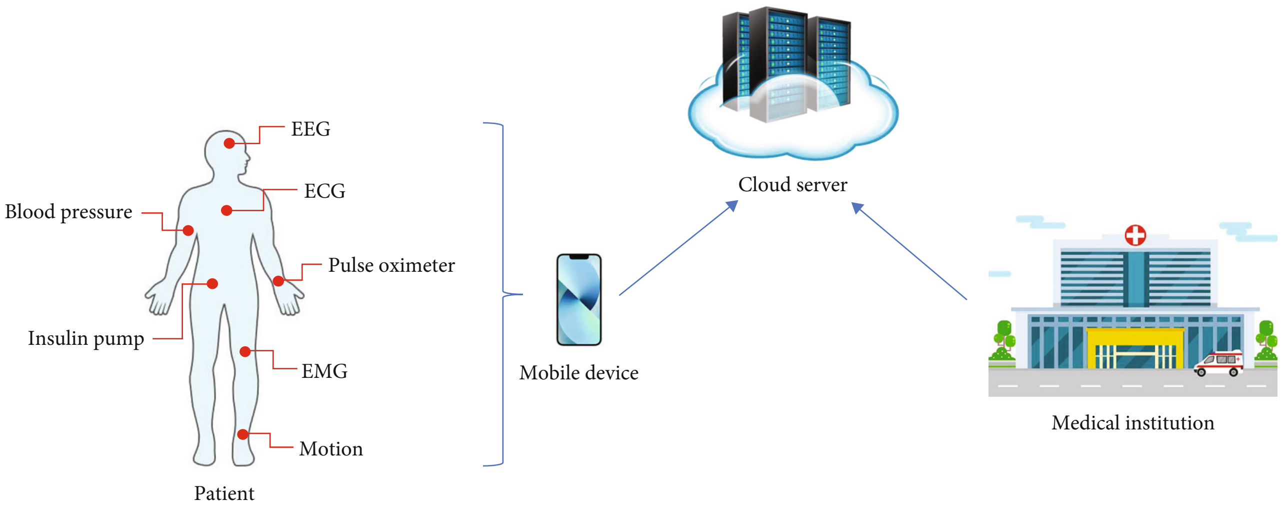


Figure 1: The architecture of WBAN.

patient's encrypted health data to the cloud server. The ciphertexts can be tested for equality without knowing the health data of the patient by the cloud server. If the patient's health data are different from the medical institution's health data, it means that the patient's health data are abnormal.

Indeed, revoking illegal or expired users on any cryptosystem is an important issue. In the traditional public key cryptosystem (PKC), public key infrastructures (PKI) must be established to manage each user's certificate which links the user's identity and public key. In addition, the certificate revocation list [18] is also included in the PKI to revoke illegal or expired users. In identity-based public key cryptosystems (ID-PKC), the first IBE was presented by Boneh and Franklin [5] in which a user can be revoked by the PKG, who sends new private keys for all nonrevoked users at each period, if the user did not receive the new private key. So far, many literatures related to revocable IBE [19–26] have been published. To the best of our knowledge, there is little research on design mechanism of user revocation in the IBEET. In this paper, we propose a novel notion of revocable identity-based encryption with equality test, called RIBEET. Based on the notion, we present the first RIBEET scheme. Meanwhile, the scheme will be proven to be secure under the bilinear Diffie-Hellman (BDH) assumption.

1.1. Related Work. In the era of advanced network communication, cloud computing is an indispensable part. The terminal devices on the user side usually do not have high-performance computing power. However, users can entrust large computing tasks to the cloud. Then, the cloud will return the corresponding results to users after finishing the tasks. Indeed, the cloud can assist each user in performing tasks that require a lot of computation, but it also means that the cloud can know each user's data if the data is not encrypted. Typically, users will encrypt data to the cloud if the data is sensitive or private. In addition, encrypted data also needs to be quickly retrieved from the cloud. To achieve this function, several schemes [3, 4, 27, 28] related to public key encryption with a keyword search were proposed. Although these schemes can retrieve encrypted data, only data encrypted under the same public key can be retrieved.

To support searchable encrypted data under different public keys, Yang et al. [29] proposed a comparison mechanism of two ciphertexts encrypted under different public keys in the traditional public key cryptosystem, called public key encryption with equality test (PKEET). However, the traditional public key cryptosystem must rely on the public key infrastructure to manage each user's certificate which links the user's identity and her/his public key. To avoid the use of public key infrastructure and certificates, Shamir [30] introduced a new concept of ID-PKC in which a user's public key is her/his identity such as name, e-mail, or telephone number. In this way, certificates will no longer be needed in the ID-PKC since the public key is meaningful and can represent the user's identity. Combining the concepts of PKEET and ID-PKC, Ma [6] proposed the first identity-based encryption with equality test, called IBEET. To consider more types of authorizations, Li et al. [31] proposed the IBEET scheme with four types of authorizations. Unfortunately, the proposed scheme of Li et al. [31] is not suitable for the IoT environment because the performance of the scheme is not good. Immediately, Elhabob et al. [10] proposed another IBEET scheme with four types of authorizations which has higher performance.

For the issue of user revocation in the ID-PKC, Boneh and Franklin [5] suggested that the new private keys should be resent to users who have not been revoked at different periods. As a result, secure channels will be established to send these private keys, and the PKG's workload will also increase. To reduce the PKG's workload, Boldyreva et al. [19] hired a binary tree to propose an IBE scheme with user revocation, named revocable IBE (RIBE). However, Boldyreva et al.'s scheme [19] only satisfied the selective-ID security. Later, Libert and Vergnaud [20] proposed another

Table 1: Comparisons between the existing schemes and our RIBEET scheme.

Schemes	Public key setting	Avoiding the use of certificates	Supporting equality test of ciphertexts	Providing user revocation
PKEET [29]	PKI-based	No	Yes	Yes
IBE [5]	ID-based	Yes	No	No
RIBE [21]	ID-based	Yes	No	Yes
IBEET [6]	ID-based	Yes	Yes	No
Our RIBEET	ID-based	Yes	Yes	Yes

RIBE scheme which meets the adaptive-ID security. A mechanism for revoking users through public channels was proposed by Tseng and Tsai [21], in which each user's full private key is divided into two parts: a fixed key and a time updated key. The fixed key is delivered to the user through secure channels only once, while the time updated key is delivered to the user through public channels at different periods. Users can be revoked if they do not receive the new time updated keys. For the security of decryption key exposure, Seo and Emura [22] proposed a new RIBE scheme to enhance the security. To reduce the length of public parameters and meet the security of decryption key exposure resistance, Watanabe et al. [23] presented another RIBE scheme. In addition, several lattice-based RIBE schemes [24–26] were proposed to resist quantum attacks.

1.2. Motivation. As mentioned earlier, revoking illegal or expired users on any cryptosystem is still an important issue. In the traditional PKC, the PKEET [29] can hire the certificate revocation list [18] to revoke illegal or expired users. However, the IBE [5] cannot effectively revoke illegal or expired users in the ID-PKC, so the RIBE [21] was proposed. To the best of our knowledge, there is little research on the design mechanism of user revocation in the IBEET [6]. Table 1 shows the comparisons between the PKEET [29], the IBE [5], the RIBE [21], the IBEET [6], and our RIBEET in terms of public key setting, avoiding the use of certificates, supporting the equality test of ciphertexts, and providing user revocation. Hence, we attempt to propose the first revocable identity-based encryption with equality test, called RIBEET.

1.3. Contribution and Organization. Although the existing RIBE schemes [21–26] provide a mechanism to revoke users, they do not extend to support the equality test for ciphertexts. On the other hand, the existing IBEET schemes [6, 10, 31] do not support to revoke users. To the best of our knowledge, there is little research on the design mechanism of user revocation in the IBEET. In this paper, we propose a novel notion of revocable identity-based encryption with equality test, called RIBEET. In the following, we list specific contributions.

(i) Based on the existing syntax and security notions of IBEET, we consider the property of user revocation to define a new syntax and security notions of RIBEET

(ii) Following the syntax of RIBEET, a concrete RIBEET scheme is proposed

(iii) In the security notions of RIBEET, the proposed scheme is proven to be secure under the bilinear Diffie-Hellman (BDH) assumption

(iv) We compare the proposed scheme with the previous RIBE scheme and IBEET scheme. We demonstrate that the proposed scheme not only provides user revocation but also supports the equality test for ciphertexts

The rest of the article includes six sections. Preliminaries are given in Section 2. Section 3 presents the syntax and security notions of RIBEET. A concrete RIBEET scheme is proposed in Section 4. The security analysis of the RIBEET scheme is shown in Section 5. We compare the RIBEET scheme with other existing schemes in Section 6. The last section gives the conclusion.

2. Preliminaries

In this section, we introduce two definitions related to a mathematical tool and security assumption. We hire the bilinear pairings [5] as a mathematical tool to construct our RIBEET scheme. To prove the security of the proposed scheme, we consider the bilinear Diffie-Hellman (BDH) problem and then give a BDH assumption [6]. The definition of the bilinear pairings is given as follows.

Definition 1. Let G_1, G_2, and G_T be three multiplicative cyclic groups of a prime order q. Assume that a mapping $\widehat{e} : G_1 \times G_2 \longrightarrow G_T$ is an asymmetric bilinear map. Then, the map $\widehat{e}$ satisfies the following properties.

(1) Bilinearity: $\widehat{e}(g_1^a, g_2^b) = \widehat{e}(g_1, g_2)^{ab}$ for $g_1 \in G_1$, $g_2 \in G_2$, and $a, b \in Z_q^*$

(2) Nondegeneracy: $\widehat{e}(g_1, g_2) \neq 1$ for some $g_1 \in G_1$ and $g_2 \in G_2$

(3) Computability: $\widehat{e}(g_1, g_2)$ can be efficiently computed for $g_1 \in G_1$, $g_2 \in G_2$

For the asymmetric bilinear map, the BDH problem is to compute $\widehat{e}(g_1, g_2)^{abc}$ by given a tuple $\langle q, G_1, G_2, G_T, \widehat{e}, g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b \rangle$. We define the BDH assumption as follows.

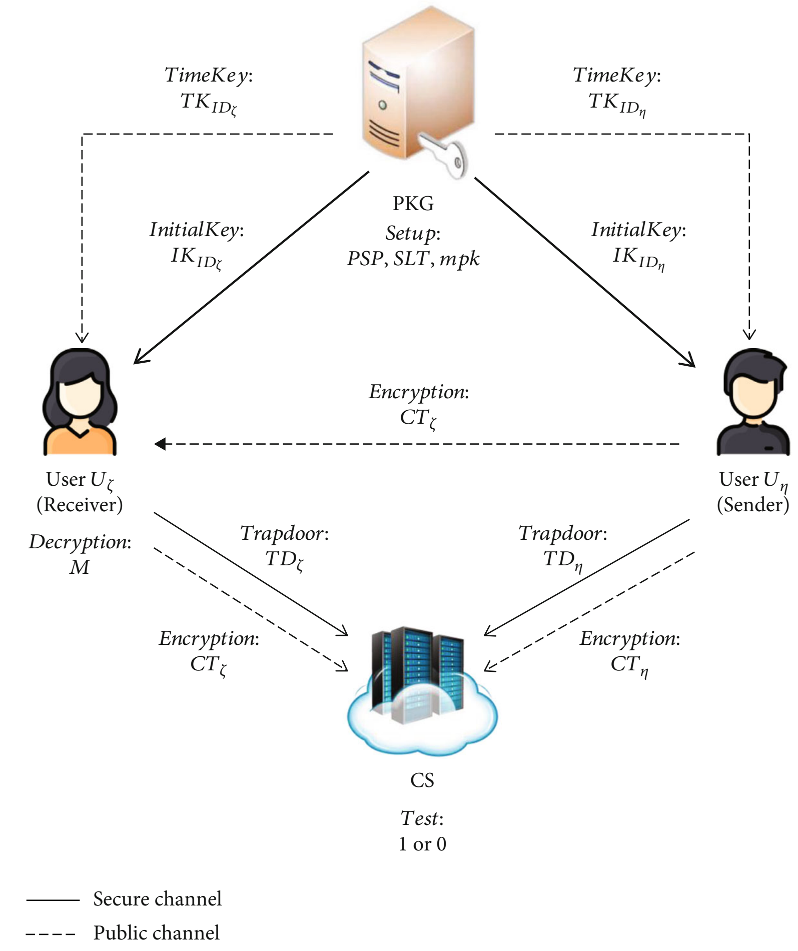
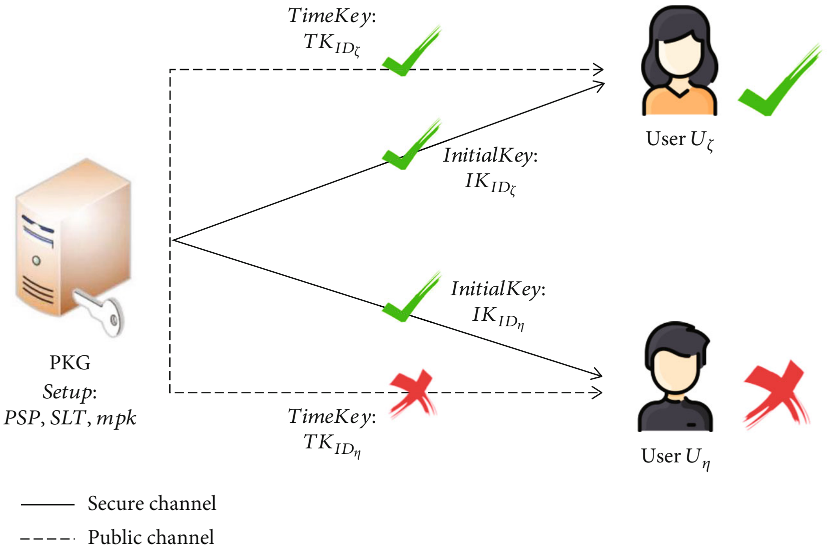


Figure 2: The syntax of RIBEET.



$TimeKey$: TK_{ID_ζ}
$InitialKey$: IK_{ID_ζ}
User U_ζ
$InitialKey$: IK_{ID_η}
PKG
$Setup$: PSP, SLT, mpk
$TimeKey$: TK_{ID_η}
User U_η
Secure channel
Public channel

Figure 3: The procedure for user revocation.

Table 2: Notations.

Notations	Meaning
PSP	The public system parameters
SLT	The system life time
mpk	The master private key
ID	The user's identity
$\mathrm{IK}_{\mathrm{ID}}$	The user initial key
$\mathrm{TK}_{\mathrm{ID}}$	The user time key
M	The message
CT	The ciphertext
$\mathrm{TD}_{\mathrm{ID}}$	The trapdoor
$(\mathrm{CT}_{\zeta}, \mathrm{TD}_{\zeta})$	The ciphertext-trapdoor pair of the user U_{ζ}
$(\mathrm{CT}_{\eta}, \mathrm{TD}_{\eta})$	The ciphertext-trapdoor pair of the user U_{η}

Definition 2. On inputting a tuple $\langle q, G_1, G_2, G_T, \hat{e}, g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b \rangle$, we say that the BDH problem holds if no algorithm $\mathscr{A}$ has nonnegligible advantage in computing $\hat{e}(g_1, g_2)^{abc}$. The advantage can be denoted as $\Pr[\mathscr{A}(g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b) = \hat{e}(g_1, g_2)^{abc}] < \varepsilon$.

3. Syntax and Security Notions

3.1. Syntax of RIBEET. Based on the syntax of IBEET schemes [6], we employ the revocation technique [21] to present a new syntax of RIBEET depicted in Figure 2 which consists of three roles and seven algorithms, namely Setup, InitialKey, TimeKey, Encryption, Decryption, Trapdoor, and Test. The first role is the private key generator (PKG) who is responsible for executing the first three algorithms, and the second role is the users who can, respectively, utilize Encryption, Decryption, and Trapdoor algorithms for encryption, decryption, and authorization. The last role is the cloud server (CS) who runs the Test algorithm to compare the two ciphertexts. For the user revocation, we use Figure 3 to illustrate how users are revoked by the PKG. If the PKG stops sending the time key to a user, it means that the user has been revoked since both initial key and time key are required to execute Decryption and Trapdoor algorithms. Here, we arrange some notations used in these algorithms in Table 2. The algorithms of RIBEET are described in detail as follows.

(i) Setup: this algorithm is performed by the PKG who takes a security parameter k and a time period t as input to produce the public system parameters PSP, the system life time SLT, and the master private key mpk

(ii) InitialKey: this algorithm is performed by the PKG who takes the public system parameter PSP, the master private key mpk, and a user's identity ID $\in \{0,1\}^*$ as input to produce user initial key $\mathrm{IK}_{\mathrm{ID}}$

(iii) TimeKey: this algorithm is performed by the PKG who takes the public system parameter PSP, the master private key mpk, a user's identity ID $\in \{0,1\}^*$, and a period $T \in$ SLT as input to produce user time key $\mathrm{TK}_{\mathrm{ID}}$

(iv) Encryption: this algorithm is performed by a user (sender) who takes the public system parameter PSP, a user's identity ID $\in \{0,1\}^*$, a period $T \in$ SLT, and a message $M \in \{0,1\}^{\lambda}$ as input to produce a ciphertext CT

(v) Decryption: this algorithm is performed by a user (receiver) who takes the public system parameter PSP, the receiver's initial key $\mathrm{IK}_{\mathrm{ID}}$, the receiver's time key $\mathrm{TK}_{\mathrm{ID}}$, and the ciphertext CT as input to produce the message M

(vi) Trapdoor: this algorithm is performed by a user who takes her/his initial key $\mathrm{IK}_{\mathrm{ID}}$ and time key $\mathrm{TK}_{\mathrm{ID}}$ as input to produce the trapdoor $\mathrm{TD}_{\mathrm{ID}}$

(vii) Test: this algorithm is performed by the CS who takes the public system parameters PSP and two ciphertext-trapdoor pairs $(\mathrm{CT}_{\zeta}, \mathrm{TD}_{\zeta})$ and $(\mathrm{CT}_{\eta}, \mathrm{TD}_{\eta})$ from any two users U_{ζ} and U_{η} as input to produce 1 or 0

3.2. Security Notions of RIBEET. In this section, we define the security notions of RIBEET which includes four types of adversaries. Two of these types are the same as the security notions of IBEET [6]. Considering the revoked users from RIBEET, we need to add two types of adversaries in the security notions. These four types of adversaries are presented as follows.

(1) Type I adversary: such an adversary can obtain all information (including time key $\mathrm{TK}_{\mathrm{ID}}$) transmitted through public channels. The adversary can be regarded as an outside attacker

(2) Type II adversary: such an adversary owns her/his initial key $\mathrm{IK}_{\mathrm{ID}}$, but he does not have the current time key $\mathrm{TK}_{\mathrm{ID}}$. The adversary can be regarded as a revoked user

(3) Type III adversary: this adversary is identical to the type I adversary, except that she/he possesses the trapdoor TD

(4) Type IV adversary: this adversary is identical to the type II adversary, except that she/he possesses the trapdoor TD

Following the security notions of IBEET [6], we consider revoked users to define the new security notions of RIBEET. Definitions 3 and 4, respectively, are given to state IND-ID-CCA and OW-ID-CCA security of an RIBEET scheme.

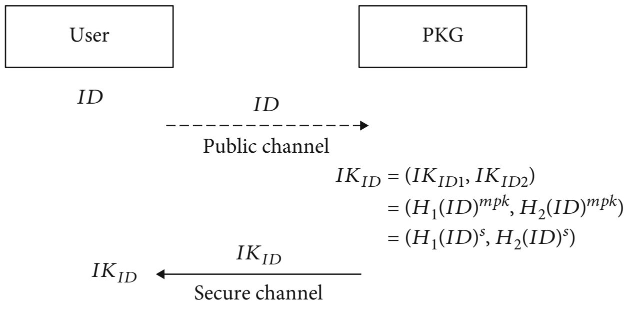

User
PKG
ID
ID
Public channel
$IK_{ID} = (IK_{ID1}, IK_{ID2})$
$= (H_1(ID)^{mpk}, H_2(ID)^{mpk})$
$= (H_1(ID)^s, H_2(ID)^s)$
IK_{ID}
IK_{ID}
Secure channel

Figure 4: InitialKey procedure.

Definition 3 (IND-ID-CCA). Let $\mathscr{A}$ be a type I or type II adversary for an RIBEET scheme and $\mathscr{B}$ be a challenger in the following game. The scheme is IND-ID-CCA secure if the advantage that $\mathscr{A}$ wins the game is negligible.

(1) Setup. The challenger $\mathscr{B}$ takes a security parameter k and a time period t as input to produce the public system parameters PSP, the system life time SLT, and the master private key mpk. The public system parameters PSP and the system life time SLT are sent to the adversary $\mathscr{A}$

(2) Phase 1. Several queries below can be issued by the adversary $\mathscr{A}$

(a) InitialKey query(ID): given an identity ID, the challenger $\mathscr{B}$ generates an initial key IK_{ID} as the response by running the InitialKey algorithm of the RIBEET scheme

(b) TimeKey query(ID, T): given an identity ID and a period T, the challenger $\mathscr{B}$ generates a time key T K_{ID} as the response by running the TimeKey algorithm of the RIBEET scheme

(c) Decryption query(ID, T, CT): given an identity ID, a period T, and a ciphertext CT, the challenger $\mathscr{B}$ generates the resulting message M as the response by running the Decryption of the RIBEET scheme

(d) Trapdoor query(ID, T): given an identity ID and a period T, the challenger $\mathscr{B}$ generates a trapdoor T D_{ID} as the response by running the Trapdoor of the RIBEET scheme

(3) Challenge. Two messages M_0^*, M_1^*, an identity ID*, and a period T^* are submitted by the adversary $\mathscr{A}$. The challenger $\mathscr{B}$ chooses $M_{\mathfrak{b}}^*$ from these two messages, where $\mathfrak{b} \in \{0,1\}$ is a random coin. The challenger $\mathscr{B}$ then generates a ciphertext CT* as the challenge one by running the Encryption of the RIBEET scheme with $(\text{ID}^*, T^*, M_{\mathfrak{b}}^*)$. Here, the following restrictions must be satisfied

(i) The adversary $\mathscr{A}$ cannot issue the Trapdoor query with ID*

(ii) The adversary $\mathscr{A}$ cannot issue the InitialKey query with ID* if it is the type I adversary

(iii) The adversary $\mathscr{A}$ cannot issue the TimeKey query with (ID^*, T^*) if it is the type II adversary

(4) Phase 2. Under the above restrictions, $\mathscr{A}$ can execute the same tasks as in phase 1

(5) Guess. The adversary $\mathscr{A}$ outputs a guess $\mathfrak{b}' \in \{0,1\}$ and wins the game if $\mathfrak{b}' = \mathfrak{b}$. The advantage that $\mathscr{A}$ wins the game can be denoted as $\text{Adv}_{\mathscr{A}}(k) = |\Pr[\mathfrak{b}' = \mathfrak{b}] - 1/2|$

Definition 4 (OW-ID-CCA). Let $\mathscr{A}$ be a type III or type IV adversary for an RIBEET scheme and $\mathscr{B}$ be a challenger in the following game. The scheme is OW-ID-CCA secure if the advantage that $\mathscr{A}$ wins the game is negligible.

(1) Setup. The challenger $\mathscr{B}$ takes a security parameter k and a time period t as input to produce the public system parameters PSP, the system life time SLT, and the master private key mpk. The public system parameters PSP and the system life time SLT are sent to the adversary $\mathscr{A}$

(2) Phase 1. Several queries below can be issued by the adversary $\mathscr{A}$

(a) InitialKey query(ID): given an identity ID, the challenger $\mathscr{B}$ generates an initial key IK_{ID} as the response by running the InitialKey algorithm of the RIBEET scheme

(b) TimeKey query(ID, T): given an identity ID and a period T, the challenger $\mathscr{B}$ generates a time key T K_{ID} as the response by running the TimeKey algorithm of the RIBEET scheme

(c) Decryption query(ID, T, CT): given an identity ID, a period T, and a ciphertext CT, the challenger $\mathscr{B}$ generates the resulting message M as the response by running the Decryption of the RIBEET scheme

(d) Trapdoor query(ID, T): given an identity ID and a period T, the challenger $\mathscr{B}$ generates a trapdoor T D_{ID} as the response by running the Trapdoor of the RIBEET scheme

(3) Challenge. An identity ID* and a period T^* are submitted by the adversary $\mathscr{A}$. The challenger $\mathscr{B}$ randomly chooses M^* and then generates a ciphertext CT* as the challenge one by running the Encryption

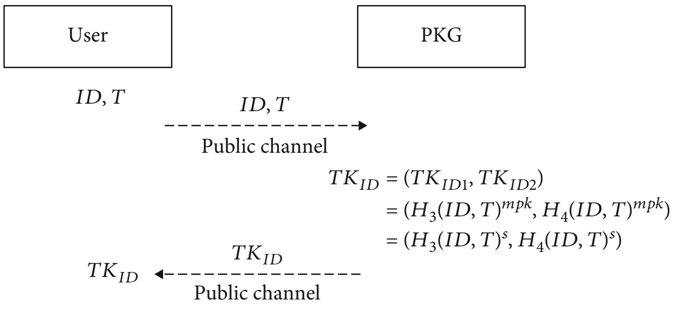

User
PKG
ID, T
ID, T
Public channel
$TK_{ID} = (TK_{ID1}, TK_{ID2})$
$= (H_3(ID, T)^{mpk}, H_4(ID, T)^{mpk})$
$= (H_3(ID, T)^s, H_4(ID, T)^s)$
TK_{ID}
TK_{ID}
Public channel

Figure 5: TimeKey procedure.

Table 3: Comparisons of the proposed RIBEET with the existing RIBE and several IBEET.

Schemes	The cost of performing encryption	The cost of performing decryption	The cost of performing equality test	Supporting equality test of ciphertexts	Providing user revocation
RIBE [21]	$1 \cdot \text{Pair} + 2 \cdot \text{Exp}$ (8.7843 ms)	$1 \cdot \text{Pair}$ (7.8351 ms)	—	No	Yes
IBEET [6]	$2 \cdot \text{Pair} + 6 \cdot \text{Exp}$ (18.5178 ms)	$2 \cdot \text{Pair} + 2 \cdot \text{Exp}$ (16.6194 ms)	$4 \cdot \text{Pair}$ (31.3404 ms)	Yes	No
IBEET [10]	$2 \cdot \text{Pair} + 4 \cdot \text{Exp}$ (17.5686 ms)	$2 \cdot \text{Pair} + 1 \cdot \text{Exp}$ (16.1448 ms)	$2 \cdot \text{Pair} + 2 \cdot \text{Exp}$ (16.6194 ms)	Yes	No
IBEET [11]	$2 \cdot \text{Pair} + 9 \cdot \text{Exp}$ (19.9416 ms)	$2 \cdot \text{Pair} + 2 \cdot \text{Exp}$ (16.6194 ms)	$2 \cdot \text{Pair} + 2 \cdot \text{Exp}$ (16.6194 ms)	Yes	No
Our RIBEET	$2 \cdot \text{Pair} + 5 \cdot \text{Exp}$ (18.0432 ms)	$2 \cdot \text{Pair} + 2 \cdot \text{Exp}$ (16.6194 ms)	$4 \cdot \text{Pair}$ (31.3404 ms)	Yes	Yes

Table 4: Comparison of communication cost.

Schemes	|PK|	|CT|	|TD|
RIBE [21]	$2\|G_1\|$	$\|G_1\| + \|Z_q\|$	—
IBEET [6]	$\|G_1\|$	$4\|G_1\| + \|Z_q\|$	$\|G_1\|$
IBEET [10]	$\|G_1\|$	$2\|G_1\| + 2\|Z_q\|$	$\|G_1\|$
IBEET [11]	$\|G_1\| + \|Z_q\|$	$3\|G_1\| + \|Z_q\|$	$4\|G_1\|$
Our RIBEET	$2\|G_2\|$	$3\|G_1\| + \|Z_q\|$	$\|G_2\|$

of the RIBEET scheme with $(\mathrm{ID}^*, T^*, M^*)$. Here, the following restrictions must be satisfied

(a) The adversary $\mathscr{A}$ cannot issue the InitialKey query with ID* if it is the type III adversary

(b) The adversary $\mathscr{A}$ cannot issue the TimeKey query with (ID^*, T^*) if it is the type IV adversary

(4) Phase 2. Under the above restrictions, $\mathscr{A}$ can execute the same tasks as in phase 1

(5) Guess. The adversary $\mathscr{A}$ outputs a guess M' and wins the game if $M^* = M'$. The advantage that $\mathscr{A}$ wins the game can be denoted as $\mathrm{Adv}_{\mathscr{A}}(k) = \Pr[M^* = M']$.

4. Concrete RIBEET Scheme

A revocable identity-based encryption with equality test scheme, which we denote by RIBEET, consists of algorithms Setup, InitialKey, TimeKey, Encryption, Decryption, Trapdoor, and Test. Each of the algorithms is described as follows.

(1) Setup: this algorithm is performed by the PKG who takes a security parameter k and a time period t as input to produce an asymmetric bilinear map $\hat{e} : G_1 \times G_2 \longrightarrow G_T$ and a system life time $\mathrm{SLT} = \{T_0, T_1, \cdots, T_t\}$, where G_1, G_2, and G_T are multiplicative cyclic groups of prime order q. The PKG first chooses two arbitrary generators $g_1 \in G_1$ and $g_2 \in G_2$ and picks eight cryptographic one-way hash functions $H_1 : \{0,1\}^* \longrightarrow G_2$, $H_2 : \{0,1\}^* \longrightarrow G_2$, $H_3 : \{0,1\}^* \longrightarrow G_2$, $H_4 : \{0,1\}^* \longrightarrow G_2$, $H_5 : G_T \times G_1 \times G_1 \longrightarrow \{0,1\}^{\lambda+l}$, $H_6 : \{0,1\}^{\lambda} \longrightarrow G_2$, $H_7 : \{0,1\}^{\lambda+l} \longrightarrow Z_q^*$, and $H_8 : G_T \longrightarrow G_2$, where λ and l are fixed lengths. Then, a random value $s \in Z_q^*$ is chosen, and $P_{\mathrm{pub}} = g_1^s$ is computed. The public system parameters are $\mathrm{PSP} = \{q, G_1$

$, G_2, G_T, \widehat{e}, g_1, g_2, P_{\text{pub}}, H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$, the system life time is $\text{SLT} = \{T_0, T_1, \cdots, T_t\}$, and the master private key is $\text{mpk} = s$

(2) InitialKey: this algorithm is performed by the PKG who takes the public system parameter PSP, the master private key mpk, and a user's identity $\text{ID} \in \{0, 1\}^*$ as input to produce user initial key

$$\begin{aligned} \text{IK}_{\text{ID}} &= (\text{IK}_{\text{ID1}}, \text{IK}_{\text{ID2}}) \\ &= \left(H_1(\text{ID})^{\text{mpk}}, H_2(\text{ID})^{\text{mpk}}\right) \\ &= (H_1(\text{ID})^s, H_2(\text{ID})^s). \end{aligned} \tag{1}$$

Here, the procedure of this algorithm is depicted in Figure 4.

(3) TimeKey: this algorithm is performed by the PKG who takes the public system parameter PSP, the master private key mpk, a user's identity $\text{ID} \in \{0, 1\}^*$, and a period $T \in \text{SLT}$ as input to produce user time key

$$\begin{aligned} T\text{K}_{\text{ID}} &= (\text{TK}_{\text{ID1}}, \text{TK}_{\text{ID2}}) \\ &= \left(H_3(\text{ID}, T)^{\text{mpk}}, H_4(\text{ID}, T)^{\text{mpk}}\right) \\ &= (H_3(\text{ID}, T)^s, H_4(\text{ID}, T)^s). \end{aligned} \tag{2}$$

Here, the procedure of this algorithm is depicted in Figure 5.

(4) Encryption: this algorithm is performed by a sender who takes the public system parameter PSP, a user's identity $\text{ID} \in \{0, 1\}^*$, a period $T \in \text{SLT}$, and a message $M \in \{0, 1\}^\lambda$ as input to produce ciphertexts $\text{CT} = (\text{CT}_1, \text{CT}_2, \text{CT}_3, \text{CT}_4)$ which are shown as follows

(a) $\text{CT}_1 = g_1^r$

(b) $\text{CT}_2 = g_1^u$

(c) $\text{CT}_3 = H_5(\widehat{e}(P_{\text{pub}}, H_1(\text{ID}) \cdot H_3(\text{ID}, T))^u, \text{CT}_1, \text{CT}_2) \oplus (M||V)$

(d) $\text{CT}_4 = H_6(M)^r \cdot H_8(\widehat{e}(P_{\text{pub}}, H_2(\text{ID}) \cdot H_4(\text{ID}, T))^u)$

Here, $r = H_7(M, V)$ and the two values $V \in \{0, 1\}^l$ and $u \in Z_q^*$ are chosen in random.

(5) Decryption: this algorithm is performed by a receiver who takes the public system parameter PSP, the receiver's initial key $\text{IK}_{\text{ID}} = (\text{IK}_{\text{ID1}}, \text{IK}_{\text{ID2}})$, the receiver's time key $\text{TK}_{\text{ID}} = (\text{TK}_{\text{ID1}}, \text{TK}_{\text{ID2}})$, and the ciphertext $\text{CT} = (\text{CT}_1, \text{CT}_2, \text{CT}_3, \text{CT}_4)$ as input to produce the message M. The detailed process is shown as follows:

(a) Compute $\text{CT}_3 \oplus H_5(\widehat{e}(\text{CT}_2, \text{IK}_{\text{ID1}} \cdot \text{TK}_{\text{ID1}}), \text{CT}_1, \text{CT}_2)$ to obtain $M'||V'$

(b) Compute $r' = H_7(M', V')$

(c) Produce the message M' as M if $\text{CT}_1 = g_1^{r'}$ and $\text{CT}_4 = H_6(M)^{r'} \cdot H_8(\widehat{e}(\text{CT}_2, \text{IK}_{\text{ID2}} \cdot \text{TK}_{\text{ID2}}))$ both hold

The correctness of obtaining $M'||V'$ can be demonstrated as follows.

$$\begin{aligned} \text{CT}_3 \oplus H_5(\widehat{e}(CT_2, \text{IK}_{\text{ID1}} \cdot \text{TK}_{\text{ID1}}), CT_1, CT_2) &= H_5\left(\widehat{e}\left(P_{\text{pub}}, H_1(\text{ID}) \cdot H_3(\text{ID}, T)\right)^u, \text{CT}_1, \text{CT}_2\right) \oplus \left(M'||V'\right) \oplus H_5(\widehat{e}(\text{CT}_2, \text{IK}_{\text{ID1}} \cdot \text{TK}_{\text{ID1}}), \text{CT}_1, \text{CT}_2) \\ &= H_5(\widehat{e}(g_1^s, H_1(\text{ID}) \cdot H_3(\text{ID}, T))^u, \text{CT}_1, \text{CT}_2) \oplus \left(M'||V'\right) \oplus H_5(\widehat{e}(g_1^u, H_1(\text{ID})^s \cdot H_3(\text{ID}, T)^s), \text{CT}_1, \text{CT}_2) \\ &= H_5(\widehat{e}(g_1, H_1(\text{ID}) \cdot H_3(\text{ID}, T))^{su}, \text{CT}_1, \text{CT}_2) \oplus \left(M'||V'\right) \oplus H_5(\widehat{e}(g_1, H_1(\text{ID}) \cdot H_3(\text{ID}, T))^{su}, \text{CT}_1, \text{CT}_2) = M'||V'. \end{aligned} \tag{3}$$

(6) Trapdoor: this algorithm is performed by a user who takes her/his initial key $\text{IK}_{\text{ID}} = (\text{IK}_{\text{ID1}}, \text{IK}_{\text{ID2}})$ and time key $\text{TK}_{\text{ID}} = (\text{TK}_{\text{ID1}}, \text{TK}_{\text{ID2}})$ as input to produce the trapdoor $\text{TD}_{\text{ID}} = \text{IK}_{\text{ID2}} \cdot \text{TK}_{\text{ID2}} = H_2(\text{ID})^s \cdot H_4(\text{ID}, T)^s$

(7) Test: this algorithm is performed by the CS who takes the public system parameters PSP and two ciphertext-trapdoor pairs $(\text{CT}_\zeta, \text{TD}_\zeta)$ and $(\text{CT}_\eta, \text{TD}_\eta)$, where $\text{CT}_\zeta = (\text{CT}_{\zeta1}, \text{CT}_{\zeta2}, \text{CT}_{\zeta3}, \text{CT}_{\zeta4})$ and $\text{CT}_\eta = (\text{CT}_{\eta1}, \text{CT}_{\eta2}, \text{CT}_{\eta3}, \text{CT}_{\eta4})$, from any two users U_ζ and U_η as input to produce 1 or 0 according to the following steps

(a) Compute R_ζ and R_η as follows:

(i) $R_\zeta = \text{CT}_{\zeta4}/H_8(\widehat{e}(\text{CT}_{\zeta2}, \text{TD}_\zeta)) = H_6(M_\zeta)^{H_7(M_\zeta, V_\zeta)}$

(ii) $R_\eta = \text{CT}_{\eta4}/H_8(\widehat{e}(\text{CT}_{\eta2}, \text{TD}_\eta)) = H_6(M_\eta)^{H_7(M_\eta, V_\eta)}$

Table 5: Comparison of energy cost.

Energy cost	RIBE [21]	IBE_{ET} [6]	IBEET [10]	IBEET [11]	Our RIBEET
Performing encryption	358.256 μJ	755.224 μJ	716.508 μJ	813.292 μJ	735.868 μJ

(b) Compute $\hat{e}(\text{CT}_{\zeta 1}, R_\eta)$ and $\hat{e}(\text{CT}_{\eta 1}, R_\zeta)$

(i) $\hat{e}(\text{CT}_{\zeta 1}, R_\eta) = \hat{e}(g_1^{H_7(M_\zeta, V_\zeta)}, H_6(M_\eta)^{H_7(M_\eta, V_\eta)}) = \hat{e}(g_1, H_6(M_\eta))^{H_7(M_\zeta, V_\zeta)\cdot H_7(M_\eta, V_\eta)}$

(ii) $\hat{e}(\text{CT}_{\eta 1}, R_\zeta) = \hat{e}(g_1^{H_7(M_\eta, V_\eta)}, H_6(M_\zeta)^{H_7(M_\zeta, V_\zeta)}) = \hat{e}(g_1, H_6(M_\zeta))^{H_7(M_\zeta, V_\zeta)\cdot H_7(M_\eta, V_\eta)}$

(c) Return 1 if $\hat{e}(\text{CT}_{\zeta 1}, R_\eta) = \hat{e}(\text{CT}_{\eta 1}, R_\zeta)$. Otherwise, return 0

In the following, we present the details of [leftmargin = 0em]

(i) $R_\zeta = \text{CT}_{\zeta 4}/H_8(\hat{e}(\text{CT}_{\zeta 2}, \text{TD}_\zeta)) = H_6(M_\zeta)^{r_\zeta} \cdot H_8(\hat{e}(P_{\text{pub}}, H_2(\text{ID}_\zeta) \cdot H_4(\text{ID}_\zeta, T))^{u_\zeta})/H_8(\hat{e}(g_1^{u_\zeta}, H_2(\text{ID}_\zeta)^s \cdot H_4(\text{ID}_\zeta, T)^s)) = H_6(M_\zeta)^{r_\zeta} \cdot H_8(\hat{e}(g_1^s, H_2(\text{ID}_\zeta) \cdot H_4(\text{ID}_\zeta, T))^{u_\zeta})/H_8(\hat{e}(g_1^{u_\zeta}, H_2(\text{ID}_\zeta)^s \cdot H_4(\text{ID}_\zeta, T)^s)) = H_6(M_\zeta)^{r_\zeta} \cdot H_8(\hat{e}(g_1, H_2(\text{ID}_\zeta) \cdot H_4(\text{ID}_\zeta, T))^{su_\zeta})/H_8(\hat{e}(g_1, H_2(ID_\zeta) \cdot H_4(\text{ID}_\zeta, T))^{su_\zeta}) = H_6(M_\zeta)^{H_7(M_\zeta, V_\zeta)}$

(ii) $R_\eta = \text{CT}_{\eta 4}/H_8(\hat{e}(\text{CT}_{\eta 2}, \text{TD}_\eta)) = H_6(M_\eta)^{r_\eta} \cdot H_8(\hat{e}(P_{\text{pub}}, H_2(\text{ID}_\eta) \cdot H_4(\text{ID}_\eta, T))^{u_\eta})/H_8(\hat{e}(g_1^{u_\eta}, H_2(\text{ID}_\eta)^s \cdot H_4(\text{ID}_\eta, T)^s)) = H_6(M_\eta)^{r_\eta} \cdot H_8(\hat{e}(g_1^s, H_2(\text{ID}_\eta) \cdot H_4(\text{ID}_\eta, T))^{u_\eta})/H_8(\hat{e}(g_1^{u_\eta}, H_2(\text{ID}_\eta)^s \cdot H_4(\text{ID}_\eta, T)^s)) = H_6(M_\eta)^{r_\eta} \cdot H_8(\hat{e}(g_1, H_2(\text{ID}_\eta) \cdot H_4(\text{ID}_\eta, T))^{su_\eta})/H_8(\hat{e}(g_1, H_2(\text{ID}_\eta) \cdot H_4(\text{ID}_\eta, T))^{su_\eta}) = H_6(M_\eta)^{H_7(M_\eta, V_\eta)}$

5. Security Analysis

In this section, we give four theorems to show that the proposed scheme has the IND-ID-CCA security for type I and II adversaries and the OW-ID-CCA security for type III and IV adversaries.

Theorem 5. *If the BDH assumption holds, the proposed RIBEET scheme satisfies the IND-ID-CCA security in the security game. More precisely, suppose that $\mathscr{A}_1$ is a PPT type 1 adversary who has at least ε advantage to break the RIBEET scheme. Then, there exists an algorithm $\mathscr{B}$ to solve the BDH problem with the advantage*

$$\varepsilon' \geq \left(\frac{1}{q_{H_5}}\right)\left[\frac{\varepsilon}{e(q_{IK} + q_T + 1)} - \frac{q_D}{q} - \frac{q_{H_8}}{q}\right], \tag{4}$$

where q_{H_5}, q_{H_8}, q_{IK}, q_T, q_D, and e, respectively, are the number of queries to random oracle H_5, random oracle H_8, Initialkey query, Trapdoor query, Decryption query, and Euler's number.

Proof. An algorithm $\mathscr{B}$ is constructed to solve the BDH problem. The algorithm $\mathscr{B}$ is given a BDH tuple $\langle q, G_1, G_2, G_T, \hat{e}, g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b\rangle$ which is defined in Section 2. The algorithm $\mathscr{B}$ can be seen as a challenger to find the answer of the BDH problem. The answer $A = \hat{e}(g_1, g_2)^{abc}$ can be found by interacting with the PPT type I adversary $\mathscr{A}_1$ in the following security game.

(1) *Setup*: the challenger $\mathscr{B}$ utilizes the BDH tuple to set $P_{\text{pub}} = g_1^a$ and then generates the public system parameters $\text{PSP} = \{q, G_1, G_2, G_T, \hat{e}, g_1, g_2, P_{\text{pub}}, H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$, where H_i is a random oracle for $i = 1, 2, \cdots, 8$. In addition, the system life time $\text{SLT} = \{T_0, T_1, \cdots, T_t\}$ can be generated by the challenger $\mathscr{B}$. Then, $\mathscr{B}$ gives $\mathscr{A}_1$ the public system parameters PSP and system life time SLT. Here, the adversary $\mathscr{A}_1$ can issue queries to each random oracle as follows

(a) H_1query(ID): $\mathscr{A}_1$ can utilize ID to obtain a response to the random oracle H_1 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_1 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, \text{U}_{\text{ID}}, u, rb\rangle$. The response is acquired from the ListH_1 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns U_{ID} as the response if ID exists in a tuple $\langle \text{ID}, U_{\text{ID}}, u, rb\rangle$ from the ListH_1

(ii) Otherwise, $\mathscr{B}$ picks a random value $u \in Z_q^*$ and a random bit $rb \in \{0, 1\}$ to compute

$$U_{\text{ID}} = \begin{cases} g_2^u, & \text{if } rb = 0, \\ g_2^{bu}, & \text{if } rb = 1, \end{cases} \tag{5}$$

where $\Pr[rb = 0] = \delta$ and $\Pr[rb = 1] = 1 - \delta$ (which will be discussed later). Then, $\mathscr{B}$ adds the tuple $\langle \text{ID}, U_{\text{ID}}, u, rb\rangle$ to the ListH_1 and returns U_{ID} to $\mathscr{A}_1$

(b) H_2query(ID): $\mathscr{A}_1$ can utilize ID to obtain a response to the random oracle H_2 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_2 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, V_{\text{ID}}, v, rb\rangle$. The response is acquired from the ListH_2 which is initially empty and can be updated by the following steps:

(a) $\mathscr{B}$ returns V_{ID} as the response if ID exists in a tuple $\langle \text{ID}, V_{\text{ID}}, v, rb\rangle$ from the ListH_2

(b) Otherwise, $\mathscr{B}$ picks a random value $v \in Z_q^*$ and utilizes ID to find rb in the ListH_2. Then, $\mathscr{B}$ computes

$$V_{ID} = \begin{cases} g_2^v, & \text{if } rb = 0, \\ g_2^{bv}, & \text{if } rb = 1, \end{cases} \quad (6)$$

and adds the tuple $\langle \text{ID}, V_{\text{ID}}, v, rb\rangle$ to ListH_2. $\mathscr{B}$ returns V_{ID} to $\mathscr{A}_1$

(c) H_3query(ID, T): $\mathscr{A}_1$ can utilize (ID, T) to obtain a response to the random oracle H_3 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_3 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, T, U_{\text{IDT}}, \eta\rangle$. The response is acquired from the ListH_3 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns $U_{\text{ID}T}$ as the response if (ID, T) exists in a tuple $\langle \text{ID}, T, U_{\text{ID}T}, \eta\rangle$ from the ListH_3

(ii) Otherwise, $\mathscr{B}$ picks a random value $\eta \in Z_q^*$ to compute $U_{\text{ID}T} = g_2^{\eta}$. Then, $\mathscr{B}$ adds the tuple $\langle \text{ID}, T, U_{\text{ID}T}, \eta\rangle$ to the ListH_3 and returns $U_{\text{ID}T}$ to $\mathscr{A}_1$

(d) H_4query(ID, T): $\mathscr{A}_1$ can utilize (ID, T) to obtain a response to the random oracle H_4 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_4 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, T, V_{\text{ID}T}, \zeta\rangle$. The response is acquired from the ListH_4 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns $V_{\text{ID}T}$ as the response if (ID, T) exists in a tuple $\langle \text{ID}, T, V_{\text{ID}T}, \zeta\rangle$ from the ListH_4

(ii) Otherwise, $\mathscr{B}$ picks a random value $\zeta \in Z_q^*$ to compute $V_{\text{ID}T} = g_2^{\zeta}$. Then, $\mathscr{B}$ adds the tuple $\langle \text{ID}, T, V_{\text{ID}T}, \zeta\rangle$ to the ListH_4 and returns $V_{\text{ID}T}$ to $\mathscr{A}_1$

(e) H_5query$(W, \text{CT}_1, \text{CT}_2)$: $\mathscr{A}_1$ can utilize $(W, \text{CT}_1, \text{CT}_2)$ to obtain a response to the random oracle H_5 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_5 which is composed of tuples, and the format of the tuple is $\langle W, \text{CT}_1, \text{CT}_2, \omega\rangle$. The response is acquired from the ListH_5 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns ω as the response if $(W, \text{CT}_1, \text{CT}_2)$ exists in a tuple $\langle W, \text{CT}_1, \text{CT}_2, \omega\rangle$ from the ListH_5

(ii) Otherwise, $\mathscr{B}$ picks a random value $\omega \in \{0, 1\}^{\lambda+l}$ and adds the tuple $\langle W, \text{CT}_1, \text{CT}_2, \omega\rangle$ to the ListH_5. Then, $\mathscr{B}$ returns ω to $\mathscr{A}_1$

(f) H_6query(M): $\mathscr{A}_1$ can utilize M to obtain a response to the random oracle H_6 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_6 which is composed of tuples, and the format of the tuple is $\langle M, Q\rangle$. The response is acquired from the ListH_6 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns Q as the response if M exists in a tuple $\langle M, Q\rangle$ from the ListH_6

(ii) Otherwise, $\mathscr{B}$ picks a random point $Q \in G_2$ and adds the tuple $\langle M, Q\rangle$ to the ListH_6. Then, $\mathscr{B}$ returns Q to $\mathscr{A}_1$

(g) H_7query(M, V): $\mathscr{A}_1$ can utilize (M, V) to obtain a response to the random oracle H_7 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_7 which is composed of tuples, and the format of the tuple is $\langle M, V, \gamma\rangle$. The response is acquired from the ListH_7 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns γ as the response if (M, V) exists in a tuple $\langle M, V, \gamma\rangle$ from the ListH_7

(ii) Otherwise, $\mathscr{B}$ picks a random value $\gamma \in Z_q^*$ and adds the tuple $\langle M, V, \gamma\rangle$ to the ListH_7. Then, $\mathscr{B}$ returns γ to $\mathscr{A}_1$

(h) H_8query(N): $\mathscr{A}_1$ can utilize N to obtain a response to the random oracle H_8 from the challenger $\mathscr{B}$.

To obtain the response, $\mathscr{B}$ maintains a list, called ListH_8 which is composed of tuples, and the format of the tuple is $\langle N, S\rangle$. The response is acquired from the ListH_8 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns S as the response if N exists in a tuple $\langle N, S\rangle$ from the ListH_8

(ii) Otherwise, $\mathscr{B}$ picks a random point $S \in G_2$ and adds the tuple $\langle N, S\rangle$ to the ListH_8. Then, $\mathscr{B}$ returns S to $\mathscr{A}_1$

(2) Phase 1: the adversary $\mathscr{A}_1$ can, respectively, utilize ID, (ID, T), $(\mathrm{ID}, T, \mathrm{CT})$ and (ID, T) to issue the InitialKey query, Timekey query, Decryption query, and Trapdoor query. The response to each query can be obtained as follows

(a) InitialKey query(ID): $\mathscr{A}_1$ utilizes ID to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle \mathrm{ID}, \mathrm{U}_{\mathrm{ID}}, u, rb\rangle$ and $\langle \mathrm{ID}, V_{\mathrm{ID}}, v, rb\rangle$ from the ListH_1 and the ListH_2 according to ID. If $rb = 1$, $\mathscr{B}$ interrupts this game. If $rb = 0$, $\mathscr{B}$ use u and v to define $I\mathrm{K}_{\mathrm{ID}} = (\mathrm{IK}_{\mathrm{ID1}}, \mathrm{IK}_{\mathrm{ID2}}) = ((g_2^a)^u, (g_2^a)^v)$. Then $\mathscr{B}$ returns $\mathrm{IK}_{\mathrm{ID}}$ as the user initial key to $\mathscr{A}_1$

(b) Timekey query(ID, T): $\mathscr{A}_1$ utilizes (ID, T) to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle \mathrm{ID}, T, U_{\mathrm{ID}T}, \eta\rangle$ and $\langle \mathrm{ID}, T, V_{\mathrm{ID}T}, \zeta\rangle$ from the ListH_3 and the ListH_4 according to (ID, T). $\mathscr{B}$ use η and ζ to define $\mathrm{TK}_{\mathrm{ID}} = (\mathrm{TK}_{\mathrm{ID1}}, \mathrm{TK}_{\mathrm{ID2}}) = ((g_2^a)^\eta, (g_2^a)^\zeta)$. Then, $\mathscr{B}$ returns $\mathrm{TK}_{\mathrm{ID}}$ as the user time key to $\mathscr{A}_1$

(c) Decryption query(ID, T, CT): $\mathscr{A}_1$ utilizes (ID, T, CT) to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle \mathrm{ID}, U_{\mathrm{ID}}, u, rb\rangle$, $\langle \mathrm{ID}, V_{\mathrm{ID}}, v, rb\rangle$, $\langle \mathrm{ID}, T, U_{\mathrm{ID}T}, \eta\rangle$, and $\langle \mathrm{ID}, T, V_{\mathrm{ID}T}, \zeta\rangle$ from the ListH_1, ListH_2, ListH_3, and the ListH_4 according to ID and T. The response of this query is acquired from these lists by performing the following tasks

(i) If $rb = 0$, $\mathscr{B}$, respectively, uses ID and (ID, T) to run InitialKeyquery and Timekey query to obtain $\mathrm{IK}_{\mathrm{ID}}$ and $\mathrm{TK}_{\mathrm{ID}}$. Then, $\mathscr{B}$ utilizes $\mathrm{IK}_{\mathrm{ID}}$, $\mathrm{TK}_{\mathrm{ID}}$, and CT to run the Decryption algorithm to produce the message M which is sent to $\mathscr{A}_1$

(ii) If $rb = 1$, $\mathscr{B}$ utilizes CT_1 and CT_2, which are from $\mathrm{CT} = (\mathrm{CT}_1, \mathrm{CT}_2, \mathrm{CT}_3, \mathrm{CT}_4)$, to find the corresponding tuple $\langle W, \mathrm{CT}_1, \mathrm{CT}_2, \omega\rangle$ from the ListH_5. Then, $M' || V' = \mathrm{CT}_3 \oplus \omega$ can be computed by using CT_3 and ω. Further, $\mathscr{B}$ utilizes M' and V' to find the corresponding tuples $\langle M, V, \gamma\rangle$ from the ListH_7 and $\langle M, Q\rangle$ from the ListH_6. Obviously, γ and Q can be obtained. If S can be found in the corresponding tuple $\langle N, S\rangle$ from the ListH_8 such that $\mathrm{CT}_4 = Q^\gamma \cdot S$ holds, $\mathscr{B}$ will confirm whether $\mathrm{CT}_1 = g_1^\gamma$ holds. If $\mathrm{CT}_1 = g_1^\gamma$, the message M' is sent to $\mathscr{A}_1$

(d) Trapdoor query(ID, T): $\mathscr{A}_1$ utilizes (ID, T) to issue the query, while $\mathscr{B}$, respectively, uses ID and (ID, T) to run InitialKey query and Timekey query to obtain $\mathrm{IK}_{\mathrm{ID}} = (\mathrm{IK}_{\mathrm{ID1}}, \mathrm{IK}_{\mathrm{ID2}})$ and $\mathrm{TK}_{\mathrm{ID}} = (\mathrm{TK}_{\mathrm{ID1}}, \mathrm{TK}_{\mathrm{ID2}})$. Then, $\mathscr{B}$ utilizes $I\mathrm{K}_{\mathrm{ID2}}$ and $\mathrm{TK}_{\mathrm{ID2}}$ to produce the trapdoor $\mathrm{TD}_{\mathrm{ID}} = \mathrm{IK}_{\mathrm{ID2}} \cdot \mathrm{TK}_{\mathrm{ID2}}$ which is sent to $\mathscr{A}_1$

(3) Challenge: when the phase 1 is over, $\mathscr{A}_1$ outputs a tuple $\langle \mathrm{ID}^*, T^*, M_0^*, M_1^*\rangle$ as the target of the challenge. $\mathscr{B}$ utilizes ID^* to find the corresponding tuples $\langle \mathrm{ID}, U_{\mathrm{ID}}, u, rb\rangle$ from the ListH_1. If $rb = 0$, $\mathscr{B}$ interrupts this game. If $rb = 1$, $\mathscr{B}$ randomly selects $\mathfrak{b} \in \{0, 1\}$ and $V \in \{0, 1\}^l$ to run H_7 query with $M_{\mathfrak{b}}^*$ and V. Then, γ can be obtained. $\mathscr{B}$ utilizes γ to set $\mathrm{CT}_1^* = g_1^\gamma$. In addition, $\mathscr{B}$ sets $\mathrm{CT}_2^* = g_1^c$, while a random value $\mathrm{CT}_3 \in \{0, 1\}^{\lambda + l}$ and a random point $\mathrm{CT}_4^* \in G_2$ are chosen. Finally, the challenge ciphertext $\mathrm{CT}^* = (\mathrm{CT}_1^*, \mathrm{CT}_2^*, \mathrm{CT}_3^*, \mathrm{CT}_4^*)$ is sent to $\mathscr{A}_1$

(4) Phase 2: $\mathscr{A}_1$ can issue the same query as phase 1, but it must be under the condition of $\mathrm{ID} \neq \mathrm{ID}^*$ and $\mathrm{CT} \neq \mathrm{CT}^*$

(5) Guess: $\mathscr{A}_1$ responds to $\mathscr{B}$ with a guess $\mathfrak{b}' \in \{0, 1\}$. If $\mathfrak{b}' \neq \mathfrak{b}$, $\mathscr{B}$ responds with failure and terminates. Otherwise, $\mathscr{A}_1$ wins the game. Then, $\mathscr{B}$ randomly selects a tuple $\langle W^*, \mathrm{CT}_1^*, \mathrm{CT}_2^*, \omega^*\rangle$ from the ListH_5 and calculates $H_5(\widehat{e}(g_1, g_2)^{abcu^*} \cdot \widehat{e}(g_1^{ac}, g_2)^{\eta^*}, CT_1^*, CT_2^*) = (M_\rho^* || V) \oplus CT_3^*$, where $\widehat{e}(g_1, g_2)^{abcu^*} \cdot \widehat{e}(g_1^{ac}, g_2)^{\eta^*} = W^*$. Hence, $\mathscr{B}$ can output the BDH solution $A = (W^* / \widehat{e}(g_1^{ac}, g_2)^{\eta^*})^{u^{*-1}}$ due to $\widehat{e}(g_1, g_2)^{abc} = (W^* / \widehat{e}(g_1^{ac}, g_2)^{\eta^*})^{u^{*-1}}$

□

Analysis. Let us start with two cases, namely, the simulation of H_i query for $i = 1, 2, \cdots, 8$ and the simulation of decryption query. For the H_1, H_2, H_3, H_4, H_6, and H_7 queries, it is obvious that the simulations are perfect because there exists no relationship between the constructions of these queries and the solution of the BDH problem. For the H_5 and H_8 queries, we consider two events $E_{H_5}^*$ and $E_{H_8}^*$ which, respectively, issues the H_5 query with $(\widehat{e}$

$(g_1, g_2)^{abcu^*}\widehat{e}(g_1^{ac}, g_2)^{\eta^*}, CT_1^*, CT_2^*)$ and the H_8 query with $\widehat{e}(g_1, g_2)^{abcv^*} \cdot \widehat{e}(g_1^{a}, g_2)^{c\zeta^*}$. We say that the simulations of H_5 and H_8 queries are perfect if $E_{H_5}^*$ and $E_{H_8}^*$ do not happen. For the decryption query, we consider an event E_{DecErr} where the challenger $\mathscr{B}$ cannot decrypt the ciphertext. Assume that q_D is the number of decryption query. Then, we obtain $\Pr[E_{\text{DecErr}}] \leq q_D/q$

Next, we discuss an event E which states that the simulation of this security game will not be interrupted. Here, we can obtain $E = (E_{H_5}^* \vee E_{H_8}^* \vee E_{\text{DecErr}}) | \neg E_{\text{Abort}}$, where E_{Abort} is defined as the event that the challenger $\mathscr{B}$ interrupts this security game. Since $\mathscr{B}$ guesses $\mathfrak{b}'$ with the advantage $\leq 1/2$, the $\Pr[\mathfrak{b} = \mathfrak{b}' | \neg E] \leq 1/2$ can be obtained if E does not occur. Further, we have

$$\Pr\left[\mathfrak{b} = \mathfrak{b}'\right] = \Pr\left[\mathfrak{b} = \mathfrak{b}' \mid E\right]\Pr[E] + \Pr\left[\mathfrak{b} = \mathfrak{b}' | \neg E\right]\Pr[\neg E] \leq \Pr[E] + \left(\frac{1}{2}\right) \cdot \Pr[\neg E] = \Pr[E] + \frac{1}{2} \cdot (1 - \Pr[E]) = \frac{1}{2} \cdot \Pr[E] + \frac{1}{2}. \quad (7)$$

According to above inequality, $\varepsilon = \Pr[\mathfrak{b} = \mathfrak{b}'] - 1/2$ and $E = (E_{H_5}^* \vee E_{H_8}^* \vee E_{\text{DecErr}}) | \neg E_{\text{Abort}}$, we have

$$\varepsilon = \Pr\left[\mathfrak{b} = \mathfrak{b}'\right] - \frac{1}{2} \leq \Pr[E] \leq \frac{\Pr\left[E_{H_5}^*\right] + \Pr\left[E_{H_8}^*\right] + \Pr[E_{\text{DecErr}}]}{\Pr[\neg E_{\text{Abort}}]}. \quad (8)$$

Moreover, we obtain

$$\Pr\left[E_{H_5}^*\right] \geq \varepsilon \cdot \Pr[\neg E_{\text{Abort}}] - \Pr[E_{\text{DecErr}}] - \Pr\left[E_{H_8}^*\right]. \quad (9)$$

Since $\Pr[\neg E_{\text{Abort}}] = \delta^{q_{\text{IK}} + q_T}(1 - \delta)$, we can gain $\Pr[\neg E_{\text{Abort}}] \geq 1/e(q_{\text{IK}} + q_T + 1)$ when $\delta = 1 - (1/(q_{\text{IK}} + q_T + 1))$. Then, we have

$$\Pr\left[E_{H_5}^*\right] \geq \frac{\varepsilon}{e(q_{\text{IK}} + q_T + 1)} - \frac{q_D}{q} - \frac{q_{H_8}}{q}. \quad (10)$$

Here, the adversary $\mathscr{A}_1$ can distinguish the target ciphertext CT^* is the real one when $E_{H_5}^*$ occurs. In addition, the tuple $\langle \widehat{e}(g_1, g_2)^{abcu^*} \cdot \widehat{e}(g_1^{ac}, g_2)^{\eta^*}, \text{CT}_1^*, \text{CT}_2^* \rangle$ has been added in the ListH_5. If the challenger $\mathscr{B}$ picks the correct tuple from the ListH_5, $\mathscr{B}$ wins this security game. Meanwhile, the advantage of solving the BDH problem is

$$\varepsilon' \geq \frac{1}{q_{H_5}} \Pr\left[E_{H_5}^*\right] \geq \frac{1}{q_{H_5}} \left[\frac{\varepsilon}{e(q_{\text{IK}} + q_T + 1)} - \frac{q_D}{q} - \frac{q_{H_8}}{q}\right]. \quad (11)$$

Theorem 6. *If the BDH assumption holds, the proposed RIBEET scheme satisfies the IND-ID-CCA security in the security game. More precisely, suppose that $\mathscr{A}_2$ is a PPT type 2 adversary who has at least ε advantage to break the RIBEET scheme. Then, there exists an algorithm $\mathscr{B}$ to solve the BDH problem with the advantage*

$$\varepsilon' \geq \left(\frac{1}{q_{H_5}}\right) \left[\frac{\varepsilon}{e(q_{TK} + q_T + 1)} - \frac{q_D}{q} - \frac{q_{H_8}}{q}\right], \quad (12)$$

where q_{H_5}, q_{H_8}, q_{TK}, q_T, q_D, and e, respectively, are the number of queries to random oracle H_5, random oracle H_8, Timekey query, Trapdoor query, Decryption query, and Euler's number.

Proof. An algorithm $\mathscr{B}$ is constructed to solve the BDH problem. The algorithm $\mathscr{B}$ is given a BDH tuple $\langle q, G_1, G_2, G_T, \widehat{e}, g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b \rangle$ which is defined in Section 2. The algorithm $\mathscr{B}$ can be seen as a challenger to find the answer of the BDH problem. The answer $A = \widehat{e}(g_1, g_2)^{abc}$ can be found by interacting with the PPT type II adversary $\mathscr{A}_2$ in the following security game.

(1) Setup: the challenger $\mathscr{B}$ utilizes the BDH tuple to set $P_{\text{pub}} = g_1^a$ and then generates the public system parameters $\text{PSP} = \{q, G_1, G_2, G_T, \widehat{e}, g_1, g_2, P_{\text{pub}}, H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$, where H_i is a random oracle for $i = 1, 2, \cdots, 8$. In addition, the system life time $\text{SLT} = \{T_0, T_1, \cdots, T_t\}$ can be generated by the challenger $\mathscr{B}$. Then, $\mathscr{B}$ gives $\mathscr{A}_2$ the public system parameters PSP and system life time SLT. Here, the adversary $\mathscr{A}_2$ can issue queries to each random oracle as follows

(a) H_1query(ID): $\mathscr{A}_2$ can utilize ID to obtain a response to the random oracle H_1 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_1 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, \text{U}_{\text{ID}}, u \rangle$. The response is acquired from the ListH_1 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns U_{ID} as the response if ID exists in a tuple $\langle \text{ID}, \text{U}_{\text{ID}}, u \rangle$ from the ListH_1

(ii) Otherwise, $\mathscr{B}$ picks a random value $u \in Z_q^*$ to compute $U_{\text{ID}} = g_2^u$. Then, $\mathscr{B}$ adds the tuple $\langle \text{ID}, \text{U}_{\text{ID}}, u \rangle$ to the ListH_1 and returns U_{ID} to $\mathscr{A}_2$

(b) H_2query(ID): $\mathscr{A}_2$ can utilize ID to obtain a response to the random oracle H_2 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_2 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, V_{\text{ID}}, v \rangle$. The response is acquired from the ListH_2 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns V_{ID} as the response if ID exists in a tuple $\langle \text{ID}, V_{\text{ID}}, v \rangle$ from the ListH_2

(ii) Otherwise, $\mathscr{B}$ picks a random value $v \in Z_q^*$ to compute $V_{\text{ID}} = g_2^v$. Then, $\mathscr{B}$ adds the tuple $\langle \text{ID}, V_{\text{ID}}, v \rangle$ to the ListH_2 and returns V_{ID} to $\mathscr{A}_2$

(c) H_3query(ID, T): $\mathscr{A}_2$ can utilize (ID, T) to obtain a response to the random oracle H_3 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_3 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, T, U_{\text{ID}T}, \eta, rb \rangle$. The response is acquired from the ListH_3 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns $U_{\text{ID}T}$ as the response if (ID, T) exists in a tuple $\langle \text{ID}, T, U_{\text{ID}T}, \eta, rb \rangle$ from the ListH_3

(ii) Otherwise, $\mathscr{B}$ picks a random value $\eta \in Z_q^*$ and a random bit $rb \in \{0, 1\}$ to compute

$$U_{\text{ID}T} = \begin{cases} g_2^{\eta}, & \text{if } rb = 0, \\ g_2^{b\eta}, & \text{if } rb = 1, \end{cases} \quad (13)$$

where $\Pr[rb = 0] = \delta$ and $\Pr[rb = 1] = 1 - \delta$ (which will be discussed later). Then, $\mathscr{B}$ adds the tuple $\langle \text{ID}, T, U_{\text{ID}T}, \eta, rb \rangle$ to the ListH_3 and returns $U_{\text{ID}T}$ to $\mathscr{A}_2$

(d) H_4query(ID, T): $\mathscr{A}_2$ can utilize (ID, T) to obtain a response to the random oracle H_4 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_4 which is composed of tuples, and the format of the tuple is $\langle \text{ID}, T, V_{\text{ID}T}, \zeta, rb \rangle$. The response is acquired from the ListH_4 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns $V_{\text{ID}T}$ as the response if (ID, T) exists in a tuple $\langle \text{ID}, T, V_{\text{ID}T}, \zeta, rb \rangle$ from the ListH_4

(ii) Otherwise, $\mathscr{B}$ picks a random value $\zeta \in Z_q^*$ and utilizes (ID, T) to find rb in the ListH_4. Then, $\mathscr{B}$ computes

$$V_{\text{ID}T} = \begin{cases} g_2^{\zeta}, & \text{if } rb = 0, \\ g_2^{b\zeta}, & \text{if } rb = 1, \end{cases} \quad (14)$$

and add the tuple $\langle \text{ID}, T, V_{\text{ID}T}, \zeta, rb \rangle$ to ListH_4. $\mathscr{B}$ returns $V_{\text{ID}T}$ to $\mathscr{A}_2$

(e) H_5query(W, CT$_1$, CT$_2$): $\mathscr{A}_2$ can utilize (W, CT$_1$, CT$_2$) to obtain a response to the random oracle H_5 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_5 which is composed of tuples, and the format of the tuple is $\langle W, \text{CT}_1, \text{CT}_2, \omega \rangle$. The response is acquired from the ListH_5 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns ω as the response if (W, CT$_1$, CT$_2$) exists in a tuple $\langle W, \text{CT}_1, \text{CT}_2, \omega \rangle$ from the ListH_5

(ii) Otherwise, $\mathscr{B}$ picks a random value $\omega \in \{0, 1\}^{\lambda+l}$ and adds the tuple $\langle W, CT_1, CT_2, \omega \rangle$ to the ListH_5. Then, $\mathscr{B}$ returns ω to $\mathscr{A}_2$

(f) H_6query(M): $\mathscr{A}_2$ can utilize M to obtain a response to the random oracle H_6 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_6 which is composed of tuples, and the format of the tuple is $\langle M, Q \rangle$. The response is acquired from the ListH_6 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns Q as the response if M exists in a tuple $\langle M, Q \rangle$ from the ListH_6

(ii) Otherwise, $\mathscr{B}$ picks a random point $Q \in G_2$ and adds the tuple $\langle M, Q \rangle$ to the ListH_6. Then, $\mathscr{B}$ returns Q to $\mathscr{A}_2$

(g) H_7query(M, V): $\mathscr{A}_2$ can utilize (M, V) to obtain a response to the random oracle H_7 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_7 which is composed of tuples, and the format of the tuple is $\langle M, V, \gamma \rangle$. The response is acquired from the ListH_7 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns γ as the response if (M, V) exists in a tuple $\langle M, V, \gamma \rangle$ from the ListH_7

(ii) Otherwise, $\mathscr{B}$ picks a random value $\gamma \in Z_q^*$ and adds the tuple $\langle M, V, \gamma \rangle$ to the ListH_7. Then, $\mathscr{B}$ returns γ to $\mathscr{A}_2$

(h) H_8query(N): $\mathscr{A}_2$ can utilize N to obtain a response to the random oracle H_8 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called ListH_8 which is composed of tuples, and the format of the tuple is $\langle N, S \rangle$. The response is acquired from the List H_8 which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns S as the response if N exists in a tuple $\langle N, S \rangle$ from the ListH_8

(ii) Otherwise, $\mathscr{B}$ picks a random point $S \in G_2$ and adds the tuple $\langle N, S \rangle$ to the ListH_8. Then, $\mathscr{B}$ returns S to $\mathscr{A}_2$

(2) Phase 1: the adversary $\mathscr{A}_2$ can, respectively, utilize ID, (ID, T), (ID, T, CT), and (ID, T) to issue the InitialKey query, Timekey query, Decryption query, and Trapdoor query. The response to each query can be obtained as follows

(a) InitialKey query(ID): $\mathscr{A}_2$ utilizes ID to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle \mathrm{ID}, \mathrm{U}_{\mathrm{ID}}, u \rangle$ and $\langle \mathrm{ID}, V_{\mathrm{ID}}, v \rangle$ from the ListH_1 and the ListH_2 according to ID. $\mathscr{B}$ use u and v to define $\mathrm{IK}_{\mathrm{ID}} = (\mathrm{IK}_{\mathrm{ID1}}, \mathrm{IK}_{\mathrm{ID2}}) = ((g_2^a)^u, (g_2^a)^v)$. Then, $\mathscr{B}$ returns $\mathrm{IK}_{\mathrm{ID}}$ as the user initial key to $\mathscr{A}_1$

(b) Timekey query(ID, T): $\mathscr{A}_2$ utilizes (ID, T) to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle \mathrm{ID}, T, U_{\mathrm{ID}T}, \eta, rb \rangle$ and $\langle \mathrm{ID}, T, V_{\mathrm{ID}T}, \zeta, rb \rangle$ from the ListH_3 and the ListH_4 according to (ID, T). If $rb = 1$, $\mathscr{B}$ interrupts this game. If $rb = 0$, $\mathscr{B}$ use η and ζ to define $\mathrm{TK}_{\mathrm{ID}} = (\mathrm{TK}_{\mathrm{ID1}}, \mathrm{TK}_{\mathrm{ID2}}) = ((g_2^a)^\eta, (g_2^a)^\zeta)$. Then, $\mathscr{B}$ returns $\mathrm{TK}_{\mathrm{ID}}$ as the user time key to $\mathscr{A}_2$

(c) Decryption query(ID, T, CT): $\mathscr{A}_2$ utilizes (ID, T, CT) to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle \mathrm{ID}, U_{\mathrm{ID}}, u \rangle$, $\langle \mathrm{ID}, V_{\mathrm{ID}}, v \rangle$, $\langle \mathrm{ID}, T, U_{\mathrm{ID}T}, \eta, rb \rangle$, and $\langle \mathrm{ID}, T, V_{\mathrm{ID}T}, \zeta, rb \rangle$ from the ListH_1, ListH_2, ListH_3, and ListH_4 according to ID and T. The response of this query is acquired from these lists by performing the following tasks

(i) If $rb = 0$, $\mathscr{B}$, respectively, uses ID and (ID, T) to run InitialKeyquery and Timekey query to obtain $\mathrm{IK}_{\mathrm{ID}}$ and $\mathrm{TK}_{\mathrm{ID}}$. Then, $\mathscr{B}$ utilizes $\mathrm{IK}_{\mathrm{ID}}$, $\mathrm{TK}_{\mathrm{ID}}$, and CT to run the Decryption algorithm to produce the message M which is sent to $\mathscr{A}_2$

(ii) If $rb = 1$, $\mathscr{B}$ utilizes CT_1 and CT_2, which are from $\mathrm{CT} = (\mathrm{CT}_1, \mathrm{CT}_2, \mathrm{CT}_3, \mathrm{CT}_4)$, to find the corresponding tuple $\langle W, \mathrm{CT}_1, \mathrm{CT}_2, \omega \rangle$ from the ListH_5. Then, $M' || V' = \mathrm{CT}_3 \oplus \omega$ can be computed by using CT_3 and ω. Further, $\mathscr{B}$ utilizes M' and V' to find the corresponding tuples $\langle M, V, \gamma \rangle$ from the ListH_7 and $\langle M, Q \rangle$ from the ListH_6. Obviously, γ and Q can be obtained. If S can be found in the corresponding tuple $\langle N, S \rangle$ from the ListH_8 such that $\mathrm{CT}_4 = Q^\gamma \cdot S$ holds, $\mathscr{B}$ will confirm whether $\mathrm{CT}_1 = g_1^\gamma$ holds. If $\mathrm{CT}_1 = g_1^\gamma$, the message M' is sent to $\mathscr{A}_2$

(1) Trapdoor query(ID, T): $\mathscr{A}_2$ utilizes (ID, T) to issue the query, while $\mathscr{B}$, respectively, uses ID and (ID, T) to run InitialKey query and Timekey query to obtain $\mathrm{IK}_{\mathrm{ID}} = (\mathrm{IK}_{\mathrm{ID1}}, \mathrm{IK}_{\mathrm{ID2}})$ and $\mathrm{TK}_{\mathrm{ID}} = (\mathrm{TK}_{\mathrm{ID1}}, \mathrm{T}$ $\mathrm{K}_{\mathrm{ID2}})$. Then, $\mathscr{B}$ utilizes $\mathrm{IK}_{\mathrm{ID2}}$ and $\mathrm{TK}_{\mathrm{ID2}}$ to produce the trapdoor $\mathrm{TD}_{\mathrm{ID}} = \mathrm{IK}_{\mathrm{ID2}} \cdot \mathrm{TK}_{\mathrm{ID2}}$ which is sent to $\mathscr{A}_2$

(3) Challenge: when phase 1 is over, $\mathscr{A}_2$ outputs a tuple $\langle \mathrm{ID}^*, T^*, M_0^*, M_1^* \rangle$ as the target of the challenge. $\mathscr{B}$ utilizes (ID^*, T^*) to find the corresponding tuples $\langle$ ID, T, $U_{\mathrm{ID}T}, \eta, rb \rangle$ from the ListH_3. If $rb = 0$, $\mathscr{B}$ interrupts this game. If $rb = 1$, $\mathscr{B}$ randomly selects $\mathfrak{b} \in \{0, 1\}$ and $V \in \{0, 1\}^l$ to run H_7 query with $M_{\mathfrak{b}}^*$ and V. Then, γ can be obtained. $\mathscr{B}$ utilizes γ to set $\mathrm{CT}_1^* = g_1^\gamma$. In addition, $\mathscr{B}$ sets $\mathrm{CT}_2^* = g_1^c$, while a random value $\mathrm{CT}_3 \in \{0, 1\}^{\lambda + l}$ and a random point $\mathrm{CT}_4^* \in G_2$ are chosen. Finally, the challenge ciphertext C $\mathrm{T}^* = (\mathrm{CT}_1^*, \mathrm{CT}_2^*, \mathrm{CT}_3^*, \mathrm{CT}_4^*)$ is sent to $\mathscr{A}_2$

(2) Phase 2: $\mathscr{A}_2$ can issue the same query as phase 1, but it must be under the condition of $\mathrm{ID} \neq \mathrm{ID}^*$ and $\mathrm{CT} \neq \mathrm{CT}^*$

(3) Guess: $\mathscr{A}_2$ responds to $\mathscr{B}$ with a guess $\mathfrak{b}' \in \{0, 1\}$. If $\mathfrak{b}' \neq \mathfrak{b}$, $\mathscr{B}$ responds with failure and terminates. Otherwise, $\mathscr{A}_2$ wins the game. Then, $\mathscr{B}$ randomly selects a tuple $\langle W^*, \mathrm{CT}_1^*, \mathrm{CT}_2^*, \omega^* \rangle$ from the ListH_5 and outputs the BDH solution $A = \left(W^* / \hat{e}(g_1^{ac}, g_2)^{u^*} \right)^{\eta^{*-1}}$ due to $\hat{e}(g_1, g_2)^{abc} = \left(W^* / \hat{e}(g_1^{ac}, g_2)^{u^*} \right)^{\eta^{*-1}}$

The security analysis is similar to Theorem 5. We obtain that $\mathscr{B}$'s advantage to solve the BDH problem is $\varepsilon' \geq (1/$

q_{H_5})$\Pr[E^*_{H_5}] \geq (1/q_{H_5})[(\varepsilon/e(q_{\mathrm{TK}} + q_T + 1)) - (q_D/q) - (q_{H_8}/q)$]. □

Theorem 7. *If the BDH assumption holds, the proposed RIBEET scheme satisfies the OW-ID-CCA security in the security game. More precisely, suppose that $\mathscr{A}_3$ is a PPT type 3 adversary who has at least ε advantage to break the RIBEET scheme. Then, there exists an algorithm $\mathscr{B}$ to solve the BDH problem with the advantage*

$$\varepsilon' \geq \left(\frac{1}{q_{H_5}}\right)\left[\frac{\varepsilon - \left(1/2^{\lambda}\right)}{e(q_{IK} + 1)}\right] - \frac{q_D}{q}, \tag{15}$$

where q_{H_5}, q_{IK}, q_D, and e, respectively, are the number of queries to random oracle H_5, Initialkey query, Decryption query, and Euler's number.

Proof. An algorithm $\mathscr{B}$ is constructed to solve the BDH problem. The algorithm $\mathscr{B}$ is given a BDH tuple $\langle q, G_1, G_2, G_T, \widehat{e}, g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b\rangle$ which is defined in Section 2. The algorithm $\mathscr{B}$ can be seen as a challenger to find the answer of the BDH problem. The answer $A = \widehat{e}(g_1, g_2)^{abc}$ can be found by interacting with the PPT type III adversary $\mathscr{A}_3$ in the following security game.

(i) Setup: The challenger $\mathscr{B}$ utilizes the BDH tuple to set $P_{pub} = g_1^a$, and then generates the public system parameters $PSP = \{q, G_1, G_2, G_T, \widehat{e}, g_1, g_2, P_{pub}, H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$, where H_i is a random oracle for $i = 1, 2, \cdots, 8$. In addition, the system life time $SLT = \{T_0, T_1, \cdots, T_t\}$ can be generated by the challenger $\mathscr{B}$. Then $\mathscr{B}$ gives $\mathscr{A}_3$ the public system parameters *PSP* and system life time *SLT*. Here, the adversary $\mathscr{A}_3$ can issue queries to each random oracle as below

(a) $H_1 query(ID)$: $\mathscr{B}$ answers $\mathscr{A}_3$ in the same form as the proof of Theorem 5

(b) $H_2 query(ID)$: $\mathscr{A}_3$ can utilize *ID* to obtain a response to the random oracle H_2 from the challenger $\mathscr{B}$. To obtain the response, $\mathscr{B}$ maintains a list, called $ListH_2$ which is composed of tuples, and the format of the tuple is $\langle ID, V_{ID}, v, rb\rangle$. The response is acquired from the $ListH_2$ which is initially empty and can be updated by the following steps

(i) $\mathscr{B}$ returns V_{ID} as the response if *ID* exists in a tuple $\langle ID, V_{ID}, v, rb\rangle$ from the $ListH_2$

(ii) Otherwise, $\mathscr{B}$ picks a random value $v \in Z_q^*$ and utilizes *ID* to find *rb* in the $ListH_2$. Then $\mathscr{B}$ computes $V_{ID} = g_2^v$ and adds the tuple $\langle ID, V_{ID}, v, rb\rangle$ to $ListH_2$. $\mathscr{B}$ returns V_{ID} to $\mathscr{A}_3$

(c) $H_3 - H_8$ *queries*: $\mathscr{B}$ answers $\mathscr{A}_3$ in the same form as the proof of Theorem 5

(ii) *Phase 1*: The adversary $\mathscr{A}_3$ can, respectively, utilize *ID*, (ID, T), (ID, T, CT) and (ID, T) to issue the *InitialKey query*, *Timekey query*, *Decryption query* and *Trapdoor query*. The response to each query can be obtained as follows

(1) *InitialKey query* (ID): $\mathscr{B}$ answers $\mathscr{A}_3$ in the same form as the proof of Theorem 5

(2) *Timekey query* (ID, T): $\mathscr{B}$ answers $\mathscr{A}_3$ in the same form as the proof of Theorem 5

(3) *Decryption query* (ID, T, CT): $\mathscr{A}_3$ utilizes (ID, T, CT) to issue the query, while $\mathscr{B}$, respectively, finds the corresponding tuples $\langle ID, U_{ID}, u, rb\rangle$, $\langle ID, V_{ID}, v, rb\rangle$, $\langle ID, T, U_{IDT}, \eta\rangle$ and $\langle ID, T, V_{IDT}, \zeta\rangle$ from the $ListH_1$, $ListH_2$, $ListH_3$ and the $ListH_4$ according to *ID* and *T*. The response of this query is acquired from these lists by performing the following tasks

(i) If $rb = 0$, $\mathscr{B}$, respectively, uses *ID* and (ID, T) to run *InitialKeyquery* and *Timekey query* to obtain IK_{ID} and TK_{ID}. Then $\mathscr{B}$ utilizes IK_{ID}, TK_{ID} and *CT* to run *Decryption* algorithm to produce the message *M* which is sent to $\mathscr{A}_3$

(ii) If $rb = 1$, $\mathscr{B}$ utilizes CT_1 and CT_2, which are from $CT = (CT_1, CT_2, CT_3, CT_4)$, to find the corresponding tuple $\langle W, CT_1, CT_2, \omega\rangle$ from the $ListH_5$. Then $M' || V' = CT_3 \oplus \omega$ can be computed by using CT_3 and ω. Further, $\mathscr{B}$ utilizes M' and V' to find the corresponding tuples $\langle M, V, \gamma\rangle$ from the $ListH_7$ and $\langle M, Q\rangle$ from the $ListH_6$. Obviously, γ and *Q* can be obtained. After that, $\mathscr{B}$ utilizes *ID* and (ID, T) to run *InitialKey query* and *Timekey query* to obtain IK_{ID2} and TK_{ID2} and computes $IK_{ID2} \cdot TK_{ID2} = g_2^{a(v+\zeta)}$. If *S* can be found in the corresponding tuple $\langle \widehat{e}(CT_2, g_2^{a(v+\zeta)}), S\rangle$ from the $ListH_8$ such that $CT_4 = Q^{\gamma} \cdot S$ holds, $\mathscr{B}$ will confirm whether $CT_1 = g_1^{\gamma}$ holds. If $CT_1 = g_1^{\gamma}$, the message M' is sent to $\mathscr{A}_3$

(d) *Trapdoor query* (ID, T): $\mathscr{A}_3$ utilizes (ID, T) to issue the query, while $\mathscr{B}$, respectively, uses *ID* and (ID, T) to run *InitialKey query* and *Timekey query* to obtain $IK_{ID} = (IK_{ID1}, IK_{ID2})$ and $TK_{ID} = (TK_{ID1}, TK_{ID2})$. Then $\mathscr{B}$ utilizes IK_{ID2} and TK_{ID2} to produce the trapdoor $TD_{ID} = IK_{ID2} \cdot TK_{ID2}$ which is sent to $\mathscr{A}_3$

(e) *Challenge*: When the phase 1 is over, $\mathscr{A}_3$ outputs a tuple $\langle ID^*, T^*, M^* \rangle$ as the target of the challenge. $\mathscr{B}$ utilizes ID^* to find the corresponding tuples $\langle ID, U_{ID}, u, rb \rangle$ from the $ListH_1$ If $rb = 0$, $\mathscr{B}$ interrupts this game. If $rb = 1$, $\mathscr{B}$ randomly selects $V \in \{0,1\}^l$ to run $H_7\,query$ with M^* and V. Then γ can be obtained. $\mathscr{B}$ utilizes γ to set $CT_1^* = g_1^\gamma$ and find H_6 $query(M^*)^\gamma$ and $H_8\,query(\widehat{e}(CT_2^*, g_2^{a(v^*+\zeta^*)}))$ to get Q and S such that $CT_4^* = Q^\gamma \cdot S$. In addition, $\mathscr{B}$ sets $CT_2^* = g_1^c$, while a random value $CT_3 \in \{0,1\}^{\lambda+l}$ is chosen. Finally, the challenge ciphertext $CT^* = (CT_1^*, CT_2^*, CT_3^*, CT_4^*)$ is sent to $\mathscr{A}_3$

(f) *Phase* 2: $\mathscr{A}_3$ can issue the same query as the phase 1, but it must be under the condition of $ID \neq ID^*$ and $CT \neq CT^*$

(g) *Guess*: $\mathscr{A}_3$ responds to $\mathscr{B}$ with a guess $M' \in \{0,1\}^\lambda$. If $M' \neq M$, $\mathscr{B}$ responds with failure and terminates. Otherwise, $\mathscr{A}_3$ wins the game. Then $\mathscr{B}$ randomly selects a tuple $\langle W^*, CT_1^*, CT_2^*, \omega^* \rangle$ from the $ListH_5$, and outputs the BDH solution A $= (W^*/\widehat{e}(g_1^{ac}, g_2)^{\eta^*})^{u^{*-1}}$ due to $\widehat{e}(g_1, g_2)^{abc} = (W^*/\widehat{e}(g_1^{ac}, g_2)^{\eta^*})^{u^{*-1}}$

The security analysis is similar to Theorem 5. We obtain that $\mathscr{B}$'s advantage to solve the BDH problem is $\varepsilon' \geq (1/q_{H_5})\Pr[E_{H_5}^*] \geq (1/q_{H_5})[(\varepsilon - (1/2^\lambda))/e(q_{\mathrm{IK}} + 1)] - q_D/q$. □

Theorem 8. *If the BDH assumption holds, the proposed RIBEET scheme satisfies the OW-ID-CCA security in the security game. More precisely, suppose that $\mathscr{A}_4$ is a PPT type 4 adversary who has at least ε advantage to break the RIBEET scheme. Then, there exists an algorithm $\mathscr{B}$ to solve the BDH problem with the advantage.*

$$\varepsilon' \geq \left(\frac{1}{q_{H_5}}\right)\left[\frac{\varepsilon - (1/2^\lambda)}{e(q_{TK} + 1)}\right] - \frac{q_D}{q}, \tag{16}$$

where q_{H_5}, q_{TK}, q_D, and e, respectively, are the number of queries to random oracle H_5, Timekey query, Decryption query, and Euler's number.

Proof. An algorithm $\mathscr{B}$ is constructed to solve the BDH problem. The algorithm $\mathscr{B}$ is given a BDH tuple $\langle q, G_1, G_2, G_T, \widehat{e}, g_1, g_1^a, g_1^c, g_2, g_2^a, g_2^b \rangle$ which is defined in Section 2. The algorithm $\mathscr{B}$ can be seen as a challenger to find the answer of the BDH problem. The answer $A = \widehat{e}(g_1, g_2)^{abc}$ can be found by interacting with the PPT type IV adversary $\mathscr{A}_4$ in the following security game.

(i) *Setup*: The challenger $\mathscr{B}$ utilizes the BDH tuple to set $P_{\text{pub}} = g_1^a$, and then generates the public system parameters $PSP = \{q, G_1, G_2, G_T, \widehat{e}, g_1, g_2, P_{pub}, H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$, where H_i is a random oracle for $i = 1, 2, \cdots, 8$. In addition, the system life time $SLT = \{T_0, T_1, \cdots, T_t\}$ can be generated by the challenger $\mathscr{B}$. Then $\mathscr{B}$ gives $\mathscr{A}_4$ the public system parameters PSP and system life time SLT. Here, the adversary $\mathscr{A}_4$ can issue queries to each random oracle as below

(a) $H_1 - H_3$ *queries*: $\mathscr{B}$ answers $\mathscr{A}_4$ in the same form as the proof of Theorem 6

(b) $H_4 - H_8$ *queries*: $\mathscr{B}$ answers $\mathscr{A}_4$ in the same form as the proof of Theorem 5

(ii) Phase 1: The adversary $\mathscr{A}_4$ can, respectively, utilize ID, (ID, T), (ID, T, CT) and (ID, T) to issue the *InitialKey query*, *Timekey query*, *Decryption query* and *Trapdoor query*. The response to each query can be obtained as follows

(a) *InitialKey query* (ID): $\mathscr{B}$ answers $\mathscr{A}_4$ in the same form as the proof of Theorem 6

(b) *Timekey query* (ID, T): $\mathscr{B}$ answers $\mathscr{A}_4$ in the same form as the proof of Theorem 6

(c) *Decryption query* (ID, T, CT): $\mathscr{B}$ answers $\mathscr{A}_4$ in the same form as the proof of Theorem 7

(d) *Trapdoor query* (ID, T): $\mathscr{B}$ answers $\mathscr{A}_4$ in the same form as the proof of Theorem 7

(1) *Challenge*: When the phase 1 is over, $\mathscr{A}_4$ outputs a tuple $\langle ID^*, T^*, M^* \rangle$ as the target of the challenge. $\mathscr{B}$ utilizes (ID^*, T^*) to find the corresponding tuples $\langle ID, T, U_{IDT}, \eta, rb \rangle$ from the $ListH_3$ If $rb = 0$, $\mathscr{B}$ interrupts this game. If $rb = 1$, $\mathscr{B}$ randomly selects $V \in \{0,1\}^l$ to run $H_7\,query$ with M^* and V. Then γ can be obtained. $\mathscr{B}$ utilizes γ to set $CT_1^* = g_1^\gamma$ and find $H_6\,query(M^*)^\gamma$ and $H_8\,query(\widehat{e}(CT_2^*, g_2^{a(v^*+\zeta^*)}))$ to get Q and S such that $CT_4^* = Q^\gamma \cdot S$. In addition, $\mathscr{B}$ sets $CT_2^* = g_1^c$, while a random value $CT_3 \in \{0,1\}^{\lambda+l}$ is chosen. Finally, the challenge ciphertext $CT^* = (CT_1^*, CT_2^*, CT_3^*, CT_4^*)$ is sent to $\mathscr{A}_4$

(2) *Phase* 2: $\mathscr{A}_4$ can issue the same query as the phase 1, but it must be under the condition of $ID \neq ID^*$ and $CT \neq CT^*$

(3) *Guess*: $\mathscr{A}_4$ responds to $\mathscr{B}$ with a guess $M' \in \{0,1\}^\lambda$. If $M' \neq M$, $\mathscr{B}$ responds with failure and terminates. Otherwise, $\mathscr{A}_4$ wins the game. Then $\mathscr{B}$ randomly

selects a tuple $\langle W^*, CT_1^*, CT_2^*, \omega^* \rangle$ from the $ListH_5$, and outputs the BDH solution A $= (W^* / \widehat{e}(g_1^{ac}, g_2)^{u^*})^{\eta^{*-1}}$ due to $\widehat{e}(g_1, g_2)^{abc} = (W^* / \widehat{e}(g_1^{ac}, g_2)^{u^*})^{\eta^{*-1}}$

The security analysis is similar to Theorem 5. We obtain that $\mathscr{B}$'s advantage to solve the BDH problem is $\varepsilon' \geq (1/q_{H_5}) \Pr[E_{H_5}^*] \geq (1/q_{H_5})[(\varepsilon - (1/2^\lambda))/e(q_{\mathrm{TK}} + 1)] - (q_D/q)$. □

Theorem 9. *The proposed RIBEET scheme is secure for brute force attacks if the discrete logarithm problem is hard.*

Proof. As mentioned in the concrete RIBEET scheme, the public system parameters are $PSP = \{q, G_1, G_2, G_T, \widehat{e}, g_1, g_2, P_{pub}, H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$, the system life time is $SLT = \{T_0, T_1, \cdots, T_t\}$ and the master private key is $mpk = s$. Based on the discrete logarithm problem, we ensure that the adversary cannot recover the master private key $mpk = s$ form $P_{pub} = g_1^s$. In addition, the security of the user initial key IK_{ID} and user time key TK_{ID} is also based on the discrete logarithm problem due to $IK_{ID} = (IK_{ID1}, IK_{ID2}) = (H_1(ID)^{mpk}, H_2(ID)^{mpk}) = (H_1(ID)^s, H_2(ID)^s)$ and $TK_{ID} = (TK_{ID1}, TK_{ID2}) = (H_3(ID, T)^{mpk}, H_4(ID, T)^{mpk}) = (H_3(ID, T)^s, H_4(ID, T)^s)$. Hence, the proposed RIBEET scheme can resist brute force attacks. □

6. Comparison

In this section, we compare the proposed RIBEET scheme with the previous RIBE scheme [21] and IBEET scheme [6]. In order to analyze the cost of performing encryption, decryption and equality test, we first define two notations as follows.

(1) Pair: time to perform a bilinear pairing $\widehat{e} : G_1 \times G_2 \longrightarrow G_T$

(2) Exp: time to perform an exponentiation in G_1, G_2 or G_T

We gain Pair = 7.8351 ms and Exp = 0.4746 ms from the literature [32]. These two execution times are obtained under the hardware device with Intel Core i7-8550U 1.80 GHz processor. Meanwhile, the prime number q selected in the cryptosystem setting phase is 256-bit. In addition, three multiplicative cyclic groups G_1, G_2, and G_T of the prime order q are chosen in the simulation.

In Table 3, we list the comparisons of our proposed RIBEET scheme with the RIBE scheme [21] and several IBEET schemes [6, 10, 11] in terms of the cost of performing encryption, decryption and equality test, and two properties related to user revocation and equality test of ciphertexts. For the cost of performing encryption and decryption, Tseng and Tsai's RIBE scheme [21] has better performance than the other two schemes. However, Tseng and Tsai's RIBE scheme does not support equality test of ciphertexts. Although the existing IBEET schemes [6, 10, 11] and our proposed RIBEET scheme support equality test of ciphertexts, the IBEET schemes does not have a mechanism to revoke users. Conversely, our proposed RIBEET scheme not only provides user revocation, but also retains the cost of encryption, decryption and equality test with the existing IBEET schemes. Additionally, Table 4 compares our RIBEET scheme with the RIBE scheme [21] and several IBEET schemes [6, 10, 11] in terms of |PK|, |CT|, and |TD| which are, respectively, denoted as the bit length of user public key, ciphertext and trapdoor. We observed that the communication cost of our RIBEET scheme is similar to that of other schemes.

As mentioned in Section 1, the data collected from sensors on the patients is finally encrypted by the mobile device and then transmitted to the cloud. For the analysis of energy cost, we employ the "ampere" app to measure the voltage and current on the mobile device. After running this app, we obtain 14.28 V and 2856 mA on the mobile device. Table 5 lists the energy cost of performing encryption on the mobile device by using the formula $W = U \cdot I \cdot t$, where W, U, I, and t, respectively, are watt, voltage, current, and time.

7. Conclusions

We considered the existing syntax of IBEET and the property of user revocation to present the new syntax of RIBEET. Under the new syntax, we proposed a concrete RIBEET scheme. Meanwhile, we demonstrated that the proposed scheme has the IND-ID-CCA security for type I and II adversaries and the OW-ID-CCA security for type III and IV adversaries. We compared the proposed scheme with the previous RIBE scheme and IBEET scheme. We showed that the proposed scheme not only supports equality test for ciphertexts but also provides user revocation.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was partially supported by the Ministry of Science and Technology, Taiwan, under contract nos. MOST 110-2222-E-019-001-MY2 and MOST 110-2221-E-019-041-MY3.

References

[1] Y. Dodis, S. Goldwasser, Y. T. Kalai, C. Peikert, and V. Vaikuntanathan, "Public-key encryption schemes with auxiliary inputs," in *Theory of Cryptography*, pp. 361–381, Springer, Berlin, Heidelberg, 2010.

[2] D. Hofheinz and T. Jager, “Tightly secure signatures and public-key encryption,” in *Advances in Cryptology-CRYPTO 2012*, pp. 590–607, Springer, Berlin, Heidelberg, 2012.

[3] R. Chen, Y. Mu, G. Yang, F. Guo, and X. Wang, “Server-aided public key encryption with keyword search,” *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 12, pp. 2833–2842, 2016.

[4] P. Xu, S. He, W. Wang, W. Susilo, and H. Jin, “Lightweight searchable public-key encryption for cloud-assisted wireless sensor networks,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3712–3723, 2018.

[5] D. Boneh and M. Franklin, “Identity-based encryption from the Weil pairing,” in *Advances in Cryptology–CRYPTO 2001*, pp. 213–229, Springer, Berlin, Heidelberg, 2001.

[6] S. Ma, “Identity-based encryption with outsourced equality test in cloud computing,” *Information Sciences*, vol. 328, pp. 389–402, 2016.

[7] X. J. Lin, L. Sun, and H. Qu, “Generic construction of public key encryption, identity-based encryption and signcryption with equality test,” *Information Sciences*, vol. 453, pp. 111–126, 2018.

[8] H. T. Lee, H. Wang, and K. Zhang, “Security analysis and modification of ID-based encryption with equality test from ACISP,” in *Information Security and Privacy*, pp. 780–786, Springer, Cham, 2018.

[9] D. H. Dung, H. Q. Le, P. S. Roy, and W. Susilo, “Lattice-based IBE with equality test in standard model,” in *Provable Security*, pp. 19–40, Springer, Cham, 2019.

[10] R. Elhabob, Y. Zhao, N. Eltayieb, A. M. Abdelgader, and H. Xiong, “Identity-based encryption with authorized equivalence test for cloud-assisted IoT,” *Cluster Computing*, vol. 23, no. 2, pp. 1085–1101, 2020.

[11] X. J. Lin, Q. Wang, L. Sun, and H. Qu, “Identity-based encryption with equality test and datestamp-based authorization mechanism,” *Theoretical Computer Science*, vol. 861, pp. 117–132, 2021.

[12] H. Alemdar and C. Ersoy, “Wireless sensor networks for healthcare: a survey,” *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[13] B. Latr, B. Braem, I. Moerman, C. Blondia, and P. Demeester, “A survey on wireless body area networks,” *Wireless Networks*, vol. 17, no. 1, pp. 1–8, 2011.

[14] J. Zhang, H. Nian, X. Ye, X. Ji, and Y. Heg, “A spatial correlation based partial coverage scheduling scheme in wireless sensor networks,” *Journal of Network Intelligence*, vol. 5, no. 2, pp. 34–43, 2020.

[15] C. H. Hsieh, J. Lin, C. M. Yu, M. H. Hung, and F. Huang, “A TSP-over-LEACH protocol for energy-efficient wireless sensor networks,” *Journal of Network Intelligence*, vol. 6, no. 4, pp. 835–846, 2021.

[16] C. M. Chen, Z. Li, S. A. Chaudhry, and L. Li, “Attacks and solutions for a two-factor authentication protocol for wireless body area networks,” *Security and Communication Networks*, vol. 2021, 12 pages, 2021.

[17] H. Xiong, Y. Hou, X. Huang, Y. Zhao, and C. M. Chen, “Heterogeneous signcryption scheme from IBC to PKI with equality test for WBANs,” *IEEE Systems Journal*, vol. 16, no. 2, pp. 2391–2400, 2021.

[18] R. Housley, W. Polk, W. Ford, and D. Solo, “Internet X.509 public key infrastructure certificate and certificate revocation list (CRL) profile,” 2002, IETF RFC 3280.

[19] A. Boldyreva, V. Goyal, and V. Kumar, “Identity-based encryption with efficient revocation,” in *Proceedings of the 15th ACM conference on Computer and communications security (CCS ‘08)*, pp. 417–426, Alexandria, Virginia, USA, 2008.

[20] B. Libert and D. Vergnaud, “Adaptive-ID secure revocable identity-based encryption,” in *Topics in Cryptology–CT-RSA 2009*, pp. 1–15, Springer, Berlin, Heidelberg, 2009.

[21] Y. M. Tseng and T. T. Tsai, “Efficient revocable ID-based encryption with a public channel,” *Computer Journal*, vol. 55, no. 4, pp. 475–486, 2012.

[22] J. H. Seo and K. Emura, “Revocable identity-based encryption revisited: security model and construction,” in *Public-Key Cryptography–PKC 2013*, pp. 216–234, Springer, Berlin, Heidelberg, 2013.

[23] Y. Watanabe, K. Emura, and J. H. Seo, “New revocable IBE in prime-order groups: adaptively secure, decryption key exposure resistant, and with short public parameters,” in *Topics in Cryptology–CT-RSA 2017*, pp. 432–449, Springer, Cham, 2017.

[24] A. Takayasu and Y. Watanabe, “Lattice-based revocable identity-based encryption with bounded decryption key exposure resistance,” in *Information Security and Privacy*, pp. 184–204, Springer, Cham, 2017.

[25] S. Katsumata, T. Matsuda, and A. Takayasu, “Lattice-based revocable (hierarchical) IBE with decryption key exposure resistance,” in *Public-Key Cryptography–PKC 2019*, pp. 441–471, Springer, Cham, 2019.

[26] A. Takayasu, “Adaptively secure lattice-based revocable IBE in the QROM: compact parameters, tight security, and anonymity,” *Designs, Codes and Cryptography*, vol. 89, no. 8, pp. 1965–1992, 2021.

[27] J. Baek, R. Safavi-Naini, and W. Susilo, “Public key encryption with keyword search revisited,” in *Computational Science and Its Applications–ICCSA 2008*, pp. 1249–1259, Springer, Berlin, Heidelberg, 2008.

[28] S. T. Hsu, C. C. Yang, and M. S. Hwang, “A study of public key encryption with keyword search,” *International Journal of Network Security*, vol. 15, no. 2, pp. 71–79, 2013.

[29] G. Yang, C. H. Tan, Q. Huang, and D. S. Wong, “Probabilistic public key encryption with equality test,” in *Topics in Cryptology–CT-RSA 2010*, pp. 119–131, Springer, Berlin, Heidelberg, 2010.

[30] A. Shamir, “Identity-based cryptosystems and signature schemes,” in *Advances in Cryptology*, pp. 47–53, Springer, Berlin, Heidelberg, 1985.

[31] H. Li, Q. Huang, S. Ma, J. Shen, and W. Susilo, “Authorized equality test on identity-based ciphertexts for secret data sharing via cloud storage,” *IEEE Access*, vol. 7, pp. 25409–25421, 2019.

[32] Y. Li, Q. Cheng, X. Liu, and X. Li, “A secure anonymous identity-based scheme in new authentication architecture for mobile edge computing,” *IEEE Systems Journal*, vol. 15, no. 1, pp. 935–946, 2021.

Hindawi
Journal of Sensors
Volume 2022, Article ID 4102552, 16 pages
https://doi.org/10.1155/2022/4102552

Research Article

A Lightweight Human Action Classification Method for Green IoT Sport Applications

Da Xiao,[1] **Tianyu Huang**,[1,2] **Yihao Li**,[1] **Chang Liu**,[1] **and Fuquan Zhang**[2,3]

[1] *Department of Computer Science and Technology, Beijing Institute of Technology, 100081, China*
[2] *Beijing Key Laboratory of Digital Performance and Simulation Technology, 100081, China*
[3] *College of Computer and Control Engineering, Minjiang University, 350108, China*

Correspondence should be addressed to Tianyu Huang; huangtianyu@bit.edu.cn

Received 3 May 2022; Accepted 15 June 2022; Published 30 June 2022

Academic Editor: Saru Kumari

This paper proposes a lightweight human action classification method for Green Internet of Things (IoT) sport applications. This method classifies the human motion data collected by wearables or other IoT devices with energy-efficient techniques, by enabling a small number of sample training and incremental classification to achieve the purpose of energy-efficient. To lessen the complexity of the model and reduce the number of samples required for parameter estimation, we propose a shared Hidden Conditional Random Field (sHCRF) model. The sHCRF model adds a shared-classification layer structure to reduce the parameter computation. In the experiments, the classification accuracy of the sHCRF model is above 95%. This paper introduces an incremental learning method based on knowledge distillation. The new model suppresses the forgetting of existing classification knowledge while fitting new data to learn new classification knowledge. In the incremental scenarios, the classification accuracy of the sHCRF model is above 70%. The experimental results show that this method can lightly implement convenient and fast automatic classification of action acquisition.

1. Introduction

In recent years, smart wearables and mobile sensors have become an integral part of human's daily lives [1] with the development of Green Internet of Things (IoT). These wearables can be used to capture the human motion data when people exercise or engage in other daily activities. By analyzing and classifying the captured motion data, it can be used to assist exercise and monitor the human health, etc., in which the results of analysis or classification can be fed back through sport applications, like sport watches and smart treadmill. Human action recognition is a higher-level work for computers to understand human motion. Among them, the motion categories are the base for human action recognition. Due to the complexity and temporal impact of human motion, high-dimensional features of motion data complicate human action classification. Modeling of human motion involves the relationship between thousands of variables. Researchers have proposed human action recognition models based on neural networks [2], like convolutional neural networks (CNN) [3], Recurrent Neural Networks (RNN) [4], and Graph Convolutional Neural (GCN) [5] networks. However, these models require more energy, and they need to be trained on a large number of manually labelled samples. The training process is time-consuming and labor-intensive. Our work is aimed at building a lightweight action classification method to achieve the purpose of energy efficiency, which enables the model training on a small number of samples while saving storage and computational overhead. The probabilistic graphical model has a strong ability to model the relationship between variables based on prior knowledge, greatly reducing model parameters and, more importantly, reducing the amount of sample data required for parameter estimation. It is potential to model motion lightweightly based on the probabilistic graphical models.

The human motion data captured by sensors is time series data, and each frame represents the posture of the human body at the current moment. It is necessary to pay attention to the spatiotemporal information so as to accurately describe the characteristics of human action classification. The Hidden Conditional Random Field (HCRF) model, as an undirected probabilistic graphical model, is a discriminative model. It can label the entire sequence of samples as actions and use hidden state variables to capture intermediate structures. For action recognition task, the existing consensus is that the ideal model should be derived and optimized on the basis of maximizing the discriminant function.

This paper proposes a lightweight human action classification method. We introduce posture base, posture change base, and posture semantic base to characterize human motion data. The features are obtained based on the data collected by sensors corresponding to the main kinematic joints. In order to reduce the computational cost of training, we propose a shared Hidden Conditional Random Field (sHCRF) model by designing a shared-classification layer structure, which reduces parameter amounts. The incremental learning and classifying methods are designed by introducing knowledge distillation. The framework of our method is shown in Figure 1. Human motion data is continuously collected by smart wearables, and then, the collected data is characterized. The features are sent to the sHCRF model for training. When the model fits new data, the model will add the distillation loss on the basis of the classification loss, so that the model can suppress forgetting owned knowledge while learning new classification knowledge. After training the model, the updated classification knowledge will be uploaded to the IoT. The contributions of this paper are as follows:

(1) A lightweight human action classification method for Green IoT sport applications is proposed. The method is based on probabilistic graphical model training on a small number of samples. In the meantime, it can realize automatic incremental learning and classifying by knowledge distillation

(2) Human action features are designed, including posture sequence, posture change sequence, and posture semantic sequence. They can describe human posture features and the temporal correlation between human postures

(3) A shared-classification layer structure is introduced to improve the HCRF probabilistic graphical model. It lessens parameter amount and achieves better classification

2. Related Work

Our lightweight action classification method is related to feature description of human motion, the human action classification model, and incremental learning methods.

2.1. Feature Representation Method of Human Motion Data. Human body motion data contain complex information due to the many multidimensional data involved. The semantic motion information of the motion data cannot be directly reflected. So, it is necessary to extract the features of the data. The extracted motion data features provide a basis for similarity measurement between different actions. Forbes and Fiume [6] used the weighted Principal Component Analysis (PCA) dimensionality reduction algorithm to extract the feature representation of motions. Still, the corresponding relationship between the features extracted by this method and the motion semantics is not apparent. Müller et al. [7] proposed a method for indexing the geometric parts of the human body. They defined 31 Boolean features to describe the geometric relationship of the human posture, ensuring the unity of logical similarity and numerical similarity of human motion. Their method has groundbreaking significance, but it had complicated feature definition. Liang et al. [8] used the technique of subspace division to represent the geometric features of human posture. They defined a set of feature vectors to represent motion data based on this method. This method is concise for the feature definition of human posture.

Combining the geometric features of the human body and the method of subspace division, the motion data features in this paper are divided into parts that describe specific human postures, posture changes, and features that characterize the primary state of posture. The first is from the static and dynamic perspectives, and the last is related to the parameter construction of the model. Semantic information in the same essential state in different actions may differ.

2.2. Human Action Classification Method. Action classification methods for human action need to model the temporal and spatial information in motion data. In recent years, researchers have proposed many action classification methods based on neural network. The methods they used include the model based on RNN [9], CNN [10, 11], GCN [12–14], and Long Short-Term Memory (LSTM) [15]. The probabilistic graphical model is one of the popular solutions to the problem of action classification. The probabilistic graphical model is divided into a directed graph model and an undirected graph model. Both of them are suitable for modeling sequence models. Ma et al. [16] used the Hidden Markov Model (HMM) to recognize human action. Samr and Nizar [17] used the Beta-Liouville HMM action classification method. Wen et al. [18] used a Hierarchy Dirichlet Process-Hidden Markov Model (HDP-HMM) to represent the action class. This method can automatically obtain the number of hidden states during the learning process. Using undirected probabilistic graphs, Vrigkas et al. [19] used the HCRF model to recognize the human activities. The modified Hidden Condition Random Field (mHCRF) model based on the HCRF model was proposed by Zhang and Gong [20]. Their method optimized the algorithm by introducing the exact hidden state sequence obtained from the HMM to prevent it from falling into a local optimum.

We use the sHCRF model for action classification, which is based on a probabilistic graph model. It is suitable for modeling the temporal correlation in sequence data, and the graph model supports sequence data input with different lengths.

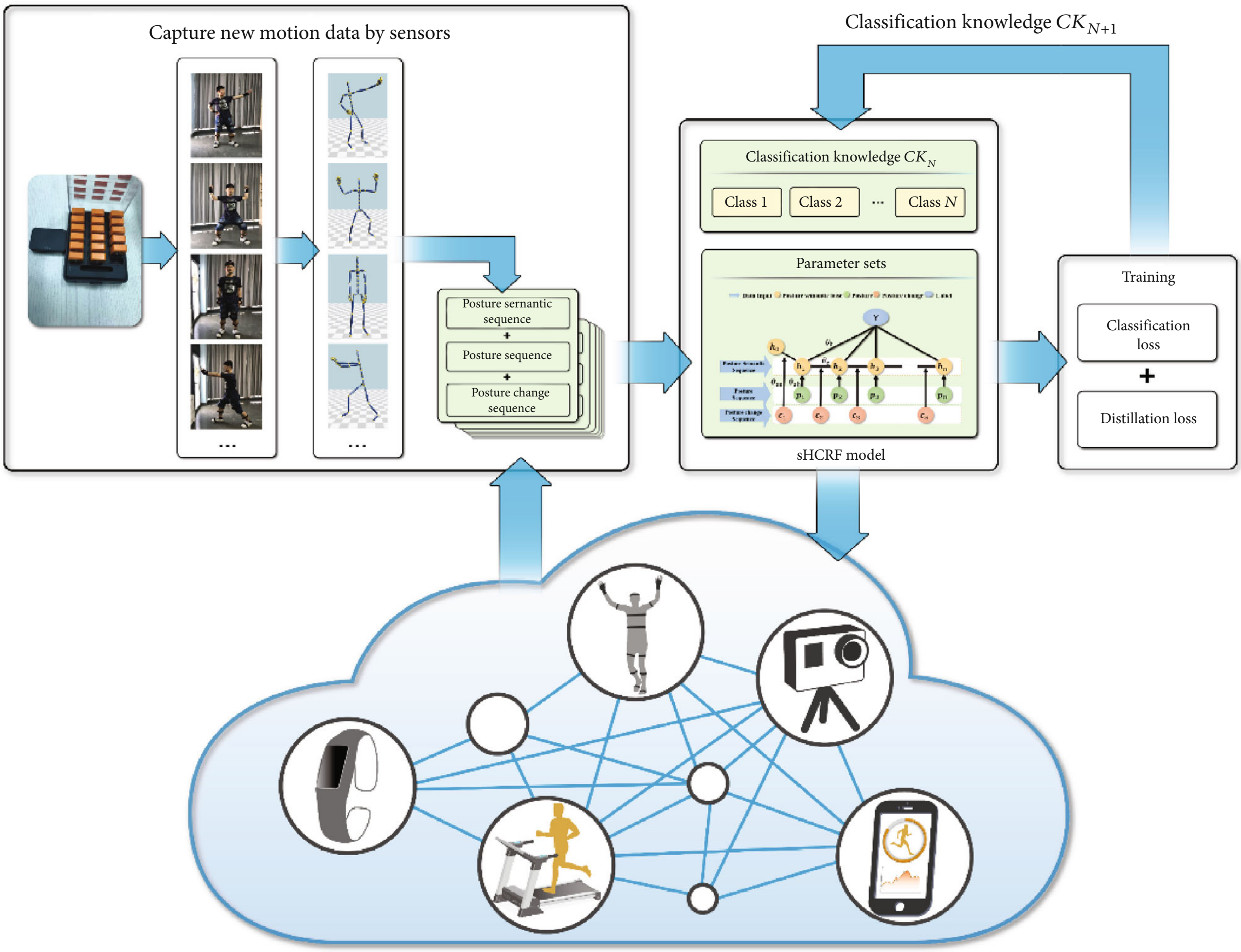
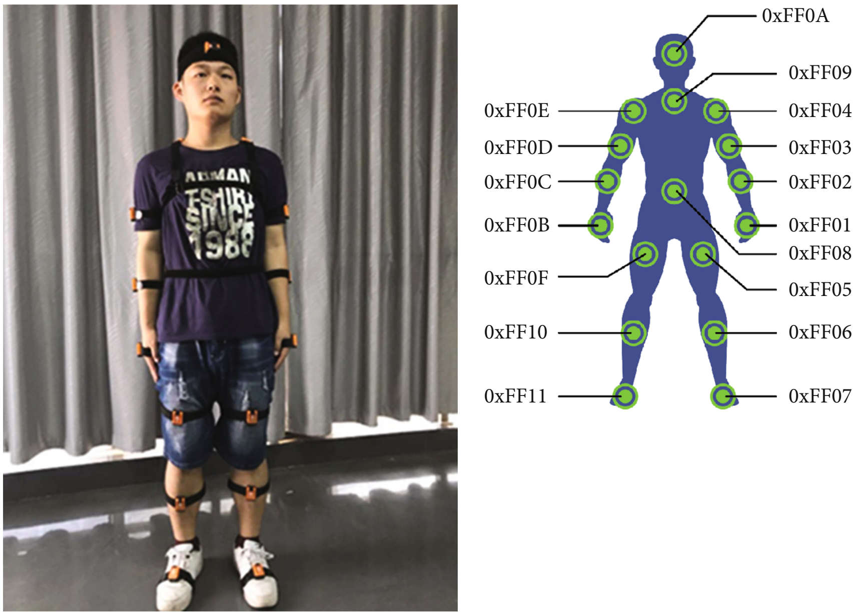


Figure 1: The framework of the action classification for Green IoT.

2.3. Incremental Learning Method. At present, among the methods for achieving category increment, the incremental learning [21] method is showing adequate results in various classification techniques. Incremental learning is a system that continuously learns from new samples and saves most previously learned knowledge. Whenever new data are added, it is not necessary to rebuild all the knowledge bases. Only the changes caused by the new data update the original knowledge bases. The incremental learning method is more closely aligned with the principles of human thinking.

Incremental learning methods based on regularization usually do not need to use old data to let the model review the tasks it has learned. The Learning without Forgetting (LwF) algorithm proposed by Li and Hoiem [22] does not need to use the data of an old task and can fit new data while suppressing the forgetting of old knowledge. The main idea of the LwF algorithm comes from the knowledge distillation method proposed by Hinton et al. [23], which predicted the new model on a new task similar to the prediction of the old model on the new task. Irfan et al. [24] proposed a model which can handle both multitask and single incremental task scenarios as opposed to various existing models that cover only the multitask scenarios. The Elastic Weight Consolidation (EWC) algorithm based on the Bayesian framework proposed by Kirkpatrick et al. [25] introduces an additional parameter-related regular loss. The loss encourages the new model parameters obtained by further task training to be as close as possible to the old model parameters according to the importance of different parameters. In the replay-based incremental learning method, when training a new task, part of the old representative data is retained and used for the model to review the old knowledge learned. It may overfit the retained old data. Lopez-Paz and Ranzato [26] proposed the Gradient Episodic Memory (GEM) algorithm for this problem. It only updates the parameters of the new task without interfering with the parameters of the old task. It modifies the gradient update direction of the new task in an inequality-constrained manner so that the model does not increase the old one. Concurrently with the loss of the task, it tries to minimize the loss value of the new task.

Most incremental learning methods are based on convolutional neural networks, and they have few applications to probabilistic graphical models. Moreover, there is no clearly applicable method in action classification. Based on the analysis and conclusions presented in Section 2.2, the action classification model proposed in this paper uses the



(a) Schematic diagram of sensor wearing

(b) Sensor position correspondence

Figure 2: Motion capture sensor set.

probability graph model as its basic model and introduces the idea of "shared parameters" in the depth model, which has the conditions for category increment.

3. Collection and Characterization of Motion Capture Data

The human action features in this paper are extracted from motion capture data. The motion capture data is represented as the skeleton hierarchy, which refers to the motion nodes as joints. They have characteristics that describe human posture, posture changes, and features that define deep semantic information of human actions. The first two are composed of a posture base and a posture change base, respectively, and describe the static and dynamic characteristics of the data, respectively. The last feature is referred to as the posture semantic base in this paper. The posture semantic base represents the essential characteristics of a human posture. The complete set of posture semantic base defined in this paper is the posture semantic base set $\mathscr{H}_D$, where D is the size of the collection. The state transition set T_C represents all posture semantic base transitions between the posture semantic bases, where C is the size of the set.

3.1. Motion Capture Data Collection. The motion acquisition device consists of 17 node sensors and a hub developed by our research group. As shown in Figure 2, there is a 9-axis sensor of each node, containing a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer. The precision of the sensor is +0.5 degrees. The original data collected are gravity acceleration, rotation rate, and magnetic force data, respectively. The reconstructed motion data is solved by a 9-axis fusion algorithm. It is identified by the specific data ID of each sensor, where the position of sensor nodes corresponds to human joints. The IDs are shown in Table 1.

Table 1: Sensor position correspondence.

MTS_ ID	Corresponding location	MTS_ ID	Corresponding location
0xFF01	Left hand	0xFF0B	Right hand
0xFF02	Left low-arm	0xFF0C	Right low-arm
0xFF03	Left up-arm	0xFF0D	Right up-arm
0xFF04	Left shoulder	0xFF0E	Right shoulder
0xFF05	Left up-leg	0xFF0F	Right up-leg
0xFF06	Left low-leg	0xFF10	Right low-leg
0xFF07	Left foot	0xFF11	Right foot
0xFF08	Waist		
0xFF09	Chest		
0xFF0A	Head		

3.2. Posture Base. Human motion capture data are sequence data of multiple frames of human posture arranged along the time axis. Each frame is composed of three-dimensional rotation data for each joint of the human body. In this paper, we define the number of primary moving joints M. The primary moving joints of the human body include the joints on the arms (including shoulders, elbows, and wrists), leg joints (including hips, knees, and ankles), and torso joints (mainly the abdomen). When M is 13, the primary moving joint is composed of all arm joints and leg joints, plus abdominal joints. This paper takes the root joint coordinate system of the initial posture of the human motion data as the absolute coordinate system of the primary moving joint. The rotation matrix of the moving joint is obtained from its Euler rotation angle. The direction vector of the moving joint is taken from a unit vector parallel to the direction vector of the coordinate axis in the absolute

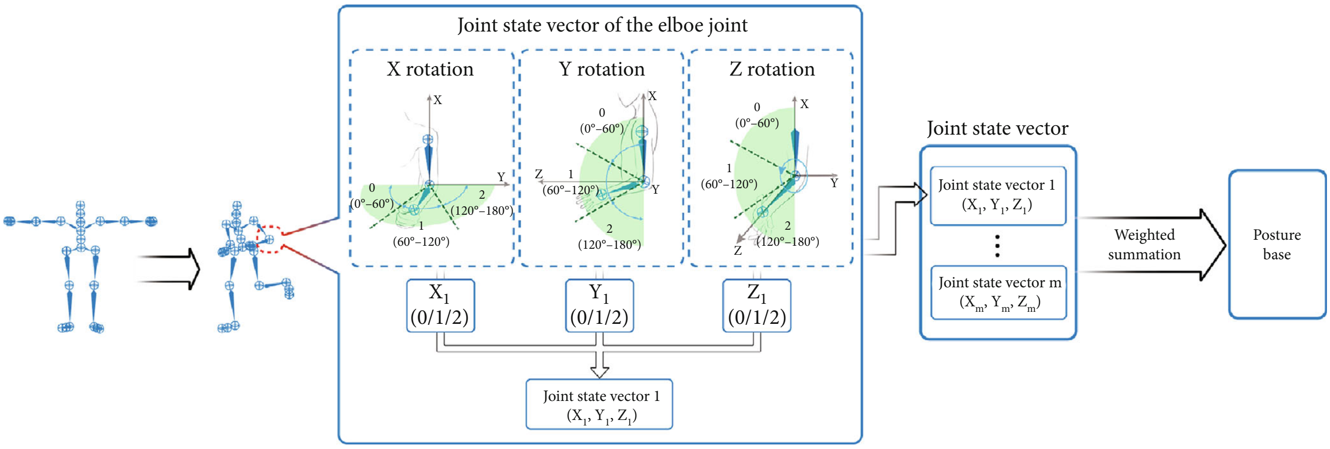


FIGURE 3: The construction process of the posture base.

coordinate system. In the local coordinate system of the joint, the spherical coordinate of the moving joint is composed of the angle between the direction vector and the axis of the local coordinate system.

To get the posture base of the joint, we first divided the rotation space of the joint into three subspaces by each axis according to the knowledge of human kinematics and related motion experience. Then, we selected a direction vector of the joint and calculate the angles between the vector and each coordinate axis. We used the angles to determine the subspace where the vector is located by each axis, and we obtained a joint state vector, whose length was 3. We also obtained other state vectors by selecting different direction vectors. Finally, we selected the required values from the obtained state vector to get a weighted sum. The weighted sum is the posture base of the joint. As shown in the construction process of the posture base of the elbow joint in Figure 3, we first select the lower arm bone as the direction vector of the elbow joint. The rotation space of the elbow joint is divided into three subspaces by each axis, which are $0^\circ \sim 60^\circ$, $60^\circ \sim 120^\circ$, and $120^\circ \sim 180^\circ$, respectively. A three-bit code is defined according to the spherical coordinates, and each bit code can take 0, 1, or 2 according to the angle interval of each dimension in the spherical coordinates. The code is a ternary code. At this point, the information about the direction vector of the joint rotating around itself is missing; so, it is necessary to introduce the spherical coordinate information of other vectors perpendicular to the direction vector to obtain other ternary code and to construct the posture base of the elbow joint in the current frame. The formula of posture base can be expressed as follows:

$$pb_i^k = f_{\mathrm{p}}\left(\mathbf{J}_i^k, \mathbf{V}_k\right) = \sum_{l=0}^{n} Z^{\alpha\left(\mathbf{J}_i^k, \mathbf{V}_k^l\right)}, \tag{1}$$

$$k \in (0, M], i \in (0, N], n \in [2, 3], l \in [1, n]. \tag{2}$$

pb_i^k represents the posture base of the kth joint in the ith frame data. $f_p(\mathbf{J}_i^k, \mathbf{V}_k)$ is the characteristic function. $\mathbf{J}_i^k$ represents the spherical coordinates of the kth joint in the ith frame data. $\mathbf{V}_k$ is the direction vector set of the kth joint. $\sum_{l=0}^{n} Z^\alpha$ is the weighted sum of joint states for the purpose of compressing information. n represents the number of joint state vectors, and its value is related to the rotational freedom of the joint. Z represents the number of divisions of the joint's subspace. Function α is the method to determine the state of the included angle in the corresponding dimension. $\mathbf{V}_k^l$ represents the lth direction vector in $\mathbf{V}_k$. Function α is defined as

$$\alpha\left(\mathbf{J}_{\mathbf{i}}^{\mathbf{k}}, \mathbf{V}_{\mathbf{k}}^{\mathbf{l}}\right) = \begin{cases} 0, & \left\langle \mathbf{J}_i^k, \mathbf{V}_k^l \right\rangle \in (\beta_1, \beta_2), \\ 1, & \left\langle \mathbf{J}_i^k, \mathbf{V}_k^l \right\rangle \in (\beta_2, \beta_3), \\ 2, & \left\langle \mathbf{J}_i^k, \mathbf{V}_k^l \right\rangle \in (\beta_3, \beta_4), \end{cases} \tag{3}$$

$$0^\circ \le \beta_1 < \beta_2 < \beta_3 < \beta_4 \le 180^\circ. \tag{4}$$

$\langle \mathbf{J}_i^k, \mathbf{V}_k^l \rangle$ represents the angle between $\mathbf{J}_i^k$ and $\mathbf{V}_k^l$. β_1, β_2, β_3, and β_4 are subspace boundary values of joint rotation space. In Figure 3, $\beta_1 = 0^\circ$, $\beta_2 = 60^\circ$, $\beta_3 = 120^\circ$, and $\beta_4 = 180^\circ$.

3.3. *Posture Sequence and Posture Change Sequence.*

The human posture sequence is a time series composed of postures. The posture $\mathbf{p}_i$ of the ith frame can be expressed as $\{pb_i^1, pb_i^2, \cdots, pb_i^M\}$. The posture sequence $\mathbf{P}^N$ is expressed as $\{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N\}$.

The posture change base is the code obtained by changes of the spherical coordinates of the corresponding joint between adjacent moments—increasing, decreasing, and unchanged—similar to the coding method of the posture base. Its code is computed by the included angle change value of each dimension. The coding formula is expressed as follows:

$$pc_i^k = f_{\mathrm{c}}\left(\mathbf{J}_i^k\right) = \sum_{l=0}^{n} 3^{d\left(\mathbf{J}_i^{k,l}, \mathbf{J}_{i+1}^{k,l}\right)}, \tag{5}$$

$$k \in (0, M], i \in (0, N], n \in [2, 4], l \in [1, 3]. \tag{6}$$

pc_i^k represents the posture change base of the kth joint between the ith and the $(i+1)$-th frame. $f_c(\mathbf{J}_i^k)$ is the characteristic function. $\sum_{l=0}^{n} 3^d$ is the weighted summation, where 3 indicates that the angle changes of the joints in three coordinate axis directions need to be considered. d is the coding function for changing spherical coordinates. It is defined as

$$d\left(\mathbf{J}_i^{k,l}, \mathbf{J}_{i+1}^{k,l}\right) = \begin{cases} 0, & \mathbf{J}_{i+1}^{k,l} - \mathbf{J}_i^{k,l} > 0, \\ 1, & \mathbf{J}_{i+1}^{k,l} - \mathbf{J}_i^{k,l} = 0, \\ 2, & \mathbf{J}_{i+1}^{k,l} - \mathbf{J}_i^{k,l} < 0. \end{cases} \quad (7)$$

$\mathbf{J}_i^{k,l}$ and $\mathbf{J}_{i+1}^{k,l}$ are the lth dimensional angles in $\mathbf{J}_i^k$ and $\mathbf{J}_{i+1}^k$, respectively. The posture change $\mathbf{c}_i$ of the ith and the $(i+1)$-th frame can be expressed as $\{pc_i^1, pc_i^2, \cdots, pc_i^M\}$. The posture change sequence $\mathbf{C}^N$ can be expressed as $\{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N\}$.

3.4. Posture Semantic Base. The posture semantic base of a frame in the human motion data represents the essential characteristics of the current posture. The posture semantic base corresponding to different postures may be the same, but it may represent different semantic information among the actions. A collection of posture semantic bases in a movement can be a feature representation of the campaign. This collection is a subset of the posture semantic base set $\mathscr{H}_D$.

The posture semantic base considers the dynamic and static characteristics of human motions. It is a K-dimensional integer vector $\mathbf{h}^K = \{\mu_1, \cdots, \mu_p, v_1, \cdots, v_q\}$, where p-dimensional data $\boldsymbol{\mu} = \{\mu_i | i = 1, \cdots, p\}$ describe the state of p motion joints of the human body and represent the state of the selected joint in the human body coordinate system. This method divides the joint rotation space into two subspaces, corresponding to the rotation angle of the joint. Taking the shoulder joint as an example, this method divides the rotation space of the shoulder joint into upper and lower subspaces, where the horizontal plane is used as the interface. The subspace is for determining the state of the joint; in addition, the q-dimensional data $\mathbf{v} = \{v_j | j = 1, \cdots, q\}$ describe the movement of the human body and are reflected by the angle between the direction of human motion and the direction of the human face in the current frame. For example, when $K = 8$, $p = 6$, and $q = 2$, $\boldsymbol{\mu}$ represents 6 states of the shoulders, hips, root joint, and abdomen joints; $\mathbf{v}$ represents the movement of the human body in the horizontal and vertical directions. The size of the posture semantic base set $\mathscr{H}_D$ is related to the definition of the posture semantic base, which is 2^K.

The posture semantic base is used as a hidden state node in the HCRF model. But in the HCRF model, the hidden state sequence is uncertain. In our method, we designed the posture semantic sequence as a certain input of the hidden state layer to optimize the model.

4. An Incremental Action Classification Method Based on the sHCRF Model

This paper proposes a new action classification model based on the human action features described above, namely, the shared Hidden Conditional Random Field model. The model was improved based on HCRF, and it introduces a structure with a shared-classification layer. To solve the problem of data storage during the learning process, this study used batched incremental learning.

4.1. The Shared-Classification Layer Structure. The sHCRF model introduces a shared-classification layer structure, which has two layers. One is the shared layer, which is used to extract the motion semantic information from the features. The other is the classification layer, which uses the information from the shared layer to classify the samples. The structure can reduce the redundancy of model parameters, as well as maintain high classification accuracy.

The shared layer is composed of two parts, one is used to process posture sequence, and it is a parameter matrix of size $M \times D$. The other is used to process posture change sequence, and it is a $M \times C$ parameter matrix. The former is constructed based on the posture semantic base set $\mathscr{H}_D$, and the latter is constructed based on the posture semantic base transition set T_C. Both the parts get the spatiotemporal information of the human action from the features. Figure 4 shows the shared layer for posture. As shown in Figure 4, in the shared layer for posture of the sHCRF model, each column vector $\mathbf{h}_i$ represents a posture semantic base. There are N posture semantic bases in the structure. Each action has its posture semantic base set, which is a subset of the posture semantic base set $\mathscr{H}_D$.

The shared layer's output is input to the classification layer. Each category corresponds to a classification layer. The number of the columns of the two-parameter matrices in the classification layer is D and C, respectively, and the number of the rows of both the two-parameter matrices is the number of the data categories. The classification layer uses these data to calculate the probability that the sample belongs to each category. At last, the category with the highest probability is used as the label of the sample.

The inference process of the sHCRF model, which uses the shared-classification structure, is as follows. When performing human action classification tasks, the model first obtains the semantic information of the human posture and posture changes in the human motion data. It then brings this information into the specific human body motion and analyzes the time-space relationship of motion again. Finally, it obtains the probability that the input sample belongs to the current motion category.

4.2. sHCRF Model Introduction. The sHCRF model further optimizes the classic HCRF model by introducing a shared-classification layer structure. This structure significantly reduces the model's computational complexity and improves its speed of motion modeling and classification accuracy. The classical HCRF model obtains the classification probability of a given input by fitting the conditional

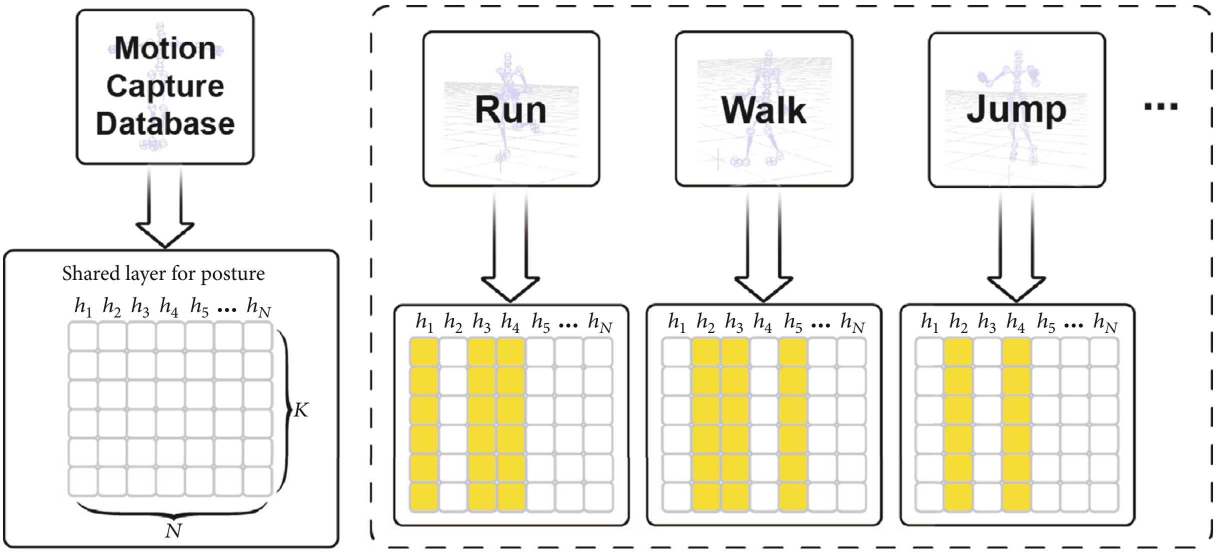

Motion Capture Database
Shared layer for posture
h_1 h_2 h_3 h_4 h_5 ... h_N
K
N
Run
Walk
Jump
...
h_1 h_2 h_3 h_4 h_5 ... h_N
h_1 h_2 h_3 h_4 h_5 ... h_N
h_1 h_2 h_3 h_4 h_5 ... h_N

Figure 4: Diagram of sHCRF for motion categories. The left part corresponds to the shared layer for posture of the sHCRF model, and the right part shows the corresponding posture semantic base set (a subset of the posture semantic base set $\mathscr{H}_D$) in different motions.

probability $P(Y \mid \mathbf{P}^N ; \Theta)$, where $\mathbf{P}^N$ is the sequence data with input length N, Y is the sample label, and Θ is the parameter of the model. The model assumes that the hidden state sequence H of the sample, which can be understood as the posture semantic sequence in the sHCRF model, is uncertain, and all hidden state arrangements are considered in the calculation. The computational complexity is $O(Q \times 2^{KN})$, Q is the number of motion categories, K is the dimension of the hidden state, and N is the length of the input sequence. The conditional probability formula calculated by the model is as follows:

$$P\left(Y \mid \mathbf{P}^N ; \Theta\right) = \frac{\sum_H e^{\phi\left(y,H,\mathbf{P}^N;\Theta\right)}}{\sum_{y,H} e^{\phi\left(y,H,\mathbf{P}^N;\Theta\right)}}. \tag{8}$$

$\mathbf{P}^N$ is the human posture sequence. ϕ is the potential function of the model, and it is defined as

$$\begin{aligned}\phi\left(y, H, \mathbf{P}^N ; \Theta\right) = {} & \sum_j f_h\left(\mathbf{h}_j\right) \times \theta_h\left(y, \mathbf{h}_j\right) + \sum_j f_e\left(e_j\right) \times \theta_e\left(y, e_j\right) \\ & + \sum_j \mathbf{f}\left(\mathbf{p}_j\right) \times \boldsymbol{\theta}\left(\mathbf{h}_j = \mathbf{a}\right).\end{aligned} \tag{9}$$

f_* is the characteristic function of the model, and both $f_h(\mathbf{h}_j)$ and $f_e(e_j)$ are the one-dimensional integers. $f_h(\mathbf{h}_j)$ is the feature corresponding to the hidden state $\mathbf{h}_j$. $f_e(e_j)$ is the feature corresponding to the jth state transition e_j. $\mathbf{f}(\mathbf{p}_j$) is a multidimensional vector, which is the feature of the posture of the jth frame $\mathbf{p}_j$. Θ is divided into $\boldsymbol{\theta}, \theta_h, \theta_e$ according to the corresponding characteristics. $\boldsymbol{\theta}$ represents the degree of correlation between the observation node and the hidden state node, which is a multidimensional vector. $\mathbf{a}$ is a current hidden state, and its length is the same as the length of the human posture at the current moment. θ_h is the weight of the corresponding hidden state in a particular category, and θ_e is the weight of the corresponding state transition in a category. The model structure is shown in Figure 5.

In classification of human action, based on the input of certain human action features, the posture semantic sequence of the input is also determined. Equations (8) and (9) can be simplified to Equations (10) and (11), respectively. The computational complexity has been reduced to $O(Q \times N)$.

$$P\left(Y \mid \mathbf{P}^N ; \Theta\right) = \frac{e^{\phi\left(y,H\left(\mathbf{P}^N\right);\Theta\right)}}{\sum_y e^{\phi\left(y,H\left(\mathbf{P}^N\right);\Theta\right)}}, \tag{10}$$

$$\phi\left(y, H\left(\mathbf{P}^N\right) ; \Theta\right) = \sum_j \mathbf{f}_h\left(\mathbf{h}_j\right) \times \boldsymbol{\theta}_h\left(y, \mathbf{h}_\mathbf{j}\right) + \sum_j \mathbf{f}_e\left(\mathbf{c}_j\right) \times \boldsymbol{\theta}_e\left(y, \mathbf{c}_j\right). \tag{11}$$

Both $\mathbf{f}_h(\mathbf{h}_j)$ and $\mathbf{f}_e(e_j)$ are the multidimensional vectors. Figure 6 shows the structure of the mHCRF model. The model has a certain hidden state sequence as the input, and there is also a posture change sequence.

The improved potential function does not consider the human body posture sequence or posture change sequence. These two feature sequences carry richer posture information than the hidden state sequence. The lack of this information reduces the effectiveness of a model. Therefore, this paper proposes an sHCRF, in which the shared parameter draws on the "sharing mechanism" of the convolutional network, which not only effectively reduces the number of model parameters but also extracts the corresponding information of the human action sequence under a hidden state (or state transition). The model considers the series of human actions while introducing shared parameters and

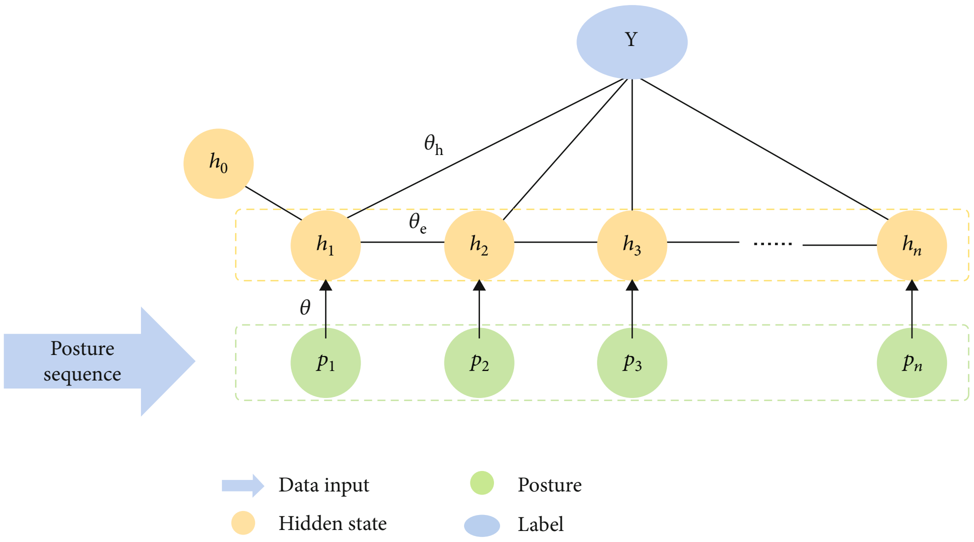


Figure 5: HCRF model structure, in which the hidden state sequence considers the combination of all hidden states.

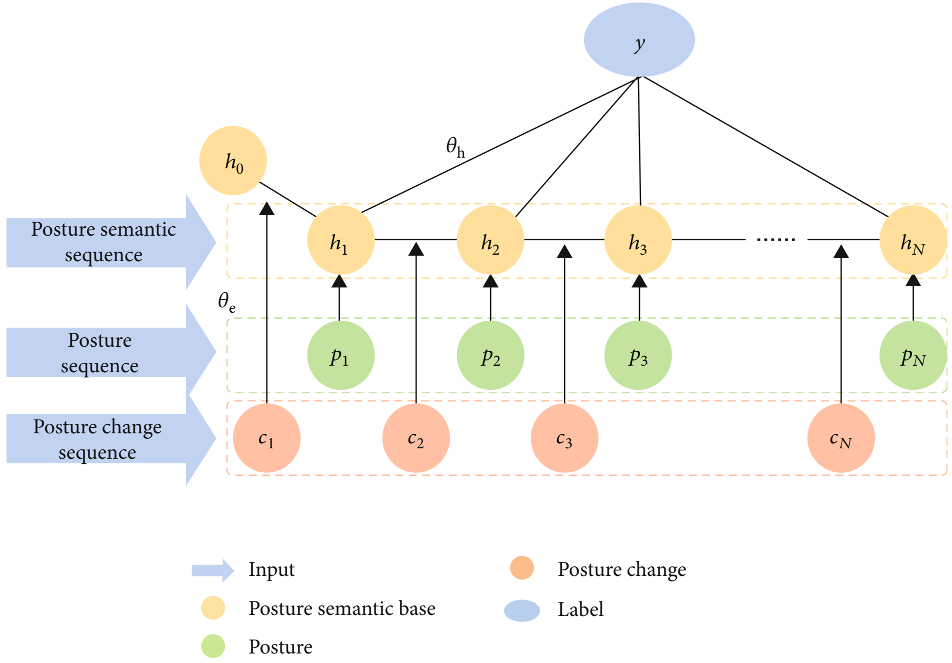


Figure 6: mHCRF model structure, in which the hidden state sequence is certain.

the determined hidden state sequence. Equations (10) and (11) are improved as follows:

$$P\left(Y \mid \mathbf{P}^N ; \Theta\right) = \frac{e^{\phi\left(y,H,\mathbf{P}^N;\Theta\right)}}{\sum_y e^{\phi\left(y,H,\mathbf{P}^N;\Theta\right)}}. \qquad (12)$$

$\boldsymbol{P}^N$ is the human posture sequence, and ϕ is defined as

$$\begin{aligned}\phi\left(y, H, \mathbf{P}^N ; \Theta\right) = {} & \sum_j \mathbf{f}\left(\mathbf{p}_j\right) \times \boldsymbol{\theta}_{\mathrm{sh}}\left(\mathbf{h}_j = \mathbf{a}\right) \times \theta_h\left(y, \mathbf{h}_j\right) \\ & + \sum_j \mathbf{f}_e\left(\mathbf{c}_j\right) \times \boldsymbol{\theta}_{\mathrm{se}}\left(\mathbf{c}_j = b\right) \times \theta_e\left(y, \mathbf{c}_j\right).\end{aligned} \qquad (13)$$

$\mathbf{p}_j$ is a human posture of the jth frame, and $\mathbf{c}_j$ is a posture change between the jth and the $(j+1)$-th frame. $\boldsymbol{\theta}_{\mathrm{sh}}$ and $\boldsymbol{\theta}_{\mathrm{se}}$ are shared parameters from the shared layer, whose function is to extract the semantic information of the input features and compress its size. The lengths of $\boldsymbol{\theta}_{\mathrm{sh}}$ and $\boldsymbol{\theta}_{\mathrm{se}}$ are the same as $\mathbf{f}(\mathbf{p}_j)$ and $\mathbf{f}_e(\mathbf{c}_j)$, respectively. θ_h and θ_e are classification parameters from the classification layer, which are used as weights to determine the category of posture and posture change. The dimension of both parameters is 1. Figure 7 shows the structure of the sHCRF model. There are three types of inputs to the model, posture semantic sequence, posture sequence, and posture change sequence. The function of posture semantic sequence is to extract the

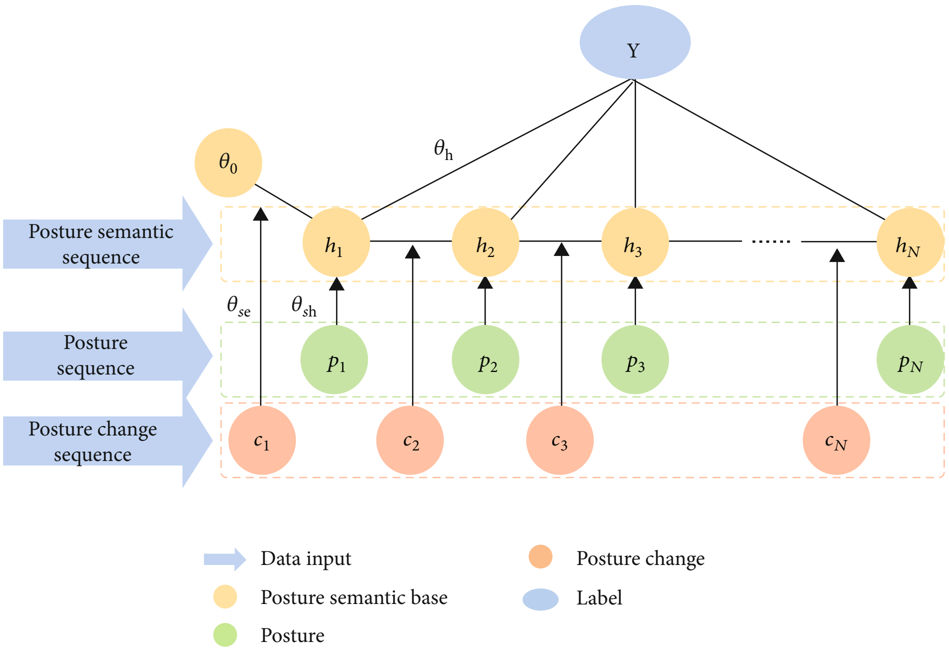


FIGURE 7: sHCRF model structure, in which the shared-classification layer structure is introduced.

information of posture sequence and posture change sequence according to some semantics. The model calculates the probability that the extracted information belongs to a certain category.

4.3. An Incremental Action Classification Based on Knowledge Distillation. To adapt to the continuous input of new categories of action samples, this paper uses the incremental learning method based on the sHCRF model to retain the old knowledge while acquiring new knowledge. In general, it is best to retrain the model by combining old data with new data, but additional storage space is required for the old data. To simulate the learning and memory mechanism of the human brain, the sHCRF model only uses new data for training. It prevents the forgetting of old knowledge without old data by adding distillation loss based on classification loss. Let $\mathscr{L}_{\mathrm{C}}$ be the classification loss function. Let $\mathscr{L}_{\mathrm{D}}$ be the distillation loss function. Let $\mathscr{L}$ be the total loss function, where λ be the custom hyperparameter. Generally, let λ be 1. They are defined as follows:

$$\mathscr{L}_{\mathrm{C}} = y_{\mathrm{n}} \times \log \hat{y}_{\mathrm{n}}, \tag{14}$$

$$\mathscr{L}_{\mathrm{D}} = y_{\mathrm{o}} \times \log \hat{y}_{\mathrm{o}}, \tag{15}$$

$$\mathscr{L} = \mathscr{L}_{\mathrm{C}} + \lambda \mathscr{L}_{\mathrm{D}} + \mathscr{R}. \tag{16}$$

y_{n} is the label of new data, which uses one-hot encoding [27]. As a hard label, it gives the accurate category of the sample. $\hat{y}_{\mathrm{n}}$, $\hat{y}_{\mathrm{o}}$, and y_{o} are the vectors of the probabilities of each label. The first two are the calculated results, which are the prediction of the new model on the new data. The last is the prediction of the old model on the new data, which is a soft label. $\mathscr{R}$ is the regular term of the parameter. To adapt to the distillation loss, we made further improvements to the conditional probability function $P(Y \mid X; \Theta, T)$ of the sHCRF model:

$$P(Y \mid X; \Theta, T) = \frac{e^{\phi(y,H,X;\Theta)/T}}{\sum_y e^{\phi(y,H,X;\Theta)/T}}. \tag{17}$$

The newly added temperature coefficient T can control the smoothness of the probability distribution of the output. y_o, $\hat{y}_n$, and $\hat{y}_o$ are defined as follows:

$$y_{\mathrm{o}} = \left(P\left(Y_i \mid X; \Theta_{\mathrm{old}}, T = t\right) \mid Y_i \in Y_{\mathrm{old}}\right), \tag{18}$$

$$\hat{y}_{\mathrm{n}} = \left(P\left(Y_i \mid X; \Theta_{\mathrm{new}}, T = 1\right) \mid Y_i \in Y_{\mathrm{new}}\right), \tag{19}$$

$$\hat{y}_{\mathrm{o}} = \left(P\left(Y_i \mid X; \Theta_{\mathrm{new}}, T = t\right) \mid Y_i \in Y_{\mathrm{old}}\right). \tag{20}$$

y_{o} and $\hat{y}_{\mathrm{o}}$ are shown in Equations (18) and (20), where the temperature coefficient $T = t$, $t > 1$. $\hat{y}_{\mathrm{n}}$ is shown in Equation (19), and the temperature coefficient $T = 1$. Y_{old} is the label set of the old data, and Y_{new} is the label set of the new data.

The knowledge distillation algorithm of the sHCRF model does not need the old data to participate in the training, and the model can retain the old knowledge while fitting the new data. The new model adds new classification parameters based on the old model. Before training, the old model is used to predict the new human motion data to obtain the soft label of the old knowledge. The temperature coefficient of the prediction function $T = t\ (t > 1)$, then the soft label is used as a kind of "pseudo label." The new data and the correct label are used as the input of the loss function. In the training process, the new model uses the prediction function of temperature coefficient $T = t\ (t > 1)$ and $T = 1$ to predict the new data. The prediction result of the former and the "pseudo label" do cross-entropy to obtain the

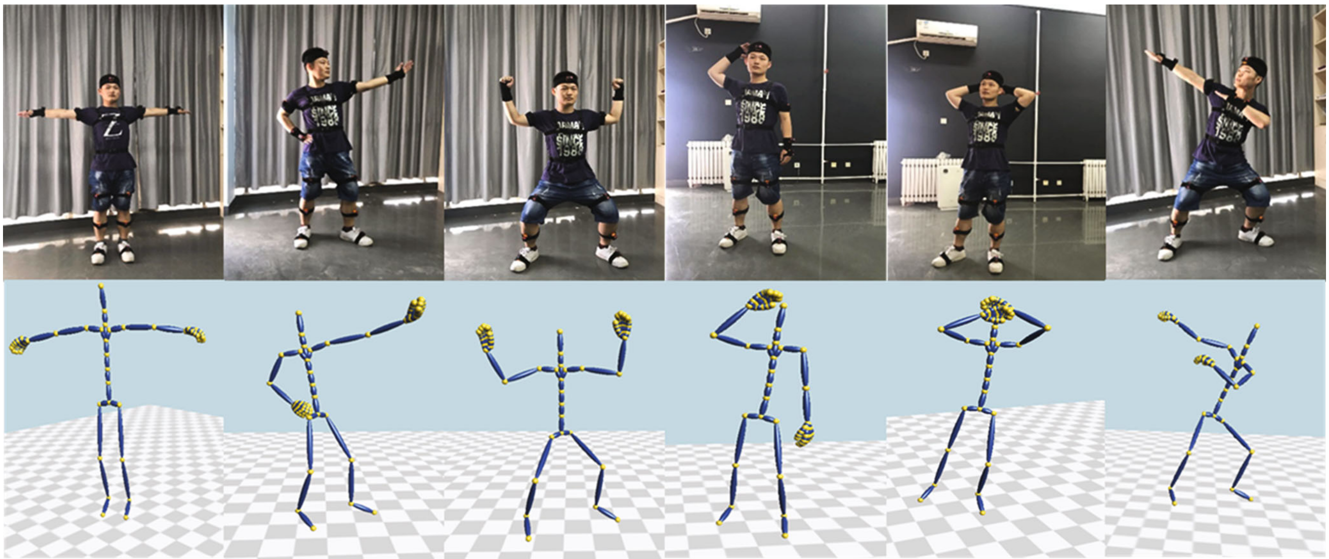


Figure 8: Motion capture and data reconstruction.

model's distillation loss, and the latter's prediction result is compared with the correct label. The cross-entropy obtains the classification loss of the model, and the two and the regular parameter terms constitute the total loss of the model.

5. Experiments

5.1. Human Motion Data Introduction. We captured the human motion data by 17-sensor motion capture equipment developed by our work group and reconstructed the captured information into skeleton data, as shown in Figure 8. The human skeleton comprises 17 joints. All the data samples take the T-Posture as the initial posture, in which the T-Posture shows that the human body is upright and the arms are held flat. Table 2 shows the dataset *Data0*, which contains all the motion data used in the experiment, including Walk, Run, Soccer, Basketball, Jump, Jump Forward, Dance, Yoga, Sit, Pick, Swing, Clean, Yan Fei, and Push-up. The dataset has 14 motion categories. The number of frames is 1,056,405, and the total number of samples is 984.

The datasets listed in Table 3 are subsets of the dataset *Data0* in Table 2. Each subset called *DataX* is divided into old and new parts, representing old data in *DataX_Old* and new data in *DataX_New*. When the datasets in Table 3 display the classes they contain, the class name is represented by a numerical label, which corresponds to the class number in *Data0*. In *Data1*, the old part contains the classes of Walk, Run, and Jump. The new part contains the classes of Soccer, Basketball, Dance, and Swing. In *Data2*, the old part contains the classes of Soccer, Basketball, Dance, and Swing. The new part contains the classes of Walk, Run, and Jump. *Data1* and *Data2* have the same data categories, and *Data2* exchanged old and new data on the basis of *Data1*. In *Data3*, the old part contains the classes of Walk, Run, and Jump. The new part contains the classes of Swing, Yan Fei, and Push-up. The degree of overlap between old and new data in *DataX* is defined by comparing the similarity of pose semantic sets of all action classes in *DataX_Old* and *DataX_New*. It is affected by two factors: one is the proportion of the intersection of posture semantic base sets in the old and new data in the old data, and the second is the proportion of the intersection of the posture semantic base sets of the new and old data in the key posture semantic base in the old data. The high overlap degree in *Data1* shows that more than 90% of posture semantic bases in the old data

Table 2: The whole motion dataset *Data0* in the experiment.

No.	Class name	Frames	Samples
1	Walk	57369	150
2	Run	23082	150
3	Soccer	13228	24
4	Basketball	51764	56
5	Jump	88572	152
6	Jump Forward	31624	72
7	Dance	246308	120
8	Yoga	80632	24
9	Sit	136066	44
10	Pick	32458	68
11	Swing	17528	40
12	Clean	109188	24
13	Yan Fei	93970	28
14	Push-up	64616	32

Table 3: Subsets of *Data0*. These subsets can be divided into *DataX_Old* and *DataX_New* sets, where *X* is the serial number of the subset to which they belong.

DataX (dataset=)	Old (*DataX_Old*=)	New (*DataX_New*=)
Data1	1, 2, 5	3, 4, 7, 11
Data2	3, 4, 7, 11	1, 2, 5
Data3	1, 2, 5	11, 13, 14

appear in the new data. The high overlap degree in *Data2* shows that the intersection of the posture semantic sets of the old and new data coincides with most of the key posture semantic bases in the old data. The overlap degree in *Data3* is less than that in the previous two datasets, indicating that more than half of the posture semantic bases in *DataX_Old* is unique to the old data.

5.2. Experiments. We designed three experiments in this paper. The first explored the sHCRF model's performance, including the model's training efficiency and the classification accuracy of the model. The second explored the memory ability of the sHCRF model. The last explored the energy consumption of the sHCRF model. In each experiment, we selected three-quarters of each type of motion samples as training samples.

The experiment of the sHCRF model's performance: this experiment used the HCRF model and the mHCRF model as the benchmarks to verify the performances of the sHCRF model in classification and training. The input of the HCRF model was only the posture sequence, while both the mHCRF model and the sHCRF model had two types of input. One was included three types of features: posture sequence, posture change sequence, and posture semantic sequence. The other lacked a posture change sequence based on the previous one.

The experiment of the sHCRF model's memory ability: this experiment had three steps. In the first step, we compared the total classification accuracy of the model with the distillation loss using different temperature T. In the second step, we compared the classification accuracy of the model with and without the distillation loss. The experiment was carried out in the datasets with different overlap degrees. We first let the model fit the old data. Then, we initialized the new classification layer parameters corresponding to the new category based on the model fitted to the old data. We used the model to fit the new data with and without distillation loss, respectively. Finally, we used the updated two new models to test on the test set containing all data, respectively. In the third step, we compared the total classification accuracy with and without distillation of the sHCRF model in incremental learning.

The experiment of energy consumption of the sHCRF model: this experiment used the HCRF model and the mHCRF model as the benchmarks to show that the sHCRF model can reduce the energy consumption. The results were presented in the form of energy consumption ratio.

Table 4 shows the settings of hyper parameters in the experience. The first column in the table represents where the hyperparameters are located.

6. Results and Discussion

6.1. Performance of the sHCRF Model. The experiment to verify the model's performance first tested the classification accuracy of the model. Then, we tested the model's training efficiency in the multiclass action classification task, and the model's parameters' amount. In the experiment, the training efficiency was represented by the single iteration time, which had been normalizing. The number of categories in the dataset ranges from 3 to 14 of Table 2. When the number of categories increased, the input samples also increased. Multiple experiments were conducted on each task by using data from random categories in the *Data0* dataset. Since the parameter optimization of the model may achieve local optimization during training, multiple experiments were needed to obtain optimal results. In the comparative experiments with the baseline: HCRF model and mHCRF model, each model's classification accuracy and the single iteration time with different multiclass tasks were compared. Among them, the mHCRF model was divided into two types according to whether its input included posture change sequences, namely, mHCRF(1) and mHCRF(2). The mHCRF(1) indicates that the input of the mHCRF model does not include the posture change sequence, while the input of the mHCRF(2) model includes the posture change sequence. In addition, the posture semantic sequence is used as input in the mHCRF model in this paper, but in the original method, the annotation of HMM is the input. The sHCRF model was divided into two types according to whether its input included posture change sequences, namely, sHCRF(1) and sHCRF(2). The sHCRF(1) indicates that the input of the sHCRF model does not include the posture change sequence, while the input of the sHCRF(2) model includes the posture change sequence.

Table 4: The settings of hyper parameters in the experience.

Parameters' location	Hyper parameters	Value
Equations (1) and (5)	M	13
$\mathbf{h}^K$	K	7
$\mathscr{H}_D$	D	128
T_C	C	16484
Equation (16)	λ	1
Equations (18) and (20)	T	2

Figure 9 shows the classification accuracy of the model in the multiclass action classification task, and Figure 10 shows the single iteration time in the multiclass action classification task. When only considering the influence of input feature, as shown in the results of sHCRF(2), sHCRF(1), mHCRF(2), mHCRF(1), and HCRF in Figures 9 and 10, the structure of the sHCRF(2) model is the same as the mHCRF(2) model. The classification accuracy of the sHCRF and the mHCRF models was higher than that of the HCRF model, the accuracy of the sHCRF and mHCRF models was more than 90%, and the accuracy of HCRF was about 80%. The mHCRF(2) model had the higher classification accuracy than the mHCRF(1), and the accuracy of the mHCRF(2) was more than 95%. The sHCRF(2) model had the higher classification accuracy than the sHCRF(1), and the accuracy of the sHCRF(2) was more than 95%. The single iteration time of the sHCRF(1) is the shortest, followed by the mHCRF(1). The HCRF took the most single iteration time. When only considering the influence of the model's structure, compare the results of sHCRF(2) and mHCRF(2) in Figures 9 and 10. The accuracy of the sHCRF(2) model

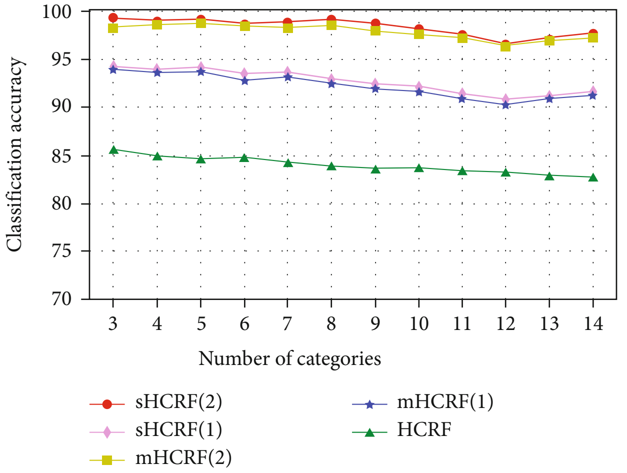


Figure 9: Classification accuracy results of each model, when the sHCRF model and the benchmark models are used to identify tasks on datasets with different numbers of classes.

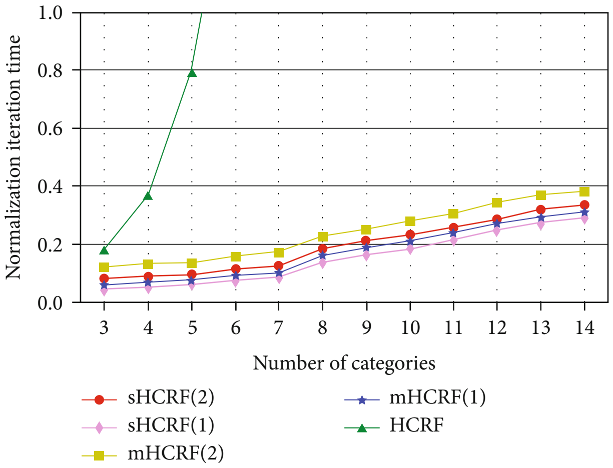


Figure 10: Normalization iteration time of each model, when the sHCRF model and the benchmark models identify tasks on datasets with different numbers of categories.

was higher than that of the mHCRF(2) model. The accuracy of the mHCRF(2) model was slightly inferior. The single iteration time of the sHCRF(2) model was shorter than that of the mHCRF(2) model. Generally, the improvement of input features contributed significantly to the classification accuracy, and the improvement of the model's structure can speed up the single iteration time.

As shown in Figure 9, when the number of categories increased, the classification accuracy of the sHCRF(2) model stayed above 95%. The classification accuracy of the mHCRF(2) model was similar to that of the sHCRF(2) model. The accuracy of the mHCRF(2) model was slightly inferior. The accuracy of the mHCRF(1) model is lower than that of the previous two models. However, the lowest accuracy is still 90%. The accuracy of the mHCRF(1) model was slightly lower than that of the sHCRF(1) model. The classification accuracy of the HCRF model remained above 80%. But as the number of categories increased, the classification accuracy rate gradually decreased. The model maintained a high classification accuracy.

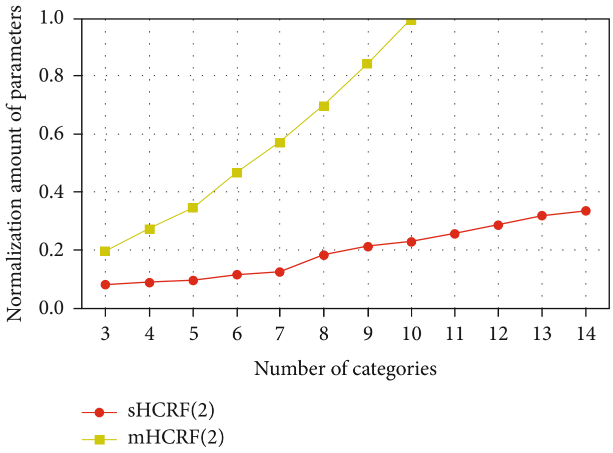


Figure 11: Normalization amount of parameters of each model, when the sHCRF model and the benchmark model identify tasks on datasets with different numbers of categories.

As shown in Figure 10, the single iteration time of the sHCRF(2), sHCRF(1), mHCRF(1), and mHCRF(2) models increased steadily with the increase in the number of categories. The single iteration time of the sHCRF(2) model was shorter than that of the mHCRF(2) model, and the single iteration time of the sHCRF(1) model was shorter than that of the mHCRF(1) model. Unlike the three models, the single iteration time of the HCRF model increases rapidly. The HCRF model spends more time on datasets with more than five categories than is displayed.

As shown in Figure 11, there were the results of the parameters' amount of the sHCRF(2) and mHCRF(2) models. The sHCRF(2) model had the minimum amount of parameters, followed by the mHCRF(2) model. When the number of categories increased, the amount of parameters of the mHCRF(2) model rapidly increased, and its parameters' amount was more than the sHCRF(2) model.

Based on the results of the first two performance indicators obtained from the comprehensive experiment, the sHCRF model and other undirected probability graph models have shown excellent performance in the comparison experiment of the motion classification task. It has good advantages in classification accuracy and training speed.

6.2. Performance of the sHCRF Model in Incremental Learning Scenarios. In the experiment of the performance of the sHCRF model in the incremental learning scenario, the study first tested the classification accuracy of the model with the distillation loss using different temperature T. The result is as shown in Figure 12. In Figure 12, when the value of T is not greater than 5, the old data classification accuracy fluctuates slightly around 80%. The new data classification accuracy is about 90%, and the total accuracy is around 85%. When the value of T increases gradually from 5, the old data classification accuracy began to decline significantly. The same applies to other results, the decline rate of the new data classification accuracy is relatively gentle. When using the distillation loss, the value of T will affect

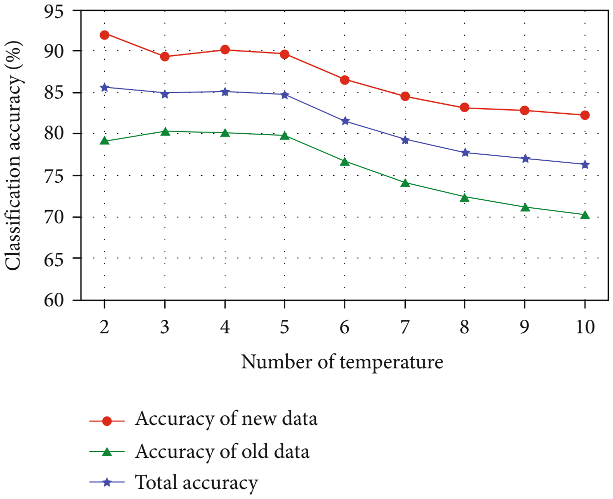


Figure 12: The classification accuracy of the model with distillation loss using different temperature T.

the classification accuracy of the model. The memory ability of old knowledge declines when the value of T exceeds a certain threshold. The threshold is related to the size of the model.

We also verified the classification accuracy of the model with and without distillation loss term, respectively. There were three datasets in the verification experiment, *Data1*, *Data2*, and *Data3*. The classification accuracy of old data is used as the experimental result in Figure 13. In the experiments on *Data1* and *Data2*, the results were 6.9% and 5.8%, respectively. Its ability to retain old knowledge is almost nonexistent. But the effects on *Data3* are not. On *Data3*, the model's classification accuracy of the old data is 33.4%, showing a specific ability to retain old knowledge, as determined by the sHCRF model's structure. When the overlap of *Data3* is low, fitting the new data has little impact on the old knowledge, which means that the parameters related to the old knowledge had not changed much. So, the old knowledge structure of the model was not completely destroyed, and the model has a particular memory ability.

After adding the distillation loss term, the model's memory ability improved further. This experiment was carried out on *Data1*, *Data2*, and *Data3*. The comparison between the classification accuracy of the model on *Data1*, *Data2*, and *Data3* with and without the distillation loss was given, as shown in Figure 13. In Figure 13, the model's classification accuracy of the old data on *Data1* has increased from 6.9% to 86.7%. The old data classification accuracy on *Data2* has increased from 5.8% to 78.8%. The old data classification accuracy on *Data3* has increased from 33.4% to 72.9%. In both *Data1* and *Data2*, the preserving rate of old knowledge of the model has been greatly improved, which effectively inhibits the forgetting of old knowledge by the model. The preserving rate of old knowledge of the model has been improved on *Data3*. The new data classification accuracy on *Data1* is from 98.9% to 94.3%. The new data classification accuracy on *Data2* is from 98.5% to 90.1%. The new data classification accuracy on *Data3* is from 96.0% to 88.0%. The accuracy still maintains high. The model has similar performance in the cognition accuracy of new and old data because of the balance between the model's classification loss and distillation loss. When the overlap degree is high, the distillation loss plays a more significant role in maintaining the old knowledge.

In Table 5, there is the comparison of the total accuracy of both the old and new data with and without distillation. On *Data1*, the total accuracy increased from 52.8% to 90.5% by adding the distillation loss. On *Data2*, the total accuracy is 52.2% without distillation, and it has improved to 84.5% by adding the distillation loss. On *Data3*, the total accuracy increased from 64.7% to 80.5% by adding the distillation loss.

We compared the total classification accuracy with and without distillation of the sHCRF model during incremental learning. The results are shown in Figure 14. We first set up the original model, which could classify three categories. When the categories were incremental, the classification accuracy of the model without distillation rapidly dropped to about 50%. The classification accuracy was around 50% as categories increased since then. The classification accuracy of the model with distillation also declined. The classification accuracy of the model was above 70% as categories increased.

6.3. Energy Consumption Analysis. We combined the training efficiency and the amount of parameters to estimate the ratio of the energy consumption of the sHCRF model to that of the baseline models. The energy consumption included the computation energy and the storage energy [28]. The computation energy was estimated from the computational complexity of the model, and the storage energy was estimated from the amount of parameters of the model. The experiment consisted of two contrasting experiments, the comparison of the sHCRF model and the mHCRF model, and the comparison of the sHCRF model and the HCRF model. In the experiment, we tested on datasets with different numbers of action categories. The results are shown in Figure 15.

As shown in Figure 15, "S/M" represents the comparison of the sHCRF model and the mHCRF model, and it represents a ratio. In the comparison of the sHCRF model and the mHCRF model, when the number of categories increased, the proportion of the storage energy dropped, and the sHCRF model consumed less storage energy than the mHCRF model. The proportion of the computation energy was less than 1. It means that the sHCRF consumed less computation energy. "S/H" represents the comparison of the sHCRF model and the HCRF model, and it represents a ratio. In the comparison of the sHCRF model and the HCRF model, the sHCRF model consumed less storage energy than the HCRF model. When the number of categories increased, the proportion of the computation energy dropped and almost dropped to 0.

Based on the results of the energy consumption, when the number of categories increased, the single iteration time and the amount of parameters of the sHCRF model were lower than those of the mHCRF model and the HCRF model. It could be estimated that the computation energy

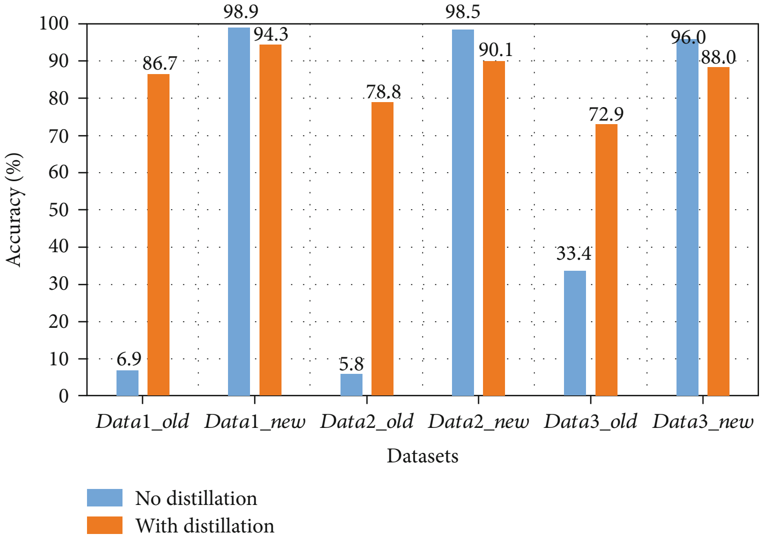


Figure 13: Comparison of the classification accuracy of the new and old data with and without distillation of the sHCRF model.

Table 5: The total accuracy of the sHCRF model of both the old and the new data with and without distillation.

Dataset name	Total accuracy (%)	
	No distillation	With distillation
Data1	52.8	90.5
Data2	52.2	84.5
Data3	64.7	80.5

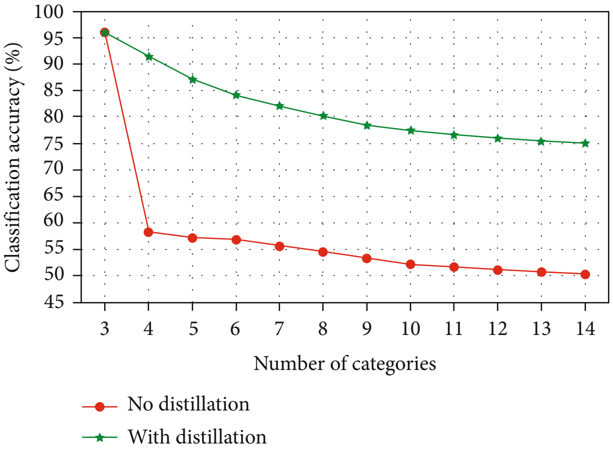


Figure 14: Comparison of the total classification accuracy with and without distillation of the sHCRF model in incremental learning process.

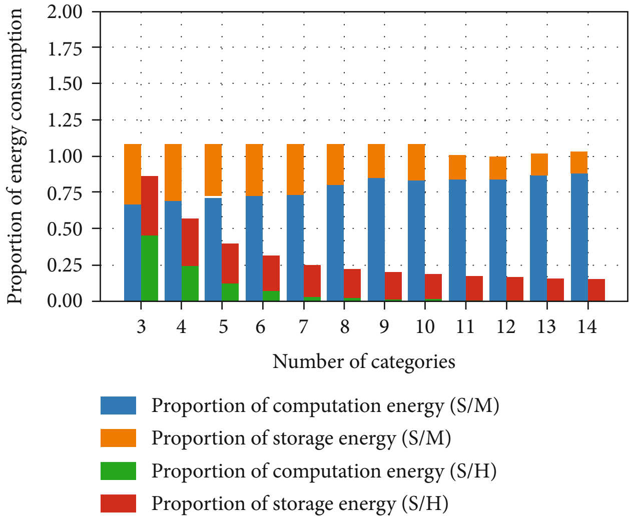


Figure 15: Energy consumption of each model, when the sHCRF model and the benchmark model identify tasks on datasets with different numbers of categories.

and the storage energy of the sHCRF model were less than those of the mHCRF model and the HCRF model. Combining the computation energy and the storage energy, it can be considered that the sHCRF model consumes less energy. The application system implemented based on the sHCRF model can save the computation and storage energy in terms of model training. The specific energy consumption also depends on the energy consumption of the chips, and related hardware. The energy consumption measurement of an application system will be the future work.

7. Conclusions

In this paper, we proposed a lightweight action classification method for Green IoT sport applications. We designed motion features which could describe the spatial and temporal information of the motion data. We proposed a human action classification model, namely, sHCRF, which can be applied to incremental learning scenarios. The model achieves the purpose of energy efficiency by reducing computation overhead and amount of sample data required for training. In general classification tasks, the model's classification accuracy is more than 95%. In the incremental learning scenario, this paper verifies that it can preserve the old knowledge. In the case of category increment, the preserving rate is about 70%. The model's learning ability balances the old and new knowledge. This can effectively control the growth rate of model capacity.

The knowledge distillation method used in this paper depends on the similarity of the training data. If the actions' distinctives are obvious, the effect of the traditional knowledge distillation algorithm is not significant. Further work will improve the model's preservation rate of old knowledge in incremental learning scenarios.

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 research was supported by the National Natural Science Foundation of China (NSFC, No. 62177005) and by the National Key Research and Development Program of China funded by the Ministry of Science and Technology (No. 2020YFC2007200 and No. 2020YFF0305200).

References

[1] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[2] F. Q. Zhang and L. J. Ma, "Multi-object crowd real-time tracking in dynamic environment based on neural network," *Journal of Network Intelligence*, vol. 7, no. 2, pp. 386–394, 2022.

[3] P. Agarwal and M. Alam, "A lightweight deep learning model for human activity recognition on edge devices," *International Conference on Computational Intelligence and Data Science*, vol. 167, pp. 2364–2373, 2020.

[4] S. Rani, H. Babbar, S. Coleman, A. Singh, and H. M. Aljahdali, "An efficient and lightweight deep learning model for human activity recognition using smartphones," *Sensors*, vol. 21, no. 11, p. 3845, 2021.

[5] Y. Jiang, X. Yang, J. Liu, and J. Zhang, "A lightweight hierarchical model with frame-level joints adaptive graph convolution for skeleton-based action recognition," *Security and Communication Networks*, vol. 2021, Article ID 2290304, 13 pages, 2021.

[6] K. Forbes and E. Fiume, "An efficient search algorithm for motion data using weighted PCA," in *Proceedings of the 2005 ACM SIGGRAPH / Eurographics symposium on Computer animation (SCA)*, pp. 67–76, Association for Computing Machinery, New York, NY, USA, 2005.

[7] M. Müller, T. Röder, and M. Clausen, "Efficient content-based retrieval of motion capture data," in *ACM SIGGRAPH 2005 Papers (SIGGRAPH)*, pp. 677–685, Association for Computing Machinery, New York, NY, USA, 2005.

[8] X. Liang, S. Zhang, Q. Li, N. Pronost, W. Geng, and F. Multon, "Intuitive motion retrieval with motion sensors," in *Proceedings of Computer Graphics International*, pp. 64–71, New York, USA, 2008.

[9] H. Ehtesham, "Learning video actions in two stream recurrent neural network," *Pattern Recognition Letters*, vol. 151, pp. 200–208, 2021.

[10] T. Teixeira, E. Granger, and A. Koerich, "Continuous emotion recognition with spatiotemporal convolutional neural networks," *Applied Sciences*, vol. 11, no. 24, article 11738, 2021.

[11] J. Li, X. Liu, M. Zhang, and D. Wang, "Spatio-temporal deformable 3D ConvNets with attention for action recognition," *Pattern Recognition*, vol. 98, pp. 107037–107037, 2020.

[12] J. Shi, C. Liu, C. Ishi, and H. Ishiguro, "Skeleton-based emotion recognition based on two-stream self-attention enhanced spatial-temporal graph convolutional network," *Sensors*, vol. 21, no. 1, pp. 205–216, 2021.

[13] K. Cheng, Y. Zhang, X. He, W. Chen, J. Cheng, and H. Lu, "Skeleton-based action recognition with shift graph convolutional network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 183–192, Seattle, WA , USA, 2020.

[14] D. Tian, Z. Lu, X. Chen, and L. H. Ma, "An attentional spatial temporal graph convolutional network with co-occurrence feature learning for action recognition," *Multimedia Tools and Applications: An International Journal*, vol. 79, no. 17-18, pp. 12679–12697, 2020.

[15] K. Muhammad, A. U. Mustaqeem, A. Ullah et al., "Human action recognition using attention based LSTM network with dilated CNN features," *Future Generation Computer Systems*, vol. 125, pp. 820–830, 2021.

[16] C. Ma, D. Yu, and H. Feng, "Recognition of badminton shot action based on the improved hidden Markov model," *Journal of healthcare engineering*, vol. 2021, Article ID 7892902, 8 pages, 2021.

[17] A. Samr and B. Nizar, "Multimodal action recognition using variational-based Beta-Liouville hidden Markov models," *IET Image Processing*, vol. 14, no. 17, pp. 4785–4794, 2020.

[18] R. Wen, Q. Wang, and Z. Li, "Human hand movement recognition using infinite hidden Markov model based sEMG classification," *Biomedical Signal Processing and Control*, vol. 68, article 102592, 2021.

[19] M. Vrigkas, E. Kazakos, C. Nikou, and I. A. Kakadiaris, "Human activity recognition using robust adaptive privileged probabilistic learning," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 915–932, 2021.

[20] J. Zhang and S. Gong, "Action categorization with modified hidden conditional random field," *Pattern Recognition*, vol. 43, no. 1, pp. 197–203, 2010.

[21] A. Zaman, F. Yangyu, M. Irfan, M. S. Ayub, L. Guoyun, and L. Shiya, "LifelongGlue: keypoint matching for 3D reconstruction with continual neural networks," *Expert Systems With Applications*, vol. 195, 2022.

[22] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 2935–2947, 2018.

[23] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, https://arxiv.org/abs/1503.02531.

[24] M. Irfan, Z. Jiangbin, M. Iqbal, Z. Masood, and M. H. Arif, "Knowledge extraction and retention based continual learning by using convolutional autoencoder-based learning classifier system," *Information Sciences*, vol. 591, pp. 287–305, 2022.

[25] J. Kirkpatrick, R. Pascanu, N. Rabinowitz et al., "Overcoming catastrophic forgetting in neural networks," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, 2017.

[26] D. Lopez-Paz and M. A. Ranzato, "Gradient episodic memory for continual learning," *Advances in Neural Information Processing Systems*, vol. 30, pp. 6467–6476, 2017.

[27] K. Raajitha, K. Meenakshi, and Y. M. Rao, "Design of thermometer coding and one-hot coding," in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV),*, pp. 601–609, Tirunelveli, India, 2021.

[28] T. J. Yang, Y. H. Chen, J. Emer, and V. Sze, "A method to estimate the energy consumption of deep neural networks," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pp. 1916–1920, Pacific Grove, CA, USA, 2017.