Security and Communication Networks

Big Data-Driven Multimedia Analytics for Cyber Security

Lead Guest Editor: Zhaoqing Pan
Guest Editors: Yuan Tian, Vidyasagar Potdar, and Md. Hasanul Kabir

Big Data-Driven Multimedia Analytics for Cyber Security

Security and Communication Networks

Big Data-Driven Multimedia Analytics for Cyber Security

Lead Guest Editor: Zhaoqing Pan
Guest Editors: Yuan Tian, Vidyasagar Potdar, and Md. Hasanul Kabir

Chief Editor

Associate Editors

Academic Editors

Contents

Research on Information Resource Sharing and Big Data of Sports Industry in the Background of OpenStack Cloud Platform
Chuan Mou and Ye Cheng
Research Article (12 pages), Article ID 2824146, Volume 2021 (2021)

Malicious Encryption Traffic Detection Based on NLP
Hao Yang, Qin He, Zhenyan Liu, and Qian Zhang
Research Article (10 pages), Article ID 9960822, Volume 2021 (2021)

SFRNet: Feature Extraction-Fusion Steganalysis Network Based on Squeeze-and-Excitation Block and RepVgg Block
Guiyong Xu, Yang Xu, Sicong Zhang, and Xiaoyao Xie
Research Article (11 pages), Article ID 3676720, Volume 2021 (2021)

Research on Calibration Method of Binocular Vision System Based on Neural Network
Hao Zhu, Mulan Wang, and Weiye Xu
Research Article (9 pages), Article ID 5542993, Volume 2021 (2021)

Quantitative Weighted Visual Cryptographic (k, m, n) Method
Yewen Wu, Shi Zeng, Bin Wu, Bin Yang, and Xianyi Chen
Research Article (13 pages), Article ID 9968652, Volume 2021 (2021)

Delegated Key-Policy Attribute-Based Set Intersection over Outsourced Encrypted Data Sets for CloudIoT
Yanfeng Shi and Shuo Qiu
Research Article (11 pages), Article ID 5595243, Volume 2021 (2021)

Improved CNN-Based Hashing for Encrypted Image Retrieval
Wenyan Pan, Meimin Wang, Jiaohua Qin, and Zhili Zhou
Research Article (8 pages), Article ID 5556634, Volume 2021 (2021)

Recognition of Disease Genetic Information from Unstructured Text Data Based on BiLSTM-CRF for Molecular Mechanisms
Lejun Gong, Xingxing Zhang, Tianyin Chen, and Li Zhang
Research Article (8 pages), Article ID 6635027, Volume 2021 (2021)

Research on Multidomain Authentication of IoT Based on Cross-Chain Technology
Dawei Li, Jia Yu, Xue Gao, and Najla Al-Nabhan
Research Article (12 pages), Article ID 6679022, Volume 2020 (2020)

Hindawi
Security and Communication Networks
Volume 2021, Article ID 2824146, 12 pages
https://doi.org/10.1155/2021/2824146

Research Article

Research on Information Resource Sharing and Big Data of Sports Industry in the Background of OpenStack Cloud Platform

Chuan Mou[1] **and Ye Cheng**[2]

[1]*College of Physical Education, Kunsan National University, Kunsan 54150, Jeonro-do, Republic of Korea*
[2]*Department of Physical Education, Hebei University of Economics and Business, Shijiazhuang 050061, Hebei, China*

Correspondence should be addressed to Ye Cheng; chengye198926@163.com

Received 10 May 2021; Revised 1 July 2021; Accepted 7 August 2021; Published 15 August 2021

Academic Editor: Yuan Tian

The rapid development of information technology and Internet makes the sports information resources retrieval service more convenient and quick; sports policy in recent years lays a foundation for the development of the Internet + sports, the development of sports industry in the process of our country economy level of development status, and the development of sports industry into the era of information and big data. This paper takes OpenStack cloud platform as the research basis (1) to realize the sharing of sports industry information resources in OpenStack cloud technology and (2) to realize big data analysis of sports industry and (3) empirical research on big data of sports industry. The main content is to realize the construction of sports resources informatization based on the OpenStack cloud platform. Through the analysis and empirical study of the big data of the sports industry, the influence of the development of the sports industry in the process of China's economic development is discussed. In this paper, the experimental results show that the sports industry showed a positive impact in the process of economic development, the sports economy for the development of the economy, the contribution rate reached 11.77%, the sports industry for the development of the economy, the pull rate of 1.056%, based on the cloud platform of information resources sharing of data analysis, sports industry for the development of the economy has a positive role in promoting.

1. Introduction

The rapid development of economy leads to the rapid growth of data. The public resource information sharing and the healthy development of big data industry have attracted great attention of the country. At present, the total amount of data appearing and stored in China exceeds 800 exabytes, and China's e-government development index is 0.6071, ranking 63rd [1]. The development and utilization of big data resources have risen to the national level, and big data technology has been applied to various fields. Government departments put forward open public data resources sharing, improve big data infrastructure construction and development of big data in various fields of application, the formation of large data product system, perfect the big data industry chain, and improve the big data security system, to ensure that big data [2].

In the context of sound economic development, the advantages of multiparty information resource sharing and acquisition have been highlighted, and people's living standards have been greatly improved, which also puts forward higher demands for the sports industry. It has become an important force to promote the sustainable development of economy and society [3]. The General Administration of Sport of China clearly proposed "Internet + sports" in the Outline of the 13th Five-Year Plan, officially opening the development process of the integration of the Internet and sports. The Opinions on Strengthening the Construction of National Fitness Facilities and Developing Mass Sports issued by the General Office of the State Council in 2020 is proposed to strengthen the construction of national fitness facilities, develop mass sports, promote the construction of fitness facilities, promote the vigorous development of mass sports, and improve the public service level of national fitness [4]. In the era of advancing with The Times of Information Technology, sports and the depth of the fusion of the Internet have become a historic turning point in the development of sports industry to encourage the development of the national fitness and sports venue construction process; the development of sports industry and sports information resource sharing becomes hotspot; under the background of big data, the analysis and processing of massive data information are a new challenge for the development of sports industry. The sharing and development of sports information resources based on cloud platform and the analysis and research of sports big data are of great significance.

Liu and Wu conducted experiments and deployment of the cloud platform for small- and medium-sized enterprises through OpenStack open source cloud framework in order to solve the construction needs of the information service platform for small- and medium-sized enterprises and get rid of the problems of decentralized management and low resource utilization of traditional enterprise information system in the study of the service cloud platform architecture for small- and medium-sized enterprises based on OpenStack. The overall framework of the service platform for small- and medium-sized enterprises is proposed, and the research results show that OpenStack is a private cloud platform more suitable for the construction of small- and medium-sized enterprises [5]. Yi et al. applied the deployment and management of OpenStack cloud platform in the course teaching reform and practice. OpenStack deployment is mainly based on private cloud computing technology, and computing, storage, and network resources are provided to users through the deployment of OpenStack cloud computing environment [6]. Tang and Huang, in the study of high availability of OpenStack-based cloud platform, analyzed high availability cluster and load balancing technology, designed Pacemaker + HAProxy technical solution, and deployed high availability of OpenStack's core component database and mirror service. The high availability of OpenStack cloud platform has been realized [7]. Zhong proposed in his study that OpenStack cloud experiment platform has a high utility ratio and a low expansion cost, which can effectively improve the efficiency of teaching and scientific research experiments. The architecture of OpenStack platform includes user interface, various services, hardware resources, and monitoring and configuration management tools. It has great potential in distributed computing, virtual desktop, and other applications and can provide flexible teaching and research experimental computing services for universities and enterprises [8]. Wang and Wu studied the construction of the innovation and entrepreneurship cloud platform for local universities based on OpenStack. Starting from the construction status of the innovation and entrepreneurship platform for local universities, they adopted the open source KVM technology to realize computing virtualization and used the storage capacity of Ceph integrated server to realize storage virtualization, thus providing various kinds of virtual innovation and entrepreneurship practice resources. Compared with traditional innovation and entrepreneurship platforms, the innovation and entrepreneurship cloud platform has the characteristics of easy expansion, high resource integration, and low construction cost, which can effectively meet the different demands of local universities for innovation and entrepreneurship resources [9]. Wen et al. analyzed the technical principles of cloud computing and open source cloud platform OpenStack in their research. First, they design and implement a platform that can virtualize and centralize the network with the method of layered design by referring to the traditional network management mode and scheme and test it. Secondly, the platform can quickly deploy and manage the virtual network and create virtual machines according to the needs of users, so as to realize the allocation of resources on demand and the efficient rate [10]. Xin et al. studied the Glance service in OpenStack platform and put forward the position of Glance service in OpenStack platform and the relationship between it and other services and proposed an OpenStack-based private cloud construction scheme suitable for small- and medium-sized enterprises [11]. Yao and Yan combined with a variety of load balancing, IP drift, network routing redundancy protocol (VRRP), database cluster (Galera), distributed cluster (Cech), and other technologies in the research of OpenStack high availability distributed storage scheme design and implementation. This paper proposes a method that can realize the high availability of distributed platform storage scheme by means of business load balancing, service instance failover, and other methods in the case of single node failure, network congestion, and high data outbreak [12].

The innovations of this paper are as follows: (1) sports information resource sharing under the OpenStack cloud computing platform is proposed; (2) empirical analysis of sports industry big data from the perspective of information resource sharing is performed; (3) the healthy development of the information resources of the sports cloud platform is proposed.

The disadvantages of this paper are as follows: (1) research orientation: this study focuses on the data analysis of sports industry; (2) limitations of the study: this paper briefly explains the OpenStack cloud platform, but how to implement the analysis is not the focus of this study. Therefore, it does not give a detailed explanation of how to implement the analysis.

2. Research Method of Big Data in Sports Industry Based on OpenStack Cloud Platform

2.1. Cloud Platform Analysis of OpenStack

2.1.1. Overview of OpenStack. OpenStack is a free software and open source project developed by NASA and Rackspace under the Apache license to provide software for the construction and management of public and private clouds. OpenStack is a simple, massively scalable, and content-rich cloud computing operating system that supports almost all types of cloud environments. At the same time, OpenStack also provides framework standards and APIs, and service calls between all modules of OpenStack are realized through standard virtualized APIs. OpenStack is written with high-quality Python code, and the quality problems and security holes of the open source code are easier to be found and corrected [13]. When establishing the OpenStack system, it mainly includes core service components such as Keystone, Glance, Nova, Neutron, Horizon, and Swift.

Keystone is an authentication service that provides authentication and authorization services for other OpenStack services and an endpoint directory for all OpenStack services. As an OpenStack identity authentication module, Keystone provides a unified identity authentication service for OpenStack service and users, mainly responsible for authentication, service rules, and service instructions [14].

Glance is a mirroring service for storing and retrieving disk images of virtual machines, which OpenStack computing uses when the instance is deployed. Glance is an important mirror query, management, registration, and transport component, which needs to be used for cloud host creation, startup, snapshot, and other services [11].

Nova is a compute service for lifecycle management of compute instances in the OpenStack environment. On-demand response includes operations such as generation, scheduling, and recycling of virtual machines. Nova service, as the core module of the whole system, completes the management and control capabilities related to virtual machine instance lifecycle management, cloud platform network management, and cloud platform storage management [15].

Neutron ensures network services that provide network connectivity to other OpenStack services. In OpenStack calculation, API is provided for users to define the network and use it. Neutron based on plug-in architecture can support numerous network providers and technologies. Neutron networks can realize the flexible division of physical networks and provide each user with an independent network environment in a multiuser and multitenant environment [16].

Horizon provides a web-based management portal that interacts with OpenStack's underlying services to allocate IP addresses and configure access control and plays the role of a dashboard in self-service and management activities. Through the web interface, users can obtain resources of the platform and simplify user operations [17, 18].

Swift's HTTP-based application interface stores and arbitrarily retrieves unstructured data objects. With data replication and extensible architecture, you can store written objects and files to multiple hard disks [19]. As a service management module of computer, SWIFT has the function of using and deploying computer information resources and realizes the control of virtual server and system [20].

2.1.2. OpenStack System Architecture. From the functional point of view of OpenStack system, all the constituent subsystems and services of the project jointly provide IaaS services, and the subsystems and services can integrate and call each other through standardized public service interface API. Each system can be independently deployed and used, and the system is internally divided into three levels: API, logical processing (Manager), and underlying Driver adaptation [21].

The management of OpenStack cloud computing management platform includes visual project management of Dashboard, command-line management of Client Console, and management of technical developers of RESTful API [22].

The three-tier service system of cloud computing includes infrastructure-as-a-service, platform-as-a-service, and software-as-a-service. Cloud computing can centralize and share resources such as servers, networks, storage, and application software and provide them to computers, other devices, or end users on demand [23].

According to user needs, OpenStack service components can be selectively installed into multiple host nodes. The method of multihost nodes is also known as multinode deployment. To put it simply, different hosts install different service components to assume different service roles. The platform of multinode deployment has the characteristics of easy expansion, rich functions, and strong maneuverability.

2.1.3. OpenStack Deployment Architecture. The common deployment methods of OpenStack's multinode deployment platform include two-node deployment architecture and three-node deployment architecture. In a two-node architecture, the configuration node consists of a control node and a compute node. In a three-node architecture, the configuration node consists of control node, network node, and compute node. There is only one control node and it can control all services including Keystone, Glance, Nova, Neutron, and API services, MySQL database, and message system [24].

Yang proposed that OpenStack provides an open source software framework for the construction and management of cloud computing. By building a private teaching cloud in schools, software and hardware resources can be shared, jointly built, and integrated, so as to improve teaching quality, service level, and management efficiency [25].

Wu et al. proposed two practical deployment modes of OpenStack nodes, including master control node and computing node. This deployment mode is characterized by (1) easy expansion: the practical application may be extended due to the expansion of the business, and whether the deployment is simple and convenient affects the difficulty of the expansion; (2) easy maintenance: when the order of magnitude of the node increases, the difficulty of maintenance needs to be considered; (3) high stability: whether the architecture has high reliability and can provide services continuously and stably [24].

Li et al. proposed an OpenStack-based private cloud platform architecture scheme in their research on the design of OpenStack-based private cloud platform and believed that PackStack could quickly deploy a set of multinode OpenStack clusters on CentOS. It is characterized by simple operation and rapid deployment [26].

2.2. Research Methods of Big Data in Sports Industry

2.2.1. Index Selection. The layout and structure of the industry are controlled by certain economic and technological conditions. In order to promote the coordinated development of the social economy, it is necessary to adjust the industrial structure appropriately, improve the conversion ability of the industrial structure, and promote the development of the industrial structure to the direction of rationalization and advancement [27]. In this paper, contribution rate and pull rate are selected as indicators to study the development of the sports industry. By analyzing the total volume and structural characteristics of the sports industry, the influence relationship of the sports industry in the process of economic development is studied.

The contribution rate is the ratio of the increment of a part of the population to the increment of the population. In order to study the contribution rate of sports industry development to economic development, since the added value of sports industry is a part of regional GDP, the contribution rate can be represented by the ratio of the incremental value of sports industry to the incremental value of GDP [28]. The calculation formula is as follows:

$$\mathrm{SR}_t = \frac{Z_t - Z_{t-1}}{\mathrm{GDP}_t - \mathrm{GDP}_{t-1}} \times 100\%. \tag{1}$$

In the formula, t is the time, SR_t is the contribution rate of sports industry, Z_t is the total revenue of the sports industry in the current year, and GDP_t is the gross domestic product of the year.

Pull rate refers to the percentage of total growth driven by the growth of a certain factor [29]. In order to study the driving rate of sports industry for economic development, the calculation formula is as follows:

$$\mathrm{SP}_t = \mathrm{SR}_t \times \mathrm{GP}_t \times 100\%. \tag{2}$$

In the formula, t is the time, SP_t is the driving rate of sports industry, and GP_t is the GDP growth rate.

2.2.2. Data Source of Sports Industry. Data demand is the added value data of the economic development of the national sports industry, the relevant information of the economic development data of the national sports industry, and the statistical classification standard of the sports industry.

Data sources: data of the added value of the national sports industry from 2015 to 2019 are from Announcement of the Data on the Scale and the added value of the national sports industry jointly issued by the National Bureau of Statistics of the General Administration of Sport, and data of the national economic development from 2015 to 2019 are from Statistical Bulletin of National Economic and Social Development of the National Bureau of Statistics. Sports Industry Statistical Classification (2019) is adopted for the accounting classification of sports industry. Based on the Classification of Industries in the National Economy (GB/T 4754-2017), it is a reclassification of related activities conforming to the characteristics of sports industry in the classification of industries in the national economy [30].

Data analysis software requirements: Excel 2016 and SPSS 24.0.

2.2.3. Data Collection and Processing. The research takes the national sports industry economic index as the research object. According to the Statistical Classification of Sports Industry, the sports industry is divided into 11 categories. It mainly includes sports management activities, sports competition performances, sports fitness and leisure activities, sports venue services, sports intermediary services, sports education and training, sports media and information services, sports goods and related product agents, other sports services, sports goods and related product manufacturing, and sports facilities. The added value data of the national sports industry include the added value data and the proportion of each year according to the statistical classification standard of the sports industry. The national economic development data include economic indicators and years such as GDP, the number of employed people, social fixed asset investment, and per capita GDP. The data are summarized and sorted into Excel 2016 software, and descriptive statistical analysis is conducted on the collected data. In order to verify the appropriateness and reliability of the measured variables, exploratory factor analysis and confirmatory factor analysis were performed on standardized data using SPSS 24.0.

2.3. Research Ideas. This paper takes the information resource sharing and sports industry big data research under the background of OpenStack cloud platform as the theme, collects data, collates data, and determines research methods. The research contents include (1) information resource sharing of sports industry based on OpenStack platform, which mainly realizes resource sharing from the aspects of OpenStack system and node deployment; (2) big data analysis of sports industry, mainly analyzing the influence of sports industry on economic development from two aspects: contribution rate and pull rate. The specific process is shown in Figure 1.

3. Research on Information Resource Sharing of Sports Industry Based on OpenStack Cloud Platform

3.1. Design of Sports Information Resource Sharing Module

3.1.1. Asynchronous Deployment of Information Resource Sharing. As the control part of cloud organization, Nova component in OpenStack realizes the configuration of resources through virtual machine, which can manage the network and control the user's access to the cloud. The information resources of Nova service are very complex. During the deployment process, API node deployment and compute node deployment are the main ones. The deployment process of Nova service is as follows: the Nova-API receives requests from the client and the creation of the virtual machine, validates them, and forwards them to the Nova-Scheduler, who forwards the creation of the virtual machine to the Nova-Compute service on the selected node. Nova-The Compute service runs to create virtual machines on compute nodes, and Nova-Compute forwards messages to libvirt, which in turn hands off tasks to KVM and Xen.

It provides API services for users to implement initial deployment and virtual interaction functions and integrates

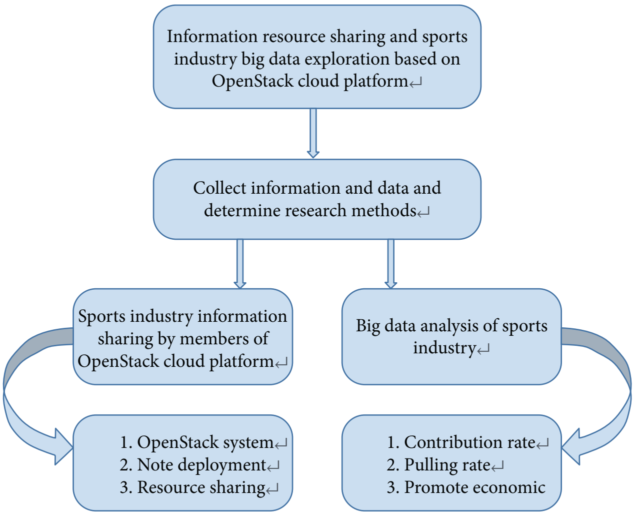

Information resource sharing and sports industry big data exploration based on OpenStack cloud platform↵
Collect information and data and determine research methods↵
Sports industry information sharing by members of OpenStack cloud platform↵
Big data analysis of sports industry↵
1. OpenStack system↵
2. Note deployment↵
3. Resource sharing↵
1. Contribution rate↵
2. Pulling rate↵
3. Promote economic

Figure 1: The research process.

sports information resources through database services and virtual network deployment. The tasks are simplified through database and virtual network services, and asynchronous deployment of message passing and information sharing tasks is realized.

3.1.2. Storage Service Deployment. In order to prevent data loss and data storage capacity problems, the durability and scalability of data can be improved through the SWIFT cloud storage service component. Even in a small deployment environment, high data persistence can be maintained in the Swift cloud storage service to prevent data loss. At the same time, Swift cloud storage can increase the storage capacity, improve the overall performance of the system, and increase the user experience. In the deployment process, in order to increase the security and stability of the system, when the geographical location is different, according to different needs, different storage nodes are allocated to different areas; even if there are multiple storage nodes, you can also ensure the acquisition of effective data.

In a fast access system, a sudden network outage may result in data, but delayed update operation will be added to the update list again. This makes up for the previous operation and allows the system to effectively pass the operation instructions.

3.1.3. Network Service Creation. OpenStack provides virtual network by Neutron, which can realize the communication between virtual machines and physical machines and provide virtual switching services and DHCP services for virtual machines. In order to enable virtual machines to access the external network and provide NAT services, it can be called by other components of OpenStack during the service process. The establishment process of the node-side network service is as follows: create Neutron service, create network service API Endpoint, install the provider network component, edit relevant files, make configuration modification, configure Linux bridge agent, configure DHCP agent, configure metadata agent, create file link, synchronize database, and restart computing service. The process of creating a compute node-side network service is as follows: install the compute node-side network component, make configuration changes, configure the Linux bridge agent, restart the compute service, and start the Linux bridge agent.

The sports big data information resources are connected to the data network, and the big data of sports information resources are turned into virtual data in cloud deployment, so as to truly realize the sports big data network system. In the network connection, both external networks and public networks can be accessed, and sports big data information resources can be shared anytime and anywhere to achieve undifferentiated network access, which is more conducive to data transmission and sharing.

3.2. OpenStack Service Inspection and Troubleshooting

3.2.1. Service Inspection. Invoke the Mova-Manage tool to verify that the OpenStack computing service is running properly, understand the state of the environment through different parameters, and detect the different components of the computing service.

Check the OpenStack web service. On the control node, execute the netstat-anlp l grep 9696 command to check whether the network Server API service is running on TCP port 9696.

Check the OpenStack object storage service. Swift, the OpenStack object storage service, has some built-in tools that check the health of the service and execute commands on the Swift node.

3.2.2. Troubleshooting Storage Services. As a storage construction system with high reliability, Swift is able to deal with most of the failure problems in the system. The client test of Swift failure operation is eliminated, and only the log operation is considered. First, there may be identity authentication problems, which are mainly caused by certificate errors in user or system configuration. A Swift system that supports OpenStack's authentication service requires you to manually set the authentication steps. Second, if a disk in the Swift storage environment fails, verify that the disk is unmounted, avoid Swift writing data, replace the disk, and readjust commands. Third, if the server is not running for a few hours, Swift will still work, and if it lasts longer, the server will need to be removed from the command.

4. Big Data Analysis of Sports Industry

4.1. Economic Development Status of Sports Industry

4.1.1. The Current Situation of Sports Industry Development. In order to study the impact of the sports industry on economic development, the sports industry accounting classification adopts the Sports Industry Statistical Classification (2019). Based on the Classification of Industries in National Economy (GB/T 4754-2017), the added value and proportion of 11 sports industry categories are sorted out and counted in the table. It can be seen from the data that in the category of sports industry from 2015 to 2019, the manufacturing of sporting goods and related products has the largest added value and proportion, and the second is the agency of sporting goods and related products. At the same time, the manufacturing of sporting goods and related products and the agency of sporting goods and related products account for the highest proportion of 78.6% in that year. The added value of the sports industry reached 431.79 billion yuan. Specific data are shown in Tables 1 and 2 and Figure 2.

Sports venue services ranked third from 2015 to 2017, with an added value of 45.81 billion yuan in 2015, accounting for 8.3%. In 2016, the added value was 56.76 billion yuan, accounting for 8.8%. In 2017, the added value was 67.82 billion yuan, accounting for 8.7 percent. Rounding out the top three in 2018–2019 was physical education and training, with an added value of 142.5 billion yuan in 2018, accounting for 14.1 percent. In 2019, the added value was 152.49 billion yuan, accounting for 13.6 percent. In 2018 and 2019, the added value of sports venue services was slightly lower than that of physical education and training. In 2018, the added value of sports venue services was 85.5 billion yuan, accounting for 8.5%, and in 2019, the added value of sports venue services was 101.22 billion yuan, accounting for 9.0%.

In 2018, the added value of other sports services was 61.48 billion yuan, accounting for 6.1% of the total. In 2019, the added value of sports, fitness, and leisure activities is 83.19 billion yuan, accounting for 7.4%, while the added value of other sports services is 70.07 billion yuan, accounting for 6.3%, which is slightly lower than that of sports, fitness, and leisure activities. The total added value of the sports industry was 549.44 billion yuan in 2015, 647.48 billion yuan in 2016, 781.14 billion yuan in 2017, 1,007.8 billion yuan in 2018, and 1,124.81 billion yuan in 2019. The added value of sports industry is increasing year by year.

4.1.2. The Current Economic Situation Related to the Development of Sports Industry. Indicators related to the development of the sports industry include the GDP, the number of employed people, the total investment in fixed assets of the whole society, and the per capita GDP. From 2015 to 2019, China's GDP and per capita GDP showed an upward trend year by year, reaching 67,6708 trillion yuan in 2015, 74,4127 trillion yuan in 2016, 8,27122 trillion yuan in 2017, and 9,0309 trillion yuan in 2018. China's GDP in 2019 is 9.9086.5 trillion yuan. In 2015, the per capita GDP was 4.9355.1 trillion yuan. From 2015 to 2019, the per capita GDP showed an upward trend year by year, reaching 7.089.2 trillion yuan in 2019. From 2015 to 2018, the investment in fixed assets of the whole society showed an upward trend. In 2019, the investment in fixed assets of the whole society declined slightly. In 2015, the investment in fixed assets of the whole society was 5.62 trillion yuan, in 2016, it was 6.06466 trillion yuan, and in 2017, it was 6.41238 trillion yuan. In 2018, the total fixed asset investment reached the highest value of 6.4567.5 trillion yuan, and in 2019, the total fixed asset investment dropped to 5.6087.4 trillion yuan. There was no significant change in the number of employed persons from 2015 to 2019. In 2017, the highest number of employed persons was 776.4 million. In 2015 to 2017, the number of employed persons increased slightly, and in 2018 to 2019, the number of employed persons decreased slightly. Specific data are shown in Table 3 and Figure 3.

4.2. Influence of Sports Industry on Economic Development

4.2.1. The Contribution of Sports Industry to Economic Development. The contribution rate of sports industry reflects the contribution of sports industry development to economic development. Can be seen from the collating of data, the "sports fitness and leisure activities, stadium services, physical education and training, sporting goods and related products agents, other sports services, sporting goods and related products manufacturing" promote larger contribution to the development of the economy, the following only for more than six industry has contributed to a detailed description. The detailed content of the contribution rate of sports industry is shown in Table 4.

In the category of sports industry, the top three contribution rates to economic development are sports goods and related products manufacturing, sports goods and related products agent, and physical education and training. Overall, the contribution rate of the manufacturing of sporting goods and related products showed a downward trend from 2015 to 2019, but the contribution rate increased slightly in 2018. The contribution rate of the manufacturing of sporting goods and related products was the highest 6.85% in 2015, followed by the contribution rate of 4.64% in 2018 and 4.25% in 2016. From 2015 to 2019, the contribution rate of sports goods and related products agency showed a downward trend. The contribution rate of sports goods and related products agency in 2015 was the highest 3.88%, and the difference between the contribution rate of sports goods and related products agency in 2016 to 2018 was small, and

Table 1: Economic added value of sports industry.

	Classification of sports industry	2019	2018	2017	2016	2015
1	Sports management activities	451.9	390	262.6	143.8	115
2	Sports competition performance activities	122.3	103	91.2	65.5	52.6
3	Sports fitness and leisure activities	831.9	477	254.9	172.9	129.4
4	Stadium services	1012.2	855	678.2	567.6	458.1
5	Sports intermediary service	117.8	106	24.6	17.8	14.0
6	Physical education and training	1524.9	1425	266.5	230.6	191.8
7	Sports media and information services	285.1	230	57.7	44.1	40.8
8	Sporting goods and related products agent	2562	2328	2615.8	2138.7	1562.4
9	Other sports services	707	614.8	197.2	179.7	139.6
10	Sporting goods and related products manufacturing	3421	3396.3	3264.6	2863.9	2755.5
11	Construction of sports facilities	211.9	151.2	97.8	50.3	35.3
	Total	11248.1	10078	7811.4	6474.8	5494.4

Source: National Bureau of Statistics (National Bureau of Statistics of General Administration of Sport of China Joint Announcement on the Data of Scale and Value Added of National Sports Industry from 2015 to 2019); unit: 100 million yuan.

Table 2: Economic added value ratio of sports industry.

	Classification of sports industry	2019	2018	2017	2016	2015
1	Sports management activities	4.0	3.9	3.4	2.2	2.1
2	Sports competition performance activities	1.1	1.0	1.2	1.0	1.0
3	Sports fitness and leisure activities	7.4	4.7	3.3	2.7	2.4
4	Stadium services	9.0	8.5	8.7	8.8	8.3
5	Sports intermediary service	1.0	1.1	0.3	0.3	0.3
6	Physical education and training	13.6	14.1	3.4	3.6	3.5
7	Sports media and information services	2.5	2.3	0.7	0.7	0.7
8	Sporting goods and related products agent	22.8	23.1	33.5	33.0	28.4
9	Other sports services	6.3	6.1	2.5	2.8	2.5
10	Sporting goods and related products manufacturing	30.4	33.7	41.8	44.2	50.2
11	Construction of sports facilities	1.9	1.5	1.3	0.8	0.6
	Total	100.0	100.0	100.0	100.0	100.0

Unit: %.

the lowest contribution rate in 2019 was 2.83%. The contribution rate of physical education and training was 1.95 percent in 2018 and 1.68 percent in 2019. There was a small difference in the contribution rate of physical education and training from 2015 to 2017. The lowest contribution rate of physical education and training in 2017 was 0.32 percent. See Figure 4 for details.

Stadium services, sports fitness and leisure activities, and other sports services also contribute significantly to economic development. From the perspective of stadium services, the contribution rate of stadium services in 2018 was the highest 1.17%, 2015 was 1.14%, 2019 was 1.12%, and 2016 and 2017 were 0.84% and 0.82%, respectively. From the perspective of physical fitness and leisure activities, the contribution rate of physical fitness and leisure activities showed an increasing trend from 2016 to 2019. The lowest contribution rate of physical fitness and leisure activities was 0.26% in 2016, the highest was 0.92% in 2019, and the contribution rate of physical fitness and leisure activities in 2015 was 0.32%. From the perspective of other sports services, the contribution rate of other sports services showed a downward trend from 2015 to 2017. In 2018, the contribution rate of other sports services rose to the highest value of 0.84%. In 2019, the contribution rate of other sports services decreased slightly compared with 2018, with 0.78% in 2019 and 0.24% in 2017. See Figure 5 for details.

4.2.2. The Driving Effect of Sports Industry on Economic Development. The driving rate of sports industry reflects the driving effect of sports industry on economic development. From the data, it can be seen that sports products and related products manufacturing, sports goods and related products agency, physical education and training, and sports venue services have a high pull rate index. The following is a specific description of the pull rate of the above four industries. The detailed content of sports industry pulling rate is shown in Table 5.

The manufacturing of sports products and related products has the most obvious driving effect on economic development. From 2015 to 2017, the driving effect of the manufacturing of sports products and related products is on the rise. From 2017 to 2019, the driving effect of the manufacturing of sports products and related products is decreasing year by year, but the overall change range is small. The highest pull rate was 0.432% in 2017, and the lowest was 0.378% in 2019. The driving rate of sports goods and related products agent increased year by year from 2015 to 2017, the

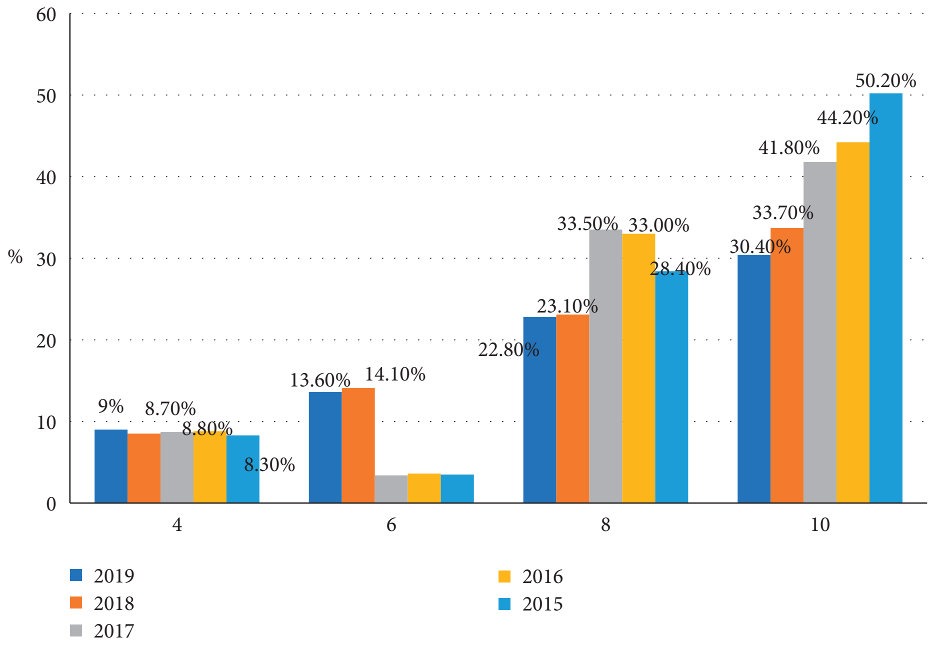


Figure 2: Economic added value of sports industry.

Table 3: National economic development data.

	Economic indicators	2019	2018	2017	2016	2015
1	Gross domestic product	990865	900309	827122	744127	676708
2	Number of employed persons	77471	77586	77640	77603	77451
3	Investment in fixed assets throughout the country	560874	645675	641238	606466	562000
4	Per capita GDP	70892	64644	59660	53980	49351

Source: National Bureau of Statistics (2015–2019 Statistical Bulletin on National Economic and Social Development); unit: 100 million yuan, ten thousand people.

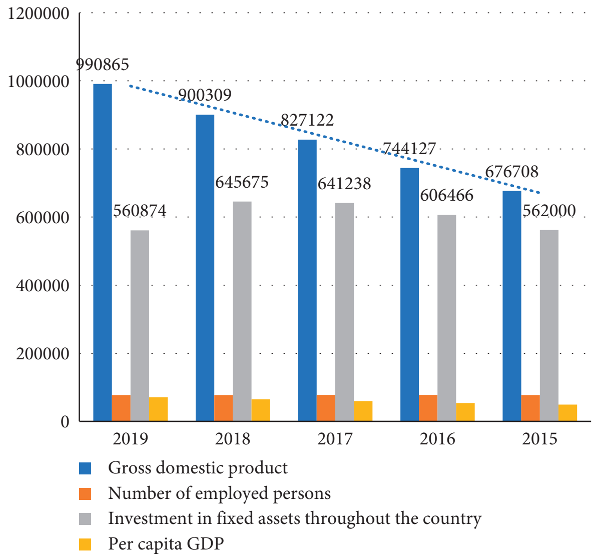


Figure 3: National economic development data.

driving rate reached the maximum of 0.347% in 2017 and the lowest of 0.233% in 2015. The driving rate of sports goods and related products agent was 0.283% in 2018 and 2019, and the driving rate was reduced compared with that of 2018 and 2017. From 2015 to 2019, the pull rate of physical education and training showed an upward trend, while from 2015 to

Table 4: Contribution rate of sports industry.

	Classification of sports industry	SR$_t$ (%)				
		2019	2018	2017	2016	2015
1	Sports management activities	0.50	0.53	0.32	0.21	0.28
2	Sports competition performance activities	0.14	0.14	0.11	0.10	0.13
3	Sports fitness and leisure activities	0.92	0.65	0.31	0.26	0.32
4	Stadium services	1.12	1.17	0.82	0.84	1.14
5	Sports intermediary service	0.13	0.14	0.03	0.03	0.03
6	Physical education and training	1.68	1.95	0.32	0.34	0.48
7	Sports media and information services	0.31	0.31	0.07	0.07	0.10
8	Sporting goods and related products agent	2.83	3.18	3.15	3.17	3.88
9	Other sports services	0.78	0.84	0.24	0.27	0.35
10	Sporting goods and related products manufacturing	3.78	4.64	3.93	4.25	6.85
11	Construction of sports facilities	0.23	0.21	0.12	0.07	0.09
	Total	12.42	13.77	9.41	9.60	13.65

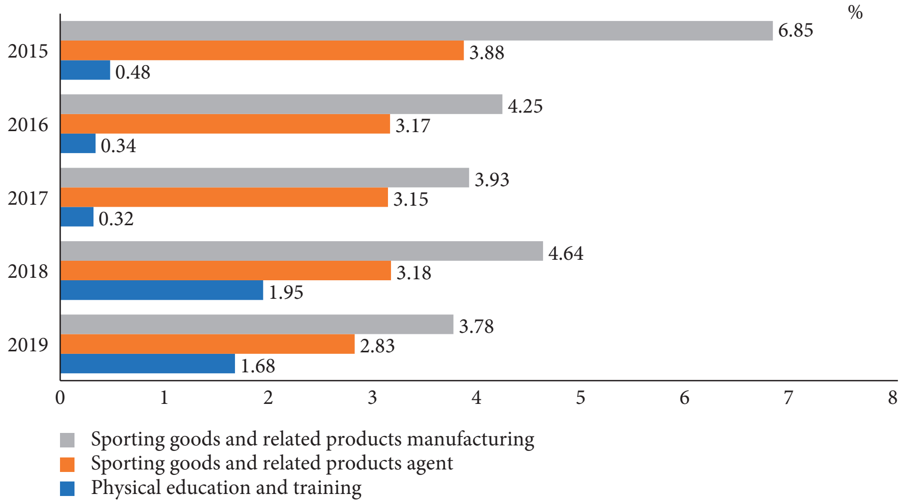


Figure 4: Contribution rate of sports industry.

2017, the rise rate was slow. The pull rate increased significantly in 2018 and 2019, with the lowest value of 0.029% in 2015 and the highest value of 0.168% in 2019. The pull rate of stadium services increased year by year from 2015 to 2019. The lowest pull rate of stadium services in 2015 was 0.068%, and the highest one in 2019 was 0.112%. See Figure 6 for details.

4.3. Contribution and Driving Effect of Sports Industry to Economy. Figure 7 shows the influence relationship between the contribution rate and pulling rate of sports industry and GDP growth rate. It can be seen from the figure that the contribution rate of sports industry to economic development was 13.65% in 2015, 9.6% in 2016, 9.41% in 2017, and the contribution rate increased significantly in 2018. The overall contribution rate of the sports industry reached 13.77%, and the contribution rate of the sports industry in 2019 was 12.42%. From the perspective of the driving effect of sports industry on economic development, from 2015 to 2019, the driving effect of sports industry on economic development is not very obvious, with the driving rate of 0.819% in 2015, 0.96% in 2016, and 1.035% in 2017. The pull rate in 2018 is 1.226%, and the pull rate in 2019 is 1.242%. The growth rate of GDP began to increase year by year in 2015 and reached a maximum of 11.15% in 2017. In 2018, the growth rate of GDP decreased, and in 2019, the growth rate of GDP rose to 10.05%. On the whole, the average contribution rate of the sports industry is 11.77%, and the average pull rate of the sports industry is 1.056%. The values of the contribution rate and pull rate of the sports industry are both positive. The sports industry has obvious contribution and pull effect on the economic development, so the development of the sports industry promotes the development of the social economy. It has a positive influence on economic development and thus promotes economic development. See Figure 7 for details.

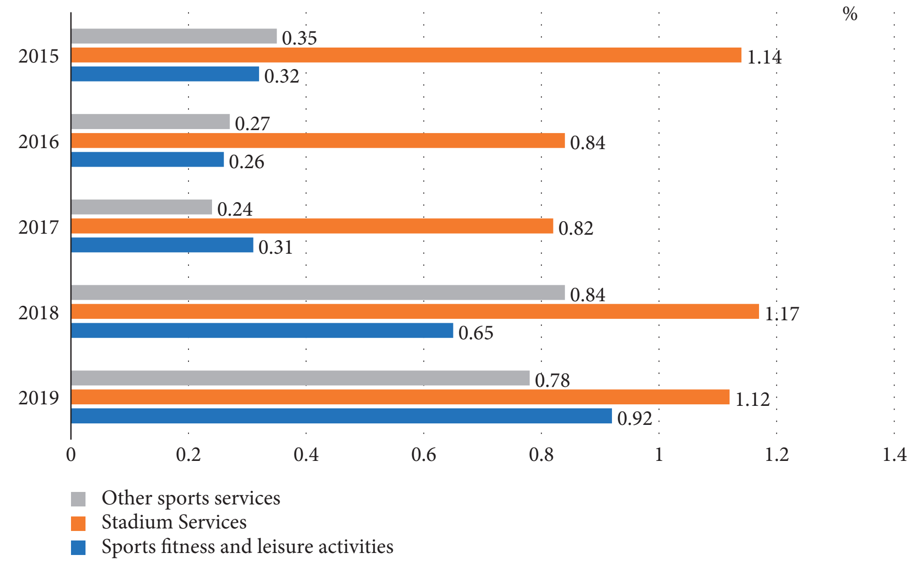


Figure 5: Contribution rate of sports industry.

Table 5: The pull rate of sports industry.

	Classification of sports industry	SP_t (%)				
		2019	2018	2017	2016	2015
1	Sports management activities	0.050	0.047	0.035	0.021	0.017
2	Sports competition performance activities	0.014	0.012	0.012	0.010	0.008
3	Sports fitness and leisure activities	0.092	0.058	0.034	0.026	0.019
4	Stadium services	0.112	0.104	0.090	0.084	0.068
5	Sports intermediary service	0.013	0.012	0.003	0.003	0.002
6	Physical education and training	0.168	0.174	0.035	0.034	0.029
7	Sports media and information services	0.031	0.028	0.008	0.007	0.006
8	Sporting goods and related products agent	0.283	0.283	0.347	0.317	0.233
9	Other sports services	0.078	0.075	0.026	0.027	0.012
10	Sporting goods and related products manufacturing	0.378	0.413	0.432	0.425	0.411
11	Construction of sports facilities	0.023	0.019	0.013	0.007	0.005
	Total	1.242	1.226	1.035	0.960	0.819

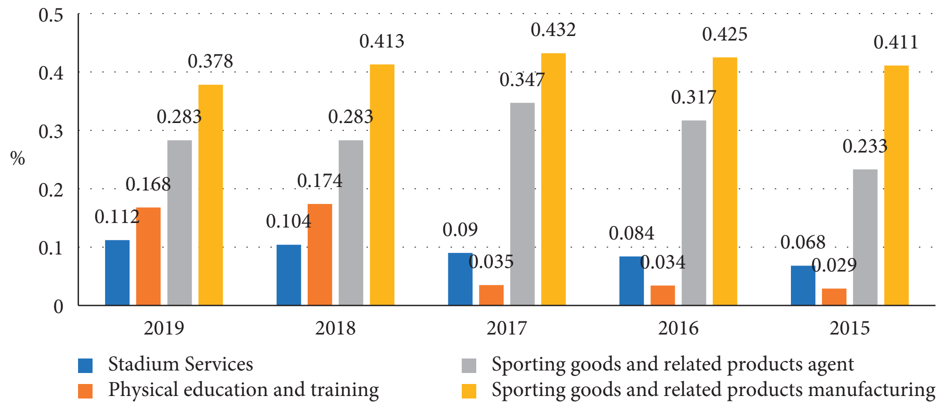


Figure 6: The pull rate of sports industry.

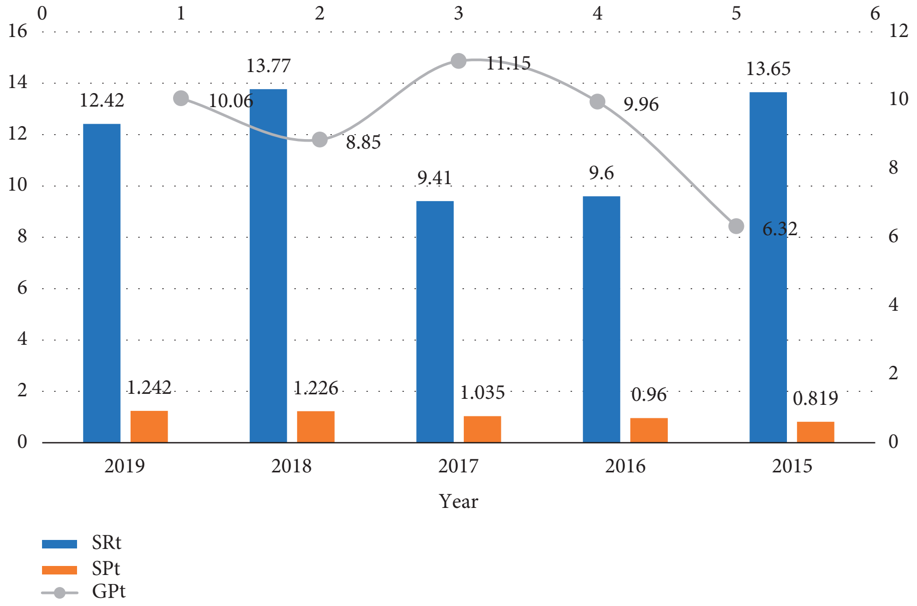


Figure 7: Contribution rate and pull rate of sports industry.

5. Conclusions

Statistical analysis according to the scale of the sports industry, sports products and related products manufacturing industry, sports products and related products agents, and physical education and training industry scale continues to expand, but has been to the direction of sports services, sports industry structure also tends to rationalization. In the development process of big data, sports industry is integrated into big data analysis to make data acquisition more intuitive and data analysis more comprehensive, grasp the development trend of sports industry economy, and predict the future development trend of sports industry, so as to promote economic and social development.

Big data are inseparable from us in modern life. In the cloud age, big data also involve various fields. Through the cloud computing platform technology, a large amount of data can be effectively acquired, stored, managed, and analyzed. The application of big data to the sports industry has the following two impacts on the sports industry. First, through the analysis of big data of the sports industry, we can intuitively understand the impact of the sports industry on economic development, which is of profound significance for promoting economic and social development. Secondly, through the design of the information resource sharing module of the cloud platform, the obstacles to the acquisition of big data resources are greatly reduced, making the acquisition of data resources and the analysis of big data more convenient.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] Smart government big data cloud platform construction scheme [EB/OL], 2018.

[2] *The State Council Issued the Action Program to Promote the Development of Big Data [EB/OL]*, Xinhua News Agency, Beijing, China, 2015.

[3] *The State Council "Several Opinions on Accelerating the Development of Sports Industry and Promoting Sports Consumption"*, State Council, Zhongnanhai, Beijing, China, 2014.

[4] *Opinions of General Office of the State Council on Strengthening the Construction of National Fitness Facilities and Developing Mass Sports [EB/OL] No*, State Council, Zhongnanhai, Beijing, China, 2020.

[5] G. C. Liu and D. Wu, *Journal of Jilin University (Engineering and Technology Edition)*, vol. 38, no. 6, pp. 709–713, 2020.

[6] H. B. Yi, R. N. Chi, Z. Nie, and X. Huang, "Teaching reform and practice of "OpenStack cloud platform deployment and management" course," *Journal of Shenzhen Vocational and Technical College*, vol. 19, no. 5, pp. 66–71, 2020.

[7] Z. J. Tang and C. Huang, "Research on high availability of cloud platform based on OpenStack," *Information & Computer (Theoretical Edition)*, vol. 2019, no. 6, pp. 21-22, 2019.

[8] Y. Zhong, "Application research of OpenStack on cloud experimental platform construction," *Electronic Design Engineering*, vol. 27, no. 8, pp. 24–28 + 33, 2019.

[9] Y. Wang and Q. L. Wu, "Construction of the cloud platform for innovation and entrepreneurship of local universities based on OpenStack," *Experimental Technology and Management*, vol. 36, no. 4, pp. 257–260+277, 2019.

[10] T. T. Wen, H. Z. Li, and S. F. Li, “Design and implementation of virtual network management platform based on OpenStack,” *Electronic Manufacturing*, vol. 2019, no. 10, pp. 47–49, 2019.

[11] Z. L. Xin, D. M. Liang, R. Ma, M. H. Zou, and J. P. Zhang, “Glance service based on OpenStack platform,” *Henan Science and Technology*, vol. 63, no. 23, pp. 33–35, 2020.

[12] J. Yao and N. Yan, “Design and implementation of OpenStack high availability distributed storage scheme,” *Computer Technology and Development*, vol. 29, no. 2, pp. 35–38, 2019.

[13] Y. Y. Zhu, *Cloud Computing Architecture and Application*, pp. 7-8, South China University of Technology Press, Guangzhou, China, 2017.

[14] X. L. Tian, Z. Yuan, and N. Zhang, “Security research and improvement of OpenStack authentication backend,” *Journal of Beijing Institute of Electronic Science and Technology*, vol. 24, no. 4, pp. 26–31, 2016.

[15] J. L. Mao, “OpenStack Nova service,” *Computer Networks*, vol. 44, no. 3, pp. 60–63, 2018.

[16] Y. Li, Y. Dong, J. Zhang, D. Liang, and J. Zhang, “Neutron services based on OpenStack platform,” *Computer Programming Skills and Maintenance*, vol. 26, no. 10, pp. 12-13+46, 2020.

[17] Z. Y. Gong, “Implementation of OpenStack Horizon automated testing based on selenium,” *Foreign Electronic Measurement Technology*, vol. 36, no. 5, pp. 45–49, 2017.

[18] C. Cheng, Z. X. Zhu, and X. J. Liang, “Research and application of OpenStack Horizon framework,” *Electronic Science and Technology*, vol. 29, no. 4, pp. 84–87, 2016.

[19] Y. Song, D. Y. Zhou, and W. C. Shi, “A method to enhance the security function of OpenStack Swift cloud storage system,” *Journal of Jilin University (Engineering and Technology Edition)*, vol. 51, no. 1, pp. 314–322, 2021.

[20] H. D. Huang, “Application of cloud computing in virtual training teaching platform for network security,” *Electronic Technology*, vol. 49, no. 12, pp. 24-25, 2020.

[21] X. L. Zhang, C. Y. Wen, and Z. Zhang, “Research and implementation of high availability storage scheme based on OpenStack Swift,” *Journal of Chengdu University of Information Technology*, vol. 34, no. 1, pp. 44–48, 2019.

[22] J. T. Yin and T. T. Lu, “Research on mobile application construction of OpenStack cloud computing platform,” *Journal of Jining Normal University*, vol. 42, no. 05, pp. 57–61, 2020.

[23] L. Y. Song, “Research and implementation of university cloud data center based on OpenStack,” *Journal of Chifeng University (Natural Science Edition)*, vol. 32, no. 15, pp. 33–35, 2016.

[24] M. L. Wu, T. H. Ren, and Y. B. Li, “Application and research of resource management technology of private cloud platform based on OpenStack,” *Industrial Technology Innovation*, vol. 2, no. 3, pp. 334–341, 2015.

[25] X. F. Yang, “Research and implementation of teaching private cloud based on OpenStack open source cloud platform,” *Information and Computers*, vol. 32, no. 4, pp. 188–190, 2020.

[26] H. Li, H. Zhou, H. Zhang, and B. Feng, “EmuStack: an OpenStack-based DTN network emulation platform,” in *Proceedings of 2016 International Conference on Networking and Network Applications (NaNA)*, vol. 35, no. 9, pp. 24–26, Hakodate City, Hokkaido, Japan, July 2016.

[27] X. M. Zhang and J. Wang, *Industrial Economics*, p. 121, Electronic Science and Technology University Press, Xi ’an, China, 2017.

[28] Y. Zhang and J. Y. Xiong, “On the calculation of “contribution rate” and “pull rate”,” *China Statistics*, vol. 64, no. 9, pp. 38–40, 2017.

[29] Y. Li, P. Y. Zhao, and M. Z. Xi, “An empirical study on the driving effect of tourism on economic growth in Xinzhou City: based on the three indicators of tourism dependence, contribution rate and driving rate,” *Journal of Mianyang Normal University*, vol. 34, no. 6, pp. 44–48, 2015.

[30] *National Economy Industry Classification Notes》(revised According to Amendment No.1)*, National Bureau of Statistics, China, 2017.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 9960822, 10 pages
https://doi.org/10.1155/2021/9960822

Research Article

Malicious Encryption Traffic Detection Based on NLP

Hao Yang,[1] **Qin He**,[1] **Zhenyan Liu**,[1] **and Qian Zhang**[2]

[1] *School of Computing Science, Chengdu University of Information Technology, Chengdu 610225, China*
[2] *School of Computer Science, University of Nottingham Jubilee Campus, Nottingham NG8 1BB, UK*

Correspondence should be addressed to Hao Yang; vhyang@foxmail.com

Received 21 March 2021; Revised 2 June 2021; Accepted 26 July 2021; Published 3 August 2021

Academic Editor: Yuan Tian

The development of Internet and network applications has brought the development of encrypted communication technology. But on this basis, malicious traffic also uses encryption to avoid traditional security protection and detection. Traditional security protection and detection methods cannot accurately detect encrypted malicious traffic. In recent years, the rise of artificial intelligence allows us to use machine learning and deep learning methods to detect encrypted malicious traffic without decryption, and the detection results are very accurate. At present, the research on malicious encrypted traffic detection mainly focuses on the characteristics' analysis of encrypted traffic and the selection of machine learning algorithms. In this paper, a method combining natural language processing and machine learning is proposed; that is, a detection method based on TF-IDF is proposed to build a detection model. In the process of data preprocessing, this method introduces the natural language processing method, namely, the TF-IDF model, to extract data information, obtain the importance of keywords, and then reconstruct the characteristics of data. The detection method based on the TF-IDF model does not need to analyze each field of the data set. Compared with the general machine learning data preprocessing method, that is, data encoding processing, the experimental results show that using natural language processing technology to preprocess data can effectively improve the accuracy of detection. Gradient boosting classifier, random forest classifier, AdaBoost classifier, and the ensemble model based on these three classifiers are, respectively, used in the construction of the later models. At the same time, CNN neural network in deep learning is also used for training, and CNN can effectively extract data information. Under the condition that the input data of the classifier and neural network are consistent, through the comparison and analysis of various methods, the accuracy of the one-dimensional convolutional network based on CNN is slightly higher than that of the classifier based on machine learning.

1. Introduction

Related principles and methods of plaintext transmission put forward higher requirements for the security of the service system, and it is an inevitable trend for Internet applications to move towards the era of comprehensive encryption [1, 2]. In recent years, the rapid increase in encrypted communications has changed the threat patterns, and many traditional methods based on conventional rules are no longer as effective as they once were. As more and more enterprises go digital, a large number of services and applications are adopting encryption as their primary means of information protection. According to NetMarketShare, the percentage of encrypted web traffic was already over 90% in October 2019. However, in the case that encrypted access can guarantee communication security, the vast majority of network devices is powerless against network attacks, malware, and other malicious encrypted traffic. A large number of malware, ransomware, proxies, remote control tools, etc., use encryption methods to avoid security protection and detection. Common security products will release unrecognized and undetectable traffic, such as Trojan, ransomware, downloaders [3], and other types of malicious software or code. In order to avoid security products and human detection, encryption is often used to disguise or hide attack behavior. Samples of malicious families that use rebound technology to bypass security devices also frequently switch back domain names and IPs and encrypt communications. The attack chain is usually divided into several steps, such as information collection, intrusion control, achievement expansion, and battlefield cleaning. The stage of the attacker can be clearly understood through staged analysis and display of events in the traffic. Therefore, it is imperative to encrypt malicious traffic detection.

In recent years, the detection of encrypted malicious traffic has been the focus of attention in the field of network security. At present, there are two mainstream attack detection methods: detection after decryption and detection without decryption. The industry gateway devices mainly use the method of decrypting traffic to detect the attack behavior, but this solution method will consume a lot of resources, the cost is very high, also it violates the original intention of encryption, and the decryption process will be strictly limited by the laws and regulations related to privacy protection. Considering the protection of user privacy, detection methods that do not require decryption of traffic are becoming an industry of concern. Researchers are usually only allowed to observe network encrypted traffic (port 443), without decryption, by using existing data resources and standard encrypted traffic for analysis [4]. For example, the method based on statistical learning can be detected without decryption, but its detection accuracy is not high, and it cannot guarantee the correct detection of most malicious traffic. With the development of machine learning and deep learning in recent years, detection methods based on artificial intelligence have become active [5]. The methods based on machine learning or deep learning can achieve high accuracy without decryption and through some processing means.

Through many verifications, artificial intelligence used in encryption traffic security detection is a very good auxiliary means. Shengbang Security Sustainable Threat Detection and Traceability System (RayEye), based on an artificial intelligence engine, can analyze the full network traffic in real time. Combined with threat intelligence data and network behavior analysis technology, it can detect suspicious behaviors in depth, help to clearly grasp the attack chain stage and success probability of the attacker, and provide customers with malicious encryption attack traffic detection solutions.

Although the detection methods based on machine learning and deep learning do not need to decrypt the encrypted data, they still need to analyze the traffic data fields and extract the features. Based on traditional detection methods that cannot detect encryption flow and machine learning methods that need to expend energy problems such as feature extraction, this paper proposes machine learning, deep learning, and natural language processing to detect malicious traffic encryption methods, the combination uses the text classification[6] method to represent encrypted traffic, so it does not need to decrypt the data, does not need to care about the meaning of the field of the traffic data itself, and does not lose the data information of the encrypted traffic. This detection method not only is applicable to encrypted malicious traffic detection but also can be used for other related detections, such as malicious code detection. It has strong generalization and high accuracy. In the later model improvement, it is not necessary to rigidly extract the information of encrypted traffic data.

2. Materials and Methods

This paper uses the TF-IDF model to calculate the TF-IDF value of each keyword in the traffic packet and does not carry out segmentation operations on the traffic packet [7]. TF-IDF model converts qualitative data in traffic packets into quantitative data and then carries out training and detection through various classifiers. The following table shows the general situation of each classifier used in this paper.

2.1. TF-IDF. TF-IDF (Term Frequency-Inverse Document Frequency) is a normally used weighting technique for message retrieval and keyword extraction [8]. The TF-IDF model is used to calculate the TF value and IDF value of a word. If a term appears frequently in a document of a class, it is a good representation of the text of that class. Such terms should be given high weight and selected as feature terms for that class of text to distinguish it from other class documents. But only using word frequency cannot effectively filter modal words or some meaningless words. IF-IDF model introduces IDF value on the basis of word frequency. TF-IDF is a statistical approach, which is used to calculate the significance of a word to a document in a document set or a corpus. The importance of a word increases in a direct proportion with the number of times it arises in the document but decreases in an inverse proportion with the frequency of its occurrence in the corpus [9]. The main idea of TF-IDF is that if a certain word or phrase appears in one article with a high TF frequency and rarely appears in other articles, that is, IDF is low, then it is considered that this word or phrase has a good ability to distinguish categories and is suitable for classification.

TF represents word frequency, that is, the frequency of keywords appearing in the text. Vectorization of text data is to take the occurrence frequency of each word in the text as the characteristic of the text [10], and IDF value is used to correct the word frequency vector represented by TF value only.

IDF stands for reverse file frequency: its size is inversely proportional to the common degree of a word; that is, if a word is included in multiple files of a corpus, then the IDF value of the word is small. The IDF is a measure of the general importance of a word. The IDF of a particular term can be obtained by dividing the total number of files by the number of files containing the term and then taking the logarithm of the resulting quotient [11].

IDF solution formula is as follows:

$$\mathrm{IDF}(x) = \log\frac{N}{N(x)}, \tag{1}$$

where N represents the total number of Chinese texts in the corpus and $N(x)$ represents the total number of texts containing the word x in the corpus.

However, there are usually some extreme cases; for example, when a rare word does not exist in the corpus,

$N(x)$ value is 0, the above calculation formula will not be valid. Therefore, the IDF calculation formula is smoothed as follows:

$$\mathrm{IDF}(x) = \log\frac{N + 1}{N(x) + 1} + 1. \tag{2}$$

Therefore, the IDF value can also be correctly calculated in the above cases. Finally, the TF value and lDF value of the word are multiplied to obtain the TF-IDF value of the word. The greater the value of the word TF-IDF, the higher the importance of the word to the article. The advantage of the TF-IDF algorithm is simple and fast, and the result is more in line with the actual situation. For traffic data, TF-IDF will not cause the loss of data information and can better extract keywords to transform data information.

2.2. Detection Method. Gradient boosting belongs to the boosting series algorithm of ensemble learning [10]. Gradient boosting boosts the combination of weak classifiers. Different from the AdaBoost algorithm, gradient boosting selects the direction of gradient descent in iteration to ensure the best final result [12, 13]. Gradient boosting generates a number of weak learners, each of which takes the negative gradient as the error measurement index of the previous round of basic learners. The goal is to fit the negative gradient of the loss function of the previous cumulative model so that the cumulative model loss after adding the weak learner can be reduced in the direction of the negative gradient. Gradient boosting, compared with AdaBoost, can use any loss function (as long as the loss function is continuously differentiable), so some relatively robust loss functions can be applied, making the model more robust in noise resistance.

Random forest is a more advanced algorithm based on a decision tree (default CART tree), which also belongs to the category of ensemble learning. Random forest is composed of multiple decision trees, and decision trees do not influence each other. The random forest algorithm allows the decision tree to construct a forest randomly. After each round, each decision tree gets its own result, and the final result is determined by voting. The category with the highest number of votes is taken as the output result of the random forest [8, 14]. The introduction of randomness makes the random forest not easy to fall into overfitting. And it has good antinoise ability. Random forest can process data with high dimensions, that is, a large number of features, and it does not need to make feature selection. It has strong adaptability to data sets: it can process both discrete and continuous data, and the data sets do not need normalization.

AdaBoost algorithm belongs to boosting series of ensemble learning algorithms, which is composed of multiple weak learners and adjusts the network by giving learners different weights each time. Its adaptability lies in that the weight of the misclassified sample (the corresponding weight of the sample) of the previous weak classifier will be strengthened, and the weight of the correctly classified sample will be reduced at the same time. The sample with updated weight will be used to train the next new weak classifier again. Finally, the linear weighted sum shows that the base learner with a small error rate has a larger weight, while the base learner with a large error rate has a smaller weight [15]. In each round of training, a new weak classifier is trained with the population (sample population) to generate new sample weights and the power of the weak classifier, and the iteration continues until it reaches the predetermined error rate or the specified maximum number of iterations. AdaBoost uses exponential loss, which has a weakness that it is very sensitive to outlier points, so AdaBoost is better than gradient boosting.

Convolutional Neural Networks (CNN) [14] is a deeply structured feedforward neural network that includes convolutional computation and is mostly used in graphics processing [16]. It is one of the representative algorithms of deep learning [17, 18]. CNN usually includes data input layer, convolution calculation layer, ReLU activation layer, pooling layer, and full connection layer. CNN obtains key information mainly through continuous extraction of feature information. Compared with a machine learning algorithm, CNN can extract key information more effectively, and the number of parameters is small without careful parameter adjustment. We only need to randomly assign a weight w and a bias term b to each neuron during initialization. In the training process, these two parameters will be continuously revised to the best quality, so as to minimize the error of the model. However, this will correspondingly increase the amount of calculation, and the model training time will also increase due to the increase in the amount of calculation.

2.3. TF-IDF-Based Detection Method. Encrypted traffic messages typically contain fields such as IP address, port number, MAC address, triple handshake protocol, and various protocols. Some of these fields may directly affect the training effect of the subsequent model, while some fields may be redundant information for model training, which requires us to reserve professional network security knowledge in advance to analyze malicious traffic data, and the size, length, and field of each traffic data in the data set are different. After the encryption traffic data analysis is completed, the encrypted data are extracted by field features. The general machine learning method is to uniformly encode the input traffic data into digital form. After some feature engineering work is done on the data, the data are input to the classification model for training detection, but this method will have the problem of information loss more or less, which may make the obtained detection accuracy fail to meet the requirements. For the problem that traffic data are not easy to be processed, a detection method combining TF-IDF model in natural language processing with machine learning or deep learning is proposed. This method does not require us to consider and analyze the meaning of packets and details in specific fields, and extract relevant features of data for coding. That is to say, the TF-IDF model is used to reconstruct the data set, extract the text information of the data set, rebuild the new features, and represent the fields in the form of vectorization when processing the data set. Using the TF-IDF model to transform traffic text will not have missing information. Moreover, better results can be obtained without further data processing. Because the data are kept in the TF-IDF model to analyze the important degree of each keyword and dealing with each keyword, instead of the need for human to deal with the analysis of data, therefore, the TF-IDF-based malicious traffic detection method can not only be used for encryption detection, but also can be applied to relevant detection that requires the use of professional technology to extract data information, such as malicious code detection.

In this paper, based on the TF-IDF model, the specific process of information text extraction and feature reconstruction for encrypted traffic data set is shown in Figure 1.

As can be seen from the figure, there are altogether 3000 pieces of encrypted traffic data in the experiment, and each piece of data is different in size, length, and field, but most of them contain text information such as transmission address, handshake information, and TCP protocol. In the experiment, a total of 906,069 keywords were obtained after the text was directly extracted from the TF-IDF model and transformed into features. The TF-IDF algorithm will calculate the TF-IDF value for each keyword. After each piece of data conversion, the original text content is converted to the corresponding keyword TF-IDF value to represent. After the conversion of the source encrypted data set, the resulting data set is a sparse matrix with the size of 3000 × 906069.

At the same time, the source encrypted traffic data set is transformed into digital text representation by thermal coding processing and then input into the classification model for training detection as a comparative experiment. One-Hot Encoding, also known as one-bit effective coding, mainly uses the bit status register to encode the states. Each state has its own independent register bit, and only one is effective at any time. Unique thermal coding uses 0 and 1 to represent some parameters and N-bit status register to encode N states. One-Hot Encoding can deal with discontinuous numerical features. It also expands the features to some extent.

After the encrypted traffic data set is processed by the TF-IDF model, the new data set obtained is input to each machine learning classifier and the convolutional neural network. Each classifier and convolutional neural network were trained and tested, respectively. The overall process of malicious encryption traffic detection method and experiment is shown in Figure 2.

In this paper, gradient boosting classifier, random forest classifier, and AdaBoost classifier are adopted, respectively, for training and detection, and the experimental results show that the detection results obtained by random forest classifier are better. The random forest itself is a kind of integrated learning method, and the results are obtained by voting. Therefore, based on this idea, this paper uses the ensemble learning method to combine multiple classifiers to form the ensemble model. Ensemble learning is to combine the above multiple weak supervised models in order to obtain a better and more comprehensive strong supervised model. The underlying idea of ensemble learning is that even if a weak classifier gets a wrong prediction, other weak classifiers can correct the error back. It is a meta-algorithm that combines several machine learning techniques into a prediction model to reduce bagging and increase or improve the prediction stack. In the experiment, gradient boosting classifier, random forest classifier, and AdaBoost classifier were adopted, and XGBoost classifier was combined with ensemble training. XGBoost has the characteristics of fast speed, good effect, and large-scale data processing, and XGBoost is an integrated learning framework with high accuracy, which is an efficient implementation of the GB algorithm. The final output of the ensemble learning model also adopts a voting scheme. The experimental results show that ensemble learning is better than single classifier training.

There are two options for convolutional neural networks: one is two-dimensional convolution, and the other is one-dimensional convolution based on eigenvectors [11, 19, 20]. The two-dimensional convolution first reconstructs the eigenvectors into single-channel matrices of the same size and then carries out the two-dimensional convolution. The experiment uses a variety of schemes from classical network structure (AlexNet, etc.) to specially designed network structure. However, in the course of the experiment, the following two points were found:

(1) When a simple network structure is adopted, its convergence speed is very fast [21], and it can achieve a very good effect on the training set (the accuracy rate tends to be close to 1), but the performance effect on the test set is very poor (the accuracy rate is less than 0.5), that is, the phenomenon of overfitting. After trying the processing methods including but not limited to dimension reduction, regularization, and simplified network, they cannot get better improvement.

(2) When a complex network structure is used, the convergence speed is slower, and even there are many iterations (more than 30 times), while the accuracy rate remains unchanged. The results obtained on the training set are similar to those obtained on the test set. However, the performance effect is generally poor compared with the ensemble learning scheme of machine learning (the accuracy rate is about 0.7).

Therefore, the two-dimensional convolution scheme is finally abandoned and one-dimensional convolution is adopted.

The encrypted traffic data are converted through the TF-IDF model, and the output is a sparse matrix with the size of 3000 × 906069. For each classifier of machine learning, including the ensemble learning model, there is no need to carry out a lot of computational fittings, so most machine learning algorithms are not limited to the size of the data set, and the general memory can meet the requirements. However, for the convolutional neural network, because each layer of the CNN network needs a lot of calculation, if

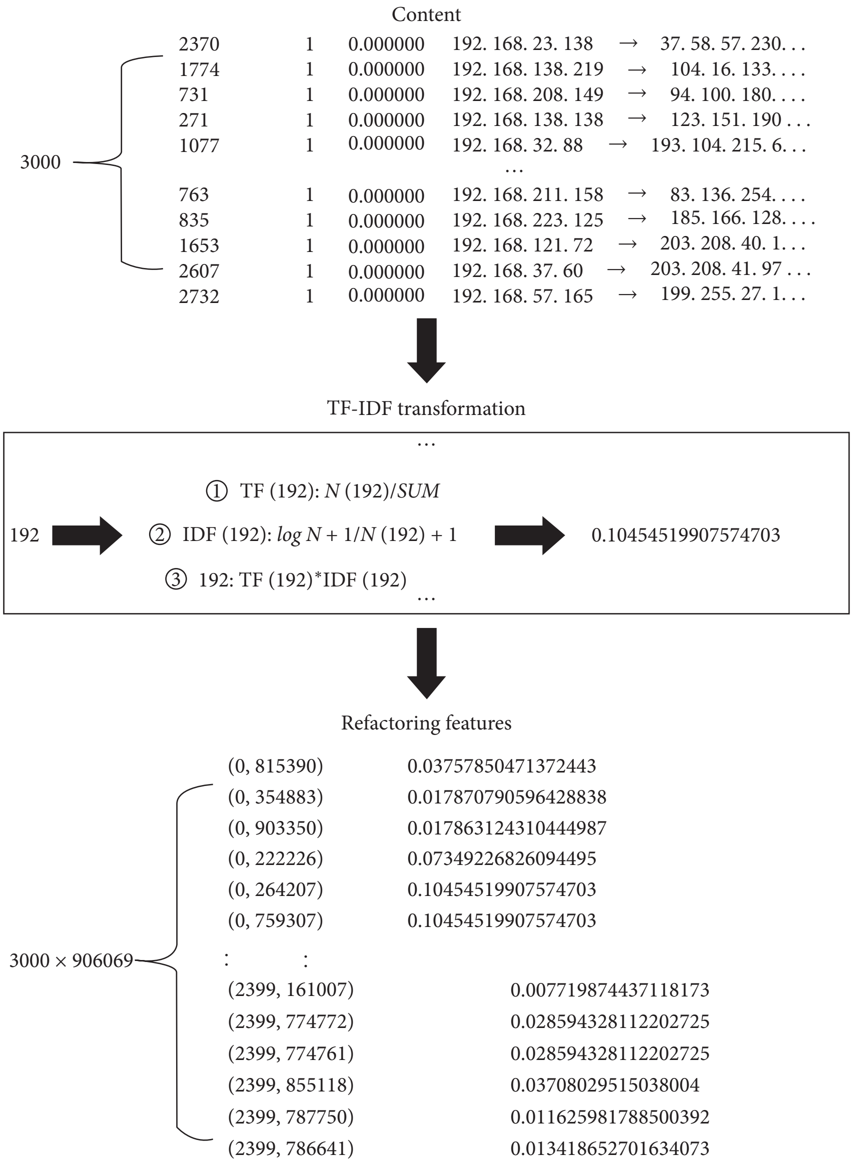


Figure 1: TF-IDF model is used to extract text information and reconstruct features from data sets.

the amount of data is too large, the general memory cannot meet the demand, and the training time will be greatly improved. Therefore, the feature dimension reduction method must be used to reduce the dimension of the data set. In this paper, Truncated SVD [22] method is used to carry out characteristic dimension reduction of data. Using Truncated SVD, the original feature matrix with size (number of texts and number of terms) is transformed into a new feature matrix with size (number of texts and number of topics). It is very suitable for data dimensionality reduction in the later stage of the TF-IDF model. The convolutional neural network structure in this paper consists of 13 layers, including the convolutional layer, the activation layer, the pooling layer, the dropout layer, the flatten layer, and the dense layer. Dropout layer is added to network results to prevent model overfitting. Add the flatten layer to convert multidimensional data to one-dimensional data. Finally, according to feature combination, the dense layer is added to classify, which greatly reduces the influence of feature position on classification.

2.4. Parameter Selection. For the ensemble learning model, it is mainly to adjust the parameters of each classifier. For the AdaBoost classifier in the ensemble model, it is mainly the decision tree classifier. Because this paper is dichotomous and the data sample is small, set max_depth to 2 and the rest to the default. Then set the maximum number of iterations of the weak learner, n_estimators, to 500. For the random forest classifier, the number of subtrees is set to 500, and the experiment shows that the effect is counterproductive when n_estimators are greater than 500. Gradient boosting classifier and XGBoost classifier are both using the default parameters. For a convolutional neural network, it is mainly about the design of iteration times and network structure. In the convolutional neural network, the convolutional layer of

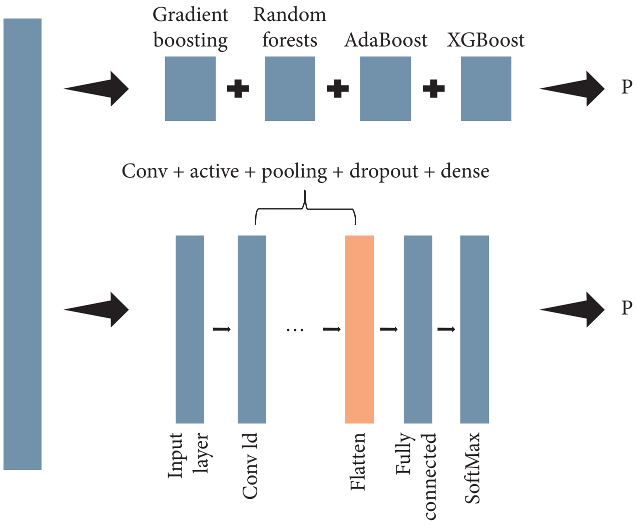


Figure 2: After the input of the training data, the results were obtained by training the data through ensemble learning and CNN, respectively.

the first few layers has a small proportion of the number of parameters, but a large proportion of the computational amount. The fully connected layer behind the network is just the opposite. Most CNN networks have this feature. Therefore, we should focus on the convolutional layer when carrying out computational acceleration optimization; when optimizing parameters and trimming weights, the focus should be on the full connection layer.

3. Results and Discussion

3.1. Data. This data set is derived from the malware and normal software collected from February to June 2020, which are operated by the sky dome sandbox of QiAnXin Technology Research Institute and filtered and generated by collecting the traffic generated. The malicious traffic defined in this data set is the encrypted traffic generated by malware (all of type exe), and the white traffic is the encrypted traffic generated by normal software (all of type exe). The traffic content is TLS/SSL packets generated by port 443. The black sample in the training set is the encrypted traffic of malware captured from February 2020 to May 2020, and the black sample in the test set is the encrypted traffic of malware captured in June 2020. All the white samples are normal software-encrypted traffic captured in 2020.

The experiment has a total of 3000 data packets, including 1500 black and white data, respectively. The experiment adopts data of 28 parts; namely, 2400 black and white data are selected as training data and 600 black and white data are selected as test data. Since the data set is a PCAP packet, the packet should be parsed first. In this experiment, Wireshark software was used to analyze the data packets. Wireshark software includes the command-line tool tshark, which can extract the desired PCAP packets by command. After analyzing PCAP data packets, data cleaning is required. Data cleaning is mainly to ensure that the collected data have a positive impact on the model. Any wrong data in the data set may have a great impact on the model construction process and the performance of the model.

When the encrypted traffic data set is large, there may be duplicate data, so we need to delete duplicate data and keep only one data. For the data set in this paper, there is basically no problem of data duplication. However, in the process of data conversion, it is inevitable that a small part of data will be wrongly copied, resulting in data duplication. This paper uses the method of matching the content of the malicious traffic field to deduplication. The parsed, deduplicated data files are then organized as DataFrame, marked with black and white labels, and shuffled out of sample order. Among them, the malicious traffic sample data are marked as 1, and the benign traffic sample data are marked as 0.

3.2. Assessment. In this paper, confusion matrix, accuracy, ROC curve, and AUC value were used to evaluate the experimental results. An obfuscation matrix is used to visually display the classification situation, and the detection results can be visually displayed for binary classification problem such as encrypted malicious traffic detection. The specific definition of the confusion matrix is shown in Table 1.

As shown in the table, TP indicates that the predicted value of the model is a benign sample, and the actual value is also a benign sample. FP indicates that the predicted value of the model is a benign sample, while the actual value is a malicious sample. FN indicates that the predicted value of the model is malicious samples, while the actual value is benign samples. TN indicates that the predicted value of the model is malicious traffic and the actual value is also malicious traffic. Accuracy, ROC curves, and AUC values are calculated on the basis of the confusion matrix.

On the basis of the confusion matrix, it can be extended to accuracy, which is our most common evaluation index, and it is easy to understand, namely, the number of samples divided by all the samples. Generally speaking, the higher the accuracy, the better the classifier. The accuracy calculation formula is as follows:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{P} + \mathrm{N}}, \tag{3}$$

ACC represents the proportion of all correctly judged results of the model to the total observed values, where P + N represents the total number of use cases. TP and TN are the number of correctly classified samples.

Receiver Operating Characteristic (ROC) is a curve drawn on a two-dimensional plane, whose abscissa is the false positive rate (FPR) and the ordinate is the true positive rate (TPR). For a classifier, we can get a TPR and FPR point pair based on its performance on the test sample. Thus, this classifier can be mapped to a point on the ROC plane. By adjusting the threshold used by the classifier, we can get a curve that goes through (0, 0) and (1, 1), which is the ROC curve of the classifier. The calculation formulas of abscissa and ordinate are as follows:

Table 1: General situation of each classifier.

Data processing method	Classifier	Characteristics
TF-IDF	Gradient boosting	Slight time, a bit poor result
	Random forest	
	AdaBoost	
	Ensemble learning	Long time, a bit good result
TF-IDF + SVD	CNN	

$$\begin{aligned} \mathrm{FPR} &= \frac{\mathrm{FP}}{N_{\mathrm{all}}}, \\ \mathrm{TPR} &= \frac{\mathrm{TP}}{P_{\mathrm{all}}}. \end{aligned} \tag{4}$$

In the formula, N_{all} represent the total number of negative samples and P_{all} represent the total number of positive samples. An example of ROC curve is shown in Figure 3,

Curves A and B in the figure represent the two classification models, respectively. It can be judged from the figure that the model represented by curve B performs better than model A. Meanwhile, the value of AUC is equal to the area of the graph under the ROC curve. Generally, the larger the AUC value is, the better the model effect is. In Figure 3, the AUC of model A is smaller than that of model B, so the realization effect of model B is better.

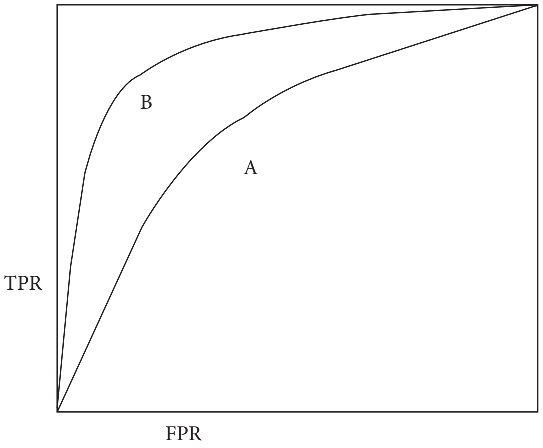


Figure 3: ROC curve diagram.

3.3. Experimental Results. After extracting keywords and reconstructing data sets from the TF-IDF model, gradient boosting classifier, random forest classifier, AdaBoost classifier, and classifier integrated with multiple classifiers are trained, respectively. Then, after the feature dimension reduction of the reconstructed data set, one-dimensional convolution CNN was input for training detection. At the same time, the source encrypted traffic data set uses One-Hot Encoding and then is input to the above classifier and CNN network training and detection. The accuracy and AUC values of the trainer and the convolutional neural network model obtained from the above experiments when testing the input test data are shown in Table 2.

It can be seen from Table 3 that, in the case of the same selected classifier and network structure, the encrypted malicious traffic detection method based on TF-IDF is significantly better than the detection method based on One-Hot Encoding. When the input data are all processed by the TF-IDF model, the detection effect of ensemble learning is better than that of other single classifiers. The detection effect of the convolutional neural network is better than that of the machine learning-based classifier, but the difference is not significant.

Table 4 shows the training time of each classifier. It can be seen from the table that the training time of the data processed by the TF-IDF model is significantly longer, and the more complex the network structure is, the more the training time is needed. Because ensemble learning is composed of many other classifiers, its training time will also increase significantly. Because of its complex network structure and a large amount of computation, the training time of CNN will also increase.

The confusion matrix obtained by TF-IDF based on the ensemble learning detection method is shown in Figure 4. After model training, 600 pieces of test data were input, among which 240 pieces of data were correctly predicted as benign samples and 317 pieces of data were correctly predicted as malicious traffic.

The ROC curve and AUC value obtained by the detection method based on TF-IDF-based ensemble learning are shown in Figure 5:

According to the ROC curve, the detection effect of the model is good, and the AUC value also reaches 0.929.

Because CNN has a better capability of feature extraction, its performance is better than ensemble learning. After the test, batch_size was set as 1000, and epoch was set as 30. In the 25th iteration, the model accuracy and loss changed little and tended to converge. Experimental results show that increasing the number of iterations will lead to a decrease in the performance of the model on the test set, and there is a tendency of slight overfitting. The changes in accuracy and loss values of its training set and test set are shown in Figure 6.

As shown in the figure, the accuracy rate of the training set tends to 1, showing a tendency of overfitting, but the accuracy rate of the test set tends to 0.933, and the effect is beyond reproach. Compared with the classifier based on machine learning, the training time of convolutional neural network is greatly increased, and it has certain requirements on the size of the data set.

Table 2: Confusion matrix definition.

		Actual category	
		True (1)	False (0)
Predicted class	True (1)	True positive (TP)	False positive (FP)
	False (0)	False negative (FN)	True negative (TN)

Table 3: The detection accuracy and AUC value obtained by different methods.

Detection method	Accuracy (TF-IDF)	Accuracy (encoding)	AUC
Gradient boosting	0.880	0.487	0.873
Random forest	0.922	0.492	0.918
AdaBoost	0.918	0.497	0.918
Ensemble learning	0.931	0.492	0.929
CNN	0.933	Huge	*

Table 4: Training time of each classifier.

Detection method	Training time (TF-IDF) (s)	Training time (encoding)
Gradient boosting	46	0.4 s
Random forest	83	6.5 s
AdaBoost	578	3.3 s
Ensemble learning	3537	11.27 s
CNN	752	Huge

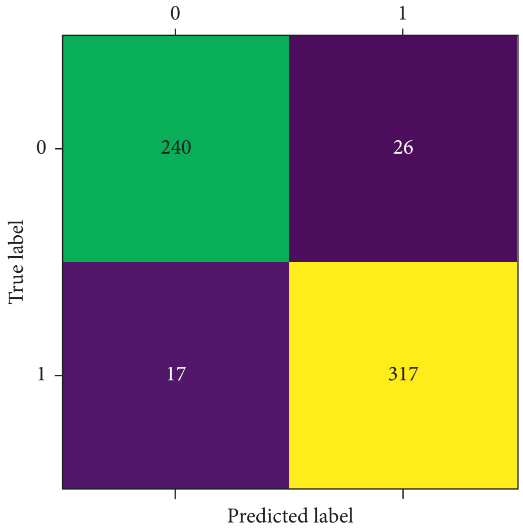


Figure 4: The confusion matrix of experimental results of ensemble learning.

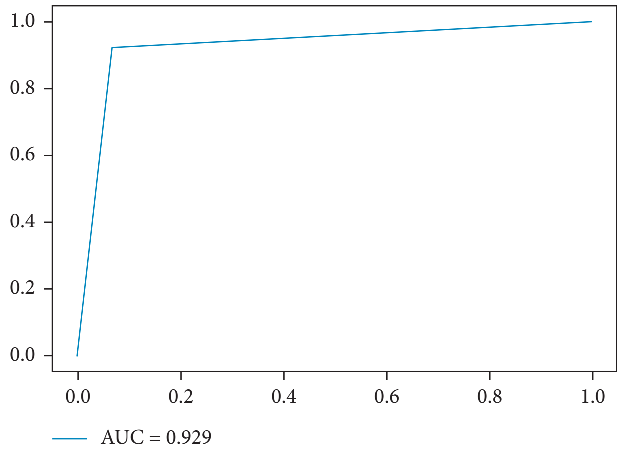


Figure 5: The ROC curve and AUC value of the experimental results of ensemble learning are obtained.

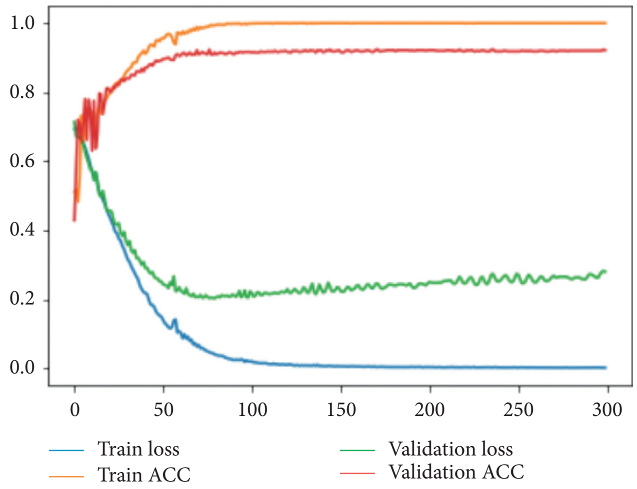


Figure 6: Loss value and accuracy value of CNN's experimental results.

4. Conclusions

Traditional detection methods, such as pattern matching, are difficult to deal with encrypted traffic data. With the development of machine learning and deep learning, the problem of encrypting data can be easily solved. At present, the detection of malicious encrypted traffic mostly adopts the method of machine learning. Although the detection method based on machine learning does not need to decrypt the encrypted traffic and is fast, it requires professionals to analyze and process the traffic data, which costs manpower and time. And the proposed detection method is based on the TF-IDF model, because the TF-IDF model does not care about the specific meaning of the data set, it replaces keywords in the source data with numbers that are calculated by their importance, so the detection method based on TF-IDF model not only applies to the malicious traffic detection but also can be used in other fields related detection, such as malicious code detection. It has strong generalization and accuracy. If the classifier model or neural network is adjusted and changed in the later stage, it is not necessary to limit the processing of the data set and information extraction. However, the detection method based on TF-IDF does not cover a comprehensive field, because the TF-IDF model simply measures the importance of a word by "word frequency," which is not comprehensive enough. Sometimes, important words may not appear many times in some data sets. In addition, the TF-IDF algorithm cannot reflect the position information of words. The words appearing in the first position and the words appearing in the second position are regarded as having the same importance, which is not accurate, and this should be taken into account in the case of different data sets.

The feature vectors reconstructed by the TF-IDF model are very sparse. This directly results in the resulting new data set being several times larger than the source data set. As the amount of data increases, memory consumption increases, leading to a significant increase in training time. For machine learning algorithms, the effect of data set size is less than that of neural network. For the convolutional neural network, due to its large amount of computation and limited by the size of the data set, there is still room for improvement in the early feature engineering processing. The experiment tried to compress the matrix, but it backfired. The following work will be improved from matrix compression, feature extraction, feature selection, and other aspects. With the further expansion of the size of the data set, if the Truncated SVD dimension reduction method cannot improve the efficiency of model construction without having a small impact on the accuracy of model recognition, then other schemes need to be reconsidered. Therefore, in the case of insufficient hardware conditions, the encrypted malicious traffic detection method based on the TF-IDF model is more suitable to use the classifier based on machine learning. Although the integrated learning model in machine learning is slightly inferior to the CNN in deep learning, the machine learning algorithm does not need to deal with the sparse matrix generated by the TF-IDF model and can retain the source data information to the maximum extent, with an accuracy rate of 0.93. Although the classifier based on ensemble learning has relatively high accuracy, the classifier based on ensemble learning has a disadvantage compared with a single classifier, which is significantly longer training time. In the case of abundant hardware resources, CNN has obvious advantages regardless of whether the sparse matrix is compressed [23–26].

Data Availability

The experimental data are real network capture packet data, provided by Qianxin Company. The data link is https://datacon.qianxin.com/opendata/maliciousstream.

Conflicts of Interest

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

Acknowledgments

This work was sponsored by the Sichuan Science and Technology Program (2020YFS0355 and 2020YFG0479).

References

[1] Computing Supercomputing, "Study data from Wuhan University of Technology update understanding of supercomputing (Ths-idpc: a three-stage hierarchical sampling method based on improved density peaks clustering algorithm for encrypted malicious traffic detection)," *Mathematics Week*, vol. 76, pp. 7489–7518, 2020.

[2] British Telecommunications Public Limited Company, *Patent Issued for Learned Profiles for Malicious Encrypted Network Traffic Identification (USPTO 10,594,707)*, Telecommunications Weekly, Beijing, China, 2020.

[3] O. L. Lyashuk, V. M. Klendii, O. Y. Gurik, and L. M. Slobodian, "Stand for investigation of the characteristics of screw downloaders," *Vìsnik Žitomirs'kogo Deržavnogo*

Tehnologìčnogo Unìversitetu: Tehnìčnì Nauki, vol. 2, no. 80, 2017.

[4] C. Michele, D. M. Mario, L. Maurizio, and S. Andrea, "Detection of encrypted multimedia traffic through extraction and parameterization of recurrence plots. Science and engineering research center," in *Proceedings of the 2016 International Conference on Sustainable Energy, Environment and Information Engineering (SEEIE 2016)*, vol. 5, Bangkok, Thailand, March 2016.

[5] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "DISTILLER: encrypted traffic classification via multimodal multitask deep learning," *Journal of Network and Computer Applications*, vol. 183-184, Article ID 102985, 2021.

[6] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," *Journal of Information Science*, vol. 44, no. 1, pp. 28–47, 2018.

[7] A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," *Expert Systems with Applications*, vol. 57, pp. 232–247, 2016.

[8] Z. Zhou, J. Qin, X. Xiang, Y. Tan, Q. Liu, and N. Xiong, "News text topic clustering optimized method based on TF-IDF algorithm on spark," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 217–231, 2020.

[9] S. Yan, H. Jia, and S. Hongping, "The study of disease symptom weight mining based on text mining word frequency inverse document frequency method," *Journal of Chengdu University of Information Technology*, vol. 29, no. 1, pp. 52–58, 2014.

[10] Z. Pan, X. Yi, Y. Zhang, H. Yuan, F. L. Wang, and S. Kwong, "Frame-level bit allocation optimization based on video content characteristics for HEVC," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 1, pp. 1–20, 2020.

[11] P. Lu and Q. Zongfeng, "Research on network public opinion detection based on improved TF-IDF algorithm," in *Proceedings of the 2019 the 9th International Workshop on Computer Science and Engineering (WCSE 2019)*, vol. 6, Hong Kong, June 2019.

[12] H. F. Jerome, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, 2001.

[13] K. Brian and B. Richard, "Small area estimation of the homeless in Los Angeles: an application of cost-sensitive stochastic gradient boosting," *Annals of Applied Statistics*, vol. 4, no. 3, 2010.

[14] J. Cheng, C. Cai, X. Tang, V. S. Sheng, and W. Guo, "A DDOS attack information fusion method based on CNN for multi-element data," *Computers, Materials & Continua*, vol. 63, no. 1, pp. 131–150, 2020.

[15] C. Guang-liang, T. Huan, Z. Fan, and Y. Sheng-liang, "AdaBoost-SVM based undergraduates evaluations," in *Proceedings of the 2019 2nd International Conference on Informatics, Control and Automation (ICA 2019)*, vol. 5, Advanced Science and Industry Research Center: Science and Engineering Research Center, Barcelona, Spain, September 2019.

[16] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, "Network traffic classifier with convolutional and recurrent neural networks for Internet of Things," *IEEE Access*, vol. 5, pp. 18042–18050, 2017.

[17] A. Onan, "Deep learning based sentiment analysis on product reviews on twitter," in *Proceedings of the International Conference on Big Data Innovations and Applications*, Springer, Istanbul, Turkey, August 2019.

[18] J. Gu, Z. Wang, J. Kuen et al., "Recent advances in convolutional neural networks," 2015, https://arxiv.org/abs/1512.07108.

[19] Z. Pan, X. Yi, Y. Zhang, B. Jeon, and S. Kwong, "Efficient in-loop filtering based on enhanced deep convolutional neural networks for HEVC," *IEEE Transactions on Image Processing*, vol. 29, pp. 5352–5366, 2020.

[20] L. Shen, X. Chen, Z. Pan, K. Fan, F. Li, and J. Lei, "No-reference stereoscopic image quality assessment based on global and local content characteristics," *Neurocomputing*, vol. 424, pp. 132–142, 2021.

[21] A. Alhussain, H. Kurdi, and L. Altoaimy, "A neural network-based trust management system for edge devices in peer-to-peer networks," *Computers, Materials & Continua*, vol. 59, no. 3, pp. 805–816, 2019.

[22] L. Shishkin Serge, A. Shalaginov, and D. Bopardikar Shaunak, "Fast approximate truncated SVD," *Numerical Linear Algebra with Applications*, vol. 26, no. 4, 2019.

[23] Z. Xiao, "Research on preprocessing method of performance monitoring data in cloud environment," in *Proceedings of the 2019 International Conference on Wireless Communication, Network and Multimedia Engineering (WCNME 2019)*, vol. 4, Advanced Science and Industry Research Center: Science and Engineering Research Center, Guilin, China, April 2019.

[24] R. Li, G. Sun, J. He et al., "Gender forecast based on the information about people who violated traffic principle," *Journal on Internet of Things*, vol. 2, no. 2, pp. 65–73, 2020.

[25] B. Mohammed and D. Naouel, "An efficient greedy traffic aware routing scheme for internet of vehicles," *Computers, Materials & Continua*, vol. 60, no. 3, pp. 959–972, 2019.

[26] A. Tahani and I. C. Alexandra, "Predicting learners' demographics characteristics deep learning ensemble architecture for learners' characteristics prediction in MOOCs," in *Proceedings of the 4th International Conference on Information and Education Innovations (ICIEI 2019)*, vol. 5, Durham, UK, July 2019.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 3676720, 11 pages
https://doi.org/10.1155/2021/3676720

Research Article

SFRNet: Feature Extraction-Fusion Steganalysis Network Based on Squeeze-and-Excitation Block and RepVgg Block

Guiyong Xu, Yang Xu, Sicong Zhang, and Xiaoyao Xie

Key Laboratory of Information and Computing Science of Guizhou Province, Guizhou Normal University, Guiyang 550001, China

Correspondence should be addressed to Yang Xu; xy@gznu.edu.cn

Received 1 May 2021; Accepted 16 July 2021; Published 26 July 2021

Academic Editor: Zhaoqing Pan

In the era of big data, convolutional neural network (CNN) has been widely used in the field of image classification and has achieved excellent performance. More and more researchers are beginning to combine deep neural networks with steganalysis to improve performance in recent years. However, most of the steganalysis algorithm based on the convolutional neural network has only run test against the WOW and S-UNIWARD algorithms; meanwhile, their versatility is insufficient due to long training time and the limit of image size. This paper proposes a new network architecture, called SFRNet, to solve these problems. The feature extraction and fusion layer can extract more features from the digital image. The RepVgg block is used to accelerate the inference and increase memory utilization. The SE block improves the detection accuracy rate because it can learn feature weights to make effective feature maps with significant weights and invalid or ineffective feature maps with small weights. Experimental results show that the SFRNet has achieved excellent performance in the detection accuracy rate against four state-of-the-art steganography algorithms in the spatial domain, e.g., HUGO, WOW, S-UNIWARD, and MiPOD, under different payloads. The SFRNet detection accuracy rate achieves 89.6% against S-UNIWARD algorithm with the payload of 0.4bpp and 72.5% at 0.2bpp. As the same time, the training time of our network is greatly reduced by 35% compared with Yedroudj-Net.

1. Introduction

The rapid development of social networks provides convenience for users to exchange data. A large number of digital images are uploaded to the Internet every day. The proliferation of digital images provides a good medium for criminals to commit crimes using steganographic algorithms. Digital image steganography is a hiding technology that takes into account data security and communication, which uses the redundancy of the cover image to embed the secret information into the public carrier and transmits it through the public channel to ensure that the secret information is not discovered and intercepted by a third party. Image steganalysis is the opposite of image steganography, which can determine whether the image contains secret information by capturing minor disturbances in the stego image that are not easily perceivable by the human visual system. They provide a basis for extracting the secret information hidden in the image. In recent years, image steganalysis has played an increasingly important role in many information security systems and has attracted many researchers [1]. At the same time, the fast-developing adaptive steganography algorithm uses syndrome-trellis code (STC) [2] to minimize distortion and retains more complex image statistical properties. The current typical spatial adaptive steganography algorithms include HUGO [3], S-UNWIWARD [4], WOW [5], and MiPOD [6]. They make the secret information more cleverly hidden in the area where it is difficult to establish a steganalysis model, which improved the security of steganography algorithms and brought significant challenges to steganalysis.

Traditional steganalysis models include subtractive pixel adjacency matrix (SPAM) [7], spatial rich model (SRM) [8], max spatial rich model (maxSRM) [9], and its variant maxSRMd2, which are all feature extraction methods based on manual design. In recent years, convolutional neural network (CNN) has been widely used in image and video processing and has achieved excellent performance [10–15].

Steganalysis can be considered as a two-class problem of images. Since convolutional neural networks can extract features in the spatial and frequency domains of images, more and more researchers are beginning to combine deep neural networks with steganalysis to improve performance. The signal noise processed by steganalysis is a weak signal which will be affected by image content so that it will be ignored by the traditional classification network. The network needs to be specially modified before it can be used in steganalysis, such as suppressing the image content and enhancing the steganographic noise signal.

Qian et al. [16] proposed a steganalysis network Qian-Net, based on a convolutional neural network, using a Gaussian activation function to replace the rectifying linear unit (ReLU) [17] activation function. Xu et al. [18] proposed Xu-Net based on the Qian-Net framework. The high-pass filter is used to extract noise residuals in the preprocessing layer. Simultaneously, the network adds the absolute value (ABS) layer and TanH-ReLU hybrid activation function. Jian et al. [19] proposed Ye-Net with a deeper network structure, using high-pass filters as the preprocessing layer and the truncated linear unit (TLU) as the activation function, introducing the selection channel. Boroumand et al. [20] proposed a 48-layer deep learning steganalysis framework SRNet, which obtains filters through learning to improve the detection accuracy rate of the network against steganography algorithms. Yedroudj et al. [21] proposed a network architecture based on the concept of Alex-Net [22], called Yedroudj-Net. In addition to using ABS layer and TLU activation function, three fully connected layers are also used. Zhang et al. [23] proposed Zhu-Net, which optimizes the filter kernel of the preprocessing layer and uses pyramid pooling [24] to obtain excellent detection accuracy rate on the S-UNIWARD and WOW algorithms.

The problem of steganographic analysis tool with neural networks is that it is impossible to analyze larger-sized images due to limitations in computer resources. And, the versatility of such steganographic analysis tools is not good. Most of the steganalysis algorithm has only run test against the WOW and S-UNIWARD algorithms. At the same time, the training time of the neural network is too long. To enhance the practicality and universality of the steganalysis network framework, we propose a feature fusion steganalysis framework based on the network structure of the RepVgg [25] and squeeze-and-excitation [26] in this paper, which is called the SFRNet. Experimental results show that the SFRNet has achieved excellent performance in the detection accuracy rate of four different steganography algorithms under some different payloads. The SFRNet detection accuracy rate achieves 89.6% against S-UNIWARD algorithm with the payload of 0.4bpp and 72.5% at 0.2bpp. In summary, we make the following contributions in this paper:

(i) Instead of using the image as an input, we extract and merge the feature of images into a feature matrix through the rich model and use the generated feature matrix as the actual input of the work, which solves the dependence of the deep neural network on the size of the input image.

(ii) We propose SFRNet, a simple architecture with favorable speed-accuracy trade-off compared to the state of the arts, which uses the RepVgg block as the convolution layer of the network and uses the squeeze-and-excitation (SE) block to improve the detection accuracy rate.

(iii) We show the effectiveness of the SFRNet in steganalysis and the efficiency and ease of implementation.

The rest of the paper is organized as follows. Section 2 introduces the prior knowledge including SRM and its several variants and the deep learning methods. In Section 3, the SFRNet is proposed. This section describes feature extraction-fusion and the detailed structure of SFRNet. In Section 4, the dataset partition, training details, and specific parameters of the SFRNet steganalysis framework are introduced. In Section 5, we validate the effective proposed model on several states-of-the-art steganographic algorithms and compare the performance of the SFRNet with several advanced steganalysis algorithms. The study is ends with the conclusion in Section 6.

2. Preliminaries

2.1. The Feature Extraction Method. Friedrich and Kodovsky [8] proposed the spatial rich model (SRM) based on the subtractive pixel adjacency matrix (SPAM) model, which designed various linear and nonlinear high-pass filters (HPF) in spatial domains. It uses these filters to filter the image to obtain a wide variety of residual images and then separately counts the frequency of occurrence of each adjacent residual sample pattern to get the co-occurrence matrix. Finally, the elements of the co-occurrence matrix are arranged into vectors as steganographic analysis features, as shown in Figure 1. The steganographic analysis features can comprehensively perceive the change of image adjacent pixel correlation caused by steganography algorithm. The SRM improves the detection accuracy rate of steganalysis algorithm, which has been used and improved by researchers of general steganalysis.

Denemark et al. [9] proposed the steganalysis method maxSRM combined with the channel selection strategy, which is a variant of the so-called SRM. The maxSRM and maxSRMd2 are built in the same manner as the SRM, but the process of forming the co-occurrence matrices is modified to consider the embedding change probabilities estimated from the analyzed image. The version of the maxSRM with all co-occurrence scan directions replaced with the oblique direction “d2,” as shown in Figure 2, is called maxSRMd2. Compared with SRM, maxSRM and maxSRMd2 have significant performance improvement.

2.2. The RepVgg Block. A classic convolutional neural network (ConvNet), VGG [27], achieved massive success in image recognition with a simple architecture composed of a stack of Conv, ReLU, and pooling. With Inception, ResNet [28], and DenseNet, many research interests were shifted to

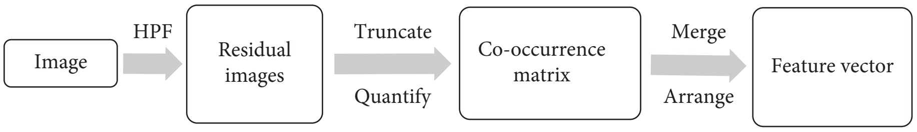
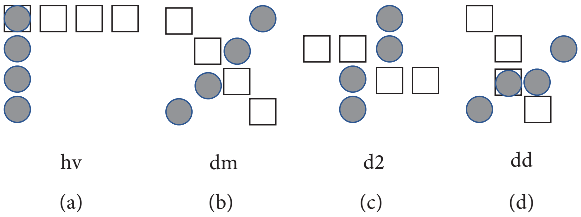


Figure 1: Feature extraction process.



(a) (b) (c) (d)

Figure 2: Four types of co-occurrence scan direction.

well-designed architectures, making the models more and more complicated. The complicated multibranch designs make the model difficult to implement, customize, slow down the inference, and reduce memory utilization. Ding et al. [25] presented RepVgg, a VGG-like inference-time body composed of nothing but a stack of 3×3 convolution and ReLU, while the training-time model has a multibranch topology.

In the SFRNet, we used the RepVgg block instead of the conventional convolution to accelerate the inference and increase memory utilization. The RepVgg block use ResNet-like identity and 1×1 branches so that the training-time information flow of a building block is $y = x + g(x) + f(x)$. It uses $W^{(3)} \in \mathbb{R}^{C_2 \times C_1 \times 3 \times 3}$ to denote the kernel of a 3×3 conv layer with C_1 input channels and C_2 output channels and $W^{(1)} \in \mathbb{R}^{C_2 \times C_2}$ for the kernel of 1×1 branch. It uses $\mu^{(3)}, \sigma^{(3)}, \gamma^{(3)}, \beta^{(3)}$ as the accumulated mean, standard deviation, learned scaling factor, and bias of the BN layer following 3×3 conv layer, $\mu^{(1)}, \sigma^{(1)}, \gamma^{(1)}, \beta^{(1)}$ for the BN layer following 1×1 conv layer, and $\mu^{(0)}, \sigma^{(0)}, \gamma^{(0)}, \beta^{(0)}$ for the identity branch. The identity branch can be viewed as a 1×1 conv layer with an identity matrix as the kernel. Let $M^{(1)} \in \mathbb{R}^{N \times C_1 \times H_1 \times W_1}$ and $M^{(2)} \in \mathbb{R}^{N \times C_2 \times H_2 \times W_2}$ be the input and output and $*$ be the convolution operator:

$$\begin{aligned} M^{(2)} = {} & bn\left(M^{(1)} * W^{(3)}, \mu^{(3)}, \sigma^{(3)}, \gamma^{(3)}, \beta^{(3)}\right) \\ & + bn\left(M^{(1)} * W^{(1)}, \mu^{(1)}, \sigma^{(1)}, \gamma^{(1)}, \beta^{(1)}\right) \\ & + bn\left(M^{(1)}, \mu^{(0)}, \sigma^{(0)}, \gamma^{(0)}, \beta^{(0)}\right). \end{aligned} \tag{1}$$

Then, it obtains the final bias by adding up the three bias vectors and the final 3×3 kernel by adding the 1×1 kernels onto the central point of 3×3 kernel, which can be easily implemented by first zero-padding the two 1×1 kernels to 3×3 and adding the three kernels up [25], as shown in Figure 3.

2.3. The Squeeze-and-Excitation Block. He et al. [26] focus on the channel relationship and propose the "Squeeze-and-Excitation" block, which can learn to use global information to emphasize informative features and suppress less useful ones selectively. Liu et al. [29] construct a new effective network with diverse filter modules (DFMs) and squeeze-and-excitation modules (SEMs), called DFSE-Net, which can better capture the embedding artifacts. The experiments presented that networks can pay more attention to critical channels by SEMs.

The squeeze-and-excitation block is not a complete network structure, which can construct a squeeze-and-excitation network by simply stacking a collection of SE blocks. The SE block can learn feature weights to make effective feature maps with significant weights and invalid or ineffective feature maps with small weights, as shown in Figure 4.

3. SFRNet

The proposed network architecture is called SFRNet: feature extraction-fusion steganalysis network via squeeze-and-excitation block and RepVgg block. Firstly, we explain the method of preprocessing, i.e., how to get the feature matrix. Then, the architecture of network is demonstrated. At the same time, we explored the values of key parameters through experiments.

3.1. Feature Extraction and Fusion Layer. The steganography algorithm modifies the original image content as little as possible when embedding secret information in the cover image to avoid detection. In other words, the steganography algorithm introduces noise in the image, which usually cannot be perceived by the human perceptual system. And, the noise is also easily ignored by those image classification networks which focus on the content of the image. At the same time, it modifies the correlation between adjacent pixels of the original image while also modifying the correlation between adjacent pixels of the residual image. The SRM and its variants are used to process the image, mainly to suppress the relevance of image content. We propose a feature information fusion block inspired by [30].

We use the following steps to extract and fuse the feature to obtain the feature matrix as input of the model. First, the residual image of the stego image and the cover image is

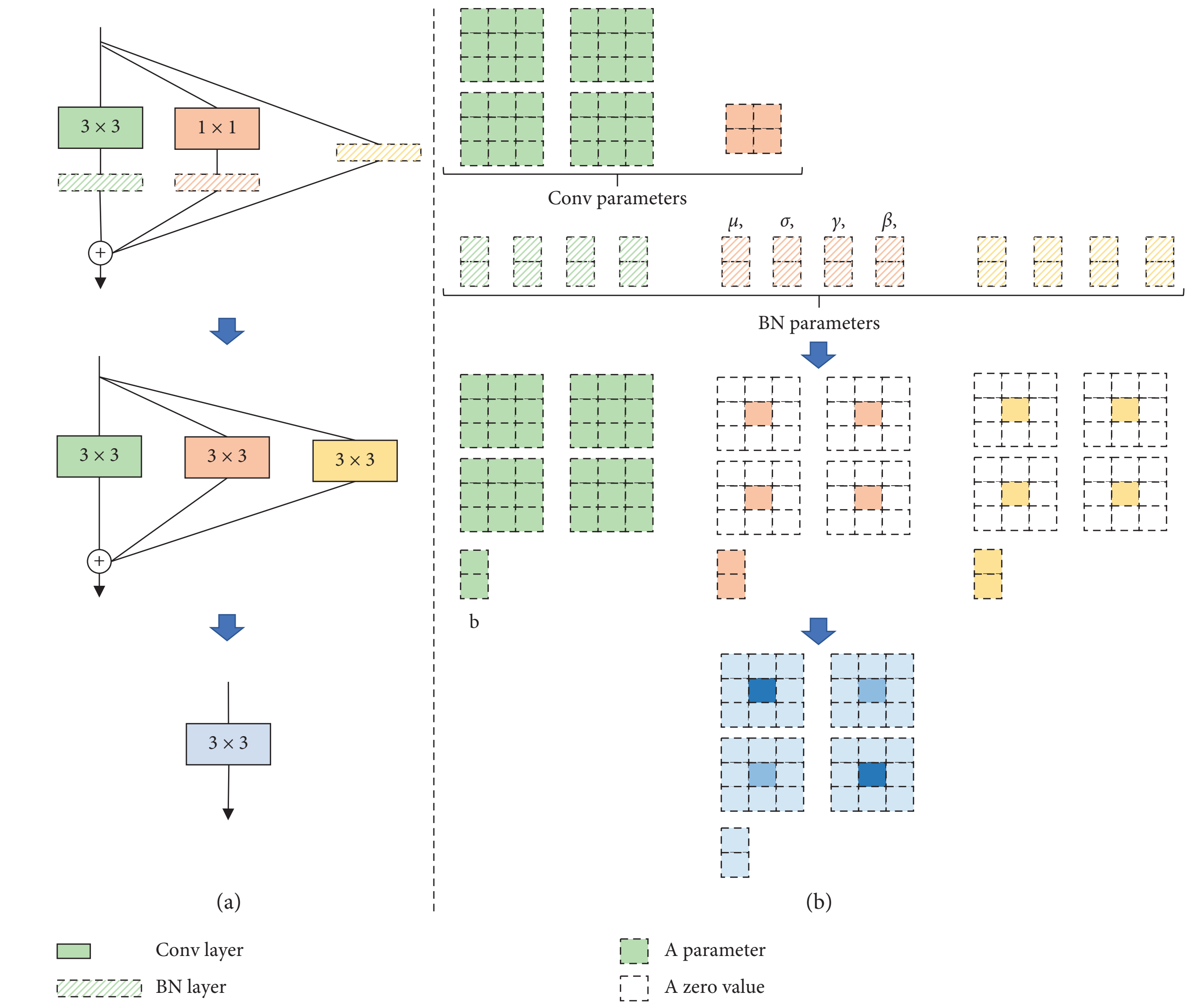


Figure 3: The RepVgg block.

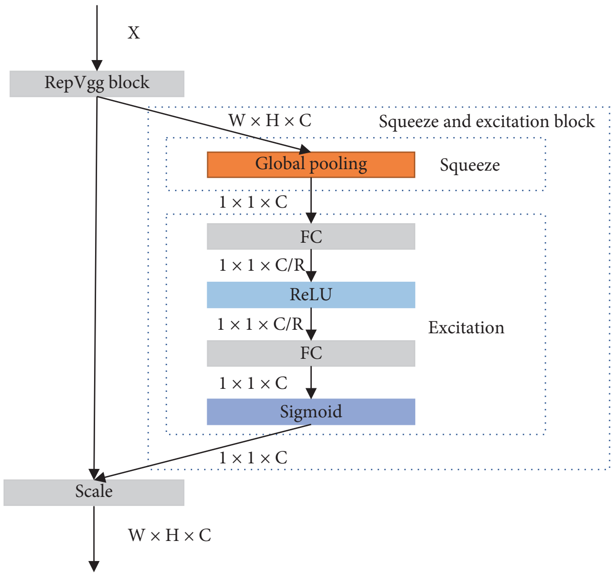


Figure 4: The squeeze-and-excitation block.

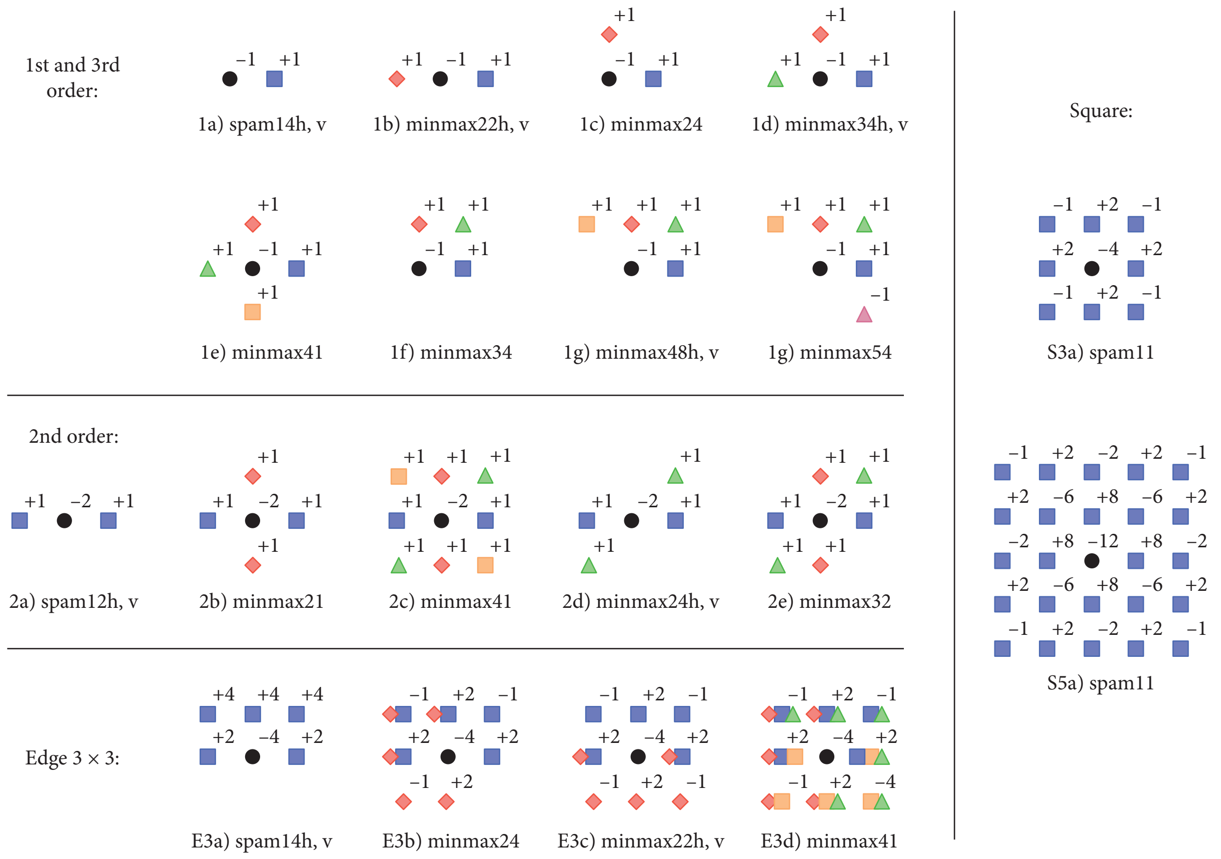
filtered by the high-pass filters to obtain submodels. Then quantize, round, and truncate each submodel and extract the co-occurrence matrix. Finally, the feature vectors are obtained by using the merging rules in SRM to process the co-occurrence matrix. The high-pass filters are shown in Figure 5.

The feature vector extraction process is defined as the follow equations:

$$R_{ij}^{k} = \hat{X}_{ij}\left(N_{ij}\right) - cX_{ij} \Leftrightarrow R^{K} = X * K^{k}, \tag{2}$$

where X_{ij} represents the i, j pixel of the cover image, N_{ij} is the adjacent pixels of X_{ij}, $X_{ij} \neq N_{ij}$, $c \in \mathbb{N}$ is residual order, $\hat{X}_{ij}(\cdot)$ is a predictor of cX_{ij} defined on N_{ij}, K^{k} is kth high-pass filters, and R^{K} is the residual filtered by the kth high-pass filter:

$$R_{ij}^{k} \leftarrow \mathrm{Trunc}_{T}\left(\mathrm{Round}\left(\frac{R_{ij}^{k}}{q}\right)\right), \tag{3}$$



1a) spam14h, v; 1b) minmax22h, v; 1c) minmax24; 1d) minmax34h, v; 1e) minmax41; 1f) minmax34; 1g) minmax48h, v; 1g) minmax54; S3a) spam11; 2a) spam12h, v; 2b) minmax21; 2c) minmax41; 2d) minmax24h, v; 2e) minmax32; S5a) spam11; E3a) spam14h, v; E3b) minmax24; E3c) minmax22h, v; E3d) minmax41

Figure 5: The high-pass filters.

where Round(·) means rounding up by element and Trunc(·) means a truncation operation by element. The purpose of truncation is to curb the dynamic range of residual to all description using co-occurrence matrices with a small T.

The SRM model extracts the co-occurrence matrix in the horizontal, defined as equation (4). The vertical co-occurrence, $C_d^{(v)}$, is defined analogically:

$$C_d^{(h)} = \sum_{i,j=1}^{n_1,n_2-3} \left[R^k_{i,j+n} = d_n, \forall n = 0,1,2,3\right]. \tag{4}$$

The maxSRM model extracts the co-occurrence matrix in the horizontal, defined by equation (5), where $\widehat{\beta}^k_{ij}$ is the embedding change probabilities in the kth high-pass filters; refer to [9], for details. The vertical co-occurrence, $C_d^{(v)}$, is defined analogically:

$$C_d^{(h)} = \sum_{i,j=1}^{n_1,n_2-3} \max_{n=0,1,2,3} \widehat{\beta}^k_{i,(j+n)} \left[\widehat{\beta}^k_{ij} = d_n, \forall n = 0,1,2,3\right]. \tag{5}$$

The scan direction of the maxSRMd2 model is different from SRM and maxSRM, which are replaced by "d2," as shown in Figure 2, so the co-occurrence matrix is defined as equations (6) and (7):

$$C_d^{(+)} = \sum_{i,j=1}^{n_{1-3},n_2} \overline{b}_{i,j} \times \left[R^k_{i,j} = d_0, R^k_{i,(j+1)} = d_1, R^k_{(i+1),(j+2)} = d_2, R^k_{(i+1),(j+3)} = d_3\right], \tag{6}$$

$$\begin{gathered} C_d^{(-)} = \sum_{i,j=1}^{n_{1-3},n_2} \widetilde{b}_{i,j} \times \left[R^k_{i,j} = d_0, R^k_{i,(j+1)} = d_1, R^k_{(i+1),(j+2)} = d_2, R^k_{(i+1),(j+3)} = d_3\right], \\ \overline{b}_{i,j} = \max\left\{\beta^k_{i,j}, \beta^k_{i,(j+1)}, \beta^k_{(i+1),(j+2)}, \beta^k_{(i+1),(j+3)}\right\}, \\ \widetilde{b}_{i,j} = \max\left\{\beta^k_{(i-1),j}, \beta^k_{(i-1),(j+1)}, \beta^k_{i,j+2}, \beta^k_{i,(j+3)}\right\}. \end{gathered} \tag{7}$$

We choose $q \in [0.5, 1, 2], T = 2$, and $d = 4$ in all extraction methods to extract feature vector, getting 106 feature vectors. Among them, 17 are 338-dimensional feature vectors and 89 are 325-dimensional features vectors, and the feature vector is defined as

$$\vec{F}_k \leftarrow \text{Range}\left(\text{Merge}\left(C_d^{(h)}, C_d^{(v)}\right)\right), \tag{8}$$

where $\vec{F}_k$ is the feature vector calculated by using the kth submodel and $\text{Merge}(\cdot)$ merges two matrices into one by combining elements with the same or similar statistical laws in the horizontal co-occurrence matrices $C_d^{(h)}$ and vertical one $C_d^{(v)}$.We use the zero vector [0, 0] as the segmentation between each feature vector to fill it into a feature vector of 34,969 dimensions, and null values after the last feature in the vector are filled with the zero vector [0,0,0, . . ., 0,0]. It is defined as equation (9), where $*$ can denote SRM, maxSRM, and maxSRMd2:

$$F_* \leftarrow \left[\vec{F}_1, 0, 0, \vec{F}_2, 0, 0, \vec{F}_3, 0, 0 \ldots . \vec{F}_{106}, 0, 0, \ldots, 0, 0, 0\right]. \tag{9}$$

Then, we obtain the finally feature matrix fused by the three feature vectors, which are defined as

$$\text{MF}_{\text{cover}} = \begin{bmatrix} \text{Reshape}\left(F_{\text{SRM}}\right) \\ \text{Reshape}\left(F_{\max \text{SRM}}\right) \\ \text{Reshape}\left(F_{\max \text{SRM}d2}\right) \end{bmatrix}, \tag{10}$$

where MF_{cover} is the feature matrix of the cover image, MF_{stego} is the feature matrix of the stego image, and $\text{Reshape}(\cdot)$ converts a feature vector of 34,969 dimensions to a feature matrix of 187×187. Finally, our goal is to use SEFNet to train a mapping $\text{Map}(\cdot)$ based on the difference between them so that the mapping satisfies equations (11) and (12):

$$\text{Map}(.) \leftarrow \text{SEFNet}\left(\text{MF}_{\text{cover}}, \text{MF}_{\text{stego}}\right), \tag{11}$$

$$\begin{cases} \text{Map}\left(\text{MF}_{\text{stego}}\right) = 1, \\ \text{Map}\left(\text{MF}_{cover}\right) = 0. \end{cases} \tag{12}$$

3.2. The SFRNet Architecture. The overall structure of the SFRNet is presented in Figure 6. The SFRNet accepts an input image of size 256×256 and outputs two-class labels (stego and cover), composed of several number of layers, including one feature extraction-fusion block, five convolution blocks with different amounts of the RepVgg block, three SE blocks, and three fully connected layers.

The layer types and parameters are displayed inside boxes in Figure 6. N×(C×W×H) means that the number of batch size is N, the number of channels is C, and the height and width of the feature matrix is W and H. RepVgg denotes the RepVgg block. The details of the RepVgg block and SE block are described below. The full name of AVG is Average Pooling. Similarly, GAP is global average pooling.

3.2.1. The SE Blocks. Squeeze is achieved by using global average polling to generate channel-wise statistics. The statistic Z is generated by squeezing the input U through its spatial dimensions H×W, and the cth element of z is calculated by

$$z_c = F_{sq}\left(u_c\right) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j). \tag{13}$$

Excitation is used to fully capture channel-wise dependencies. First, it must be capable of learning a nonlinear interaction between channels, and second, it must learn a nonmutually exclusive relationship. The operations of excitation can be defined by

$$\begin{aligned} s &= F_{ex}(z, W) \\ &= \sigma(g(z, W)) \\ &= \sigma\left(W_2 \delta\left(W_1 z\right)\right), \end{aligned} \tag{14}$$

where δ refers to the ReLU function, W_1 and W_2 are the fully connected operation, and σ refers to the sigmoid function. The final output of the SE block is obtained by rescaling U with the activations s:

$$\widetilde{x}_c = F_{\text{scale}}\left(\mathbf{u}_c, s_c\right) = \mathbf{u}_c s_c, \tag{15}$$

where $F_{\text{scale}}\left(\mathbf{u}_c, s_c\right)$ refers to channel-wise multiplication between the scalar s_c and the feature map $\mathbf{u}_c$ and $\widetilde{\mathbf{X}} = \left[\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_C\right]$ is the final output of the SE block.

In our architecture, the SE block is followed by the first three stages, as shown in Figure 6. To show the performance of the SE block against steganalysis algorithm, we conducted a comparative experiment based on the SFRNet with the SE block and without the SE block. The result in Figures 7 and 8 show that the SE block accelerates the convergence and shows better performance against WOW algorithm at 0.4 bpp.

3.2.2. Nonlinear Activation Layer. Two different activation functions, TLU and ReLU, are used in the SFRNet. The classical ReLU can prevent gradient vanishing/exploding and accelerate network convergence. The ReLU is used in “stage3,” which selectively responds to embedded signals among the input feature map and get more efficient feature. Note that the remaining layers do not use the activation function.

Compared with cover image content, the signal introduced by the embedded message is usually of low amplitude. The high-frequency stego noise adds to the cover as a weak signal, significantly impacted by the image content. Therefore, the TLU is used to reduce the dynamic range of input feature maps in “stage1” and “stage2,” suppressing image content and extract embedding signals more effectively. It can be defined as

$$\text{TLU}(x) = \begin{cases} T, & x > T, \\ x, & T \geq x \geq -T, \\ -T, & x < -T, \end{cases} \tag{16}$$

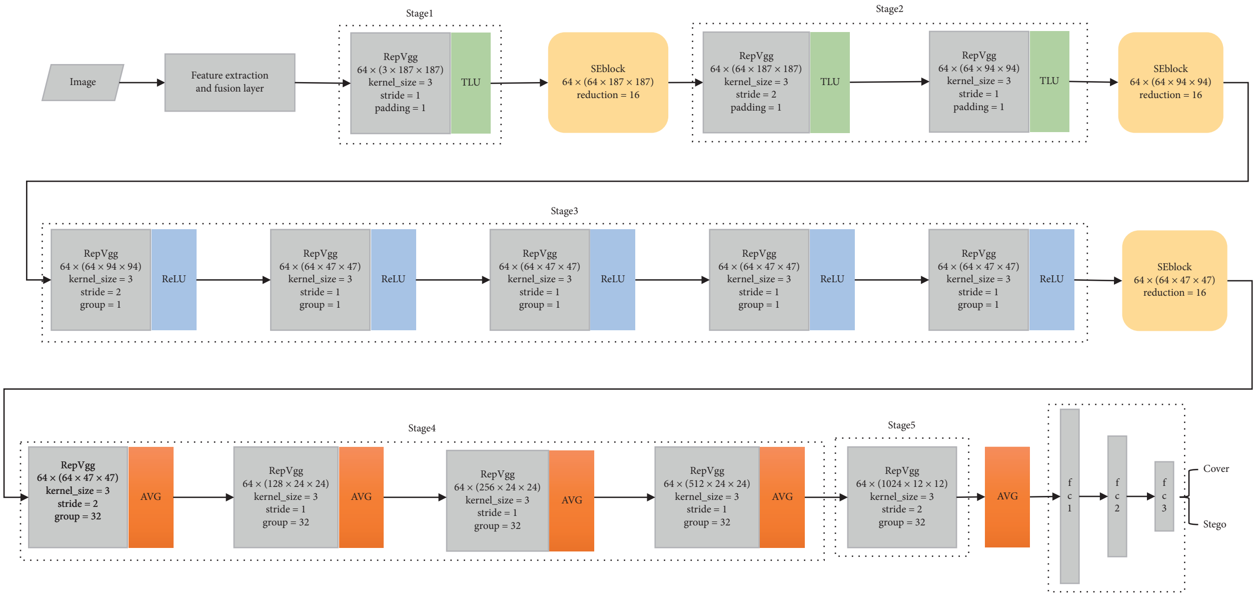


Figure 6: The SFRNet architecture.

where $T > 0$ is the threshold determined by experiments. To investigate the impact of parameter T in our network, we conduct several experiments with the SEFNet for a range of different T values. The results are shown in Table 1 and Figures 9 and 10. When the value of T is 1, the model achieves better performance and faster convergence.

4. Experiments

Python 3.8.3 was used for architecture construction, and the model was designed mainly with PyTorch 1.4.0. The operating system of the machine is Ubuntu 20.04 LTS, and the CUDA version is 11.0. The hardware of the machine has a GeForce RTX2080 SUPER with 8 GB and 250W, an Intel I7-9700k processor, and RAM with 32 GB (2 modules of 16 GB with 2666Mhz).

4.1. Dataset and the Steganographic Schemes. All experiments in this paper were evaluated and contrasted on the standard dataset BOSSBase ver. 1.01. This source contains 10,000 images acquired by seven digital cameras in the RAW format and subsequently processed by converting them to 8-bit grayscale, resizing, and central-cropping to 512 × 512 pixels. The image and camera information is shown in Table 2. The image source is widely used in research fields, such as information hiding, forensics, and steganalysis, which can be found at http://dde.binghamton.edu/download/.

Because other steganalysis algorithms use 256 × 256-size image as input, we decided to evaluate the effectiveness of all models on the images with a size of 256 × 256. To this end, we resized all the images into the size of 256 × 256 pixels using "imresize ()" in MATLAB with the default setting to generate the final datasets.

In our experiments, several state-of-the-art steganographic methods in the spatial domain, such as WOW, S-UNWARD, MiPOD, and HUGO, were employed to produce standard datasets. And, the embedding algorithms WOW and S-UNIWARD are implemented with STC simulator based on the publicly available codes, which can be found at the same URL of the BOSSBase original images. We use the MATLAB version rather than C++ implementation to avoid the problem as [31] that all images are embedded with the same key for all the steganographic algorithms. All methods were used to process the original images with two payloads: 0.2 bpp and 0.4 bpp. We use bit-per-pixel (bpp) to represent the size of secret data embedded into cover images in all experiments. For each steganography algorithm, we randomly select 5000 image pairs for training, 1000 image pairs for validating, and 4000 image pairs for testing, and the testing set was untouched during all of the training phases.

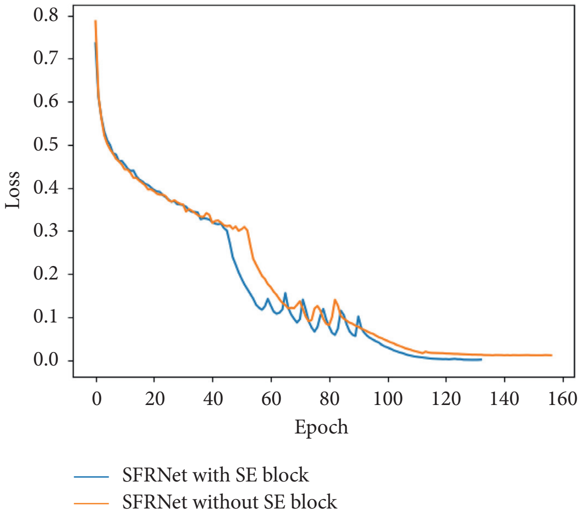


Figure 7: Comparing convergence performance of SFRNet with the SE block and SFRNet without the SE block against the WOW algorithm at 0.4 bpp.

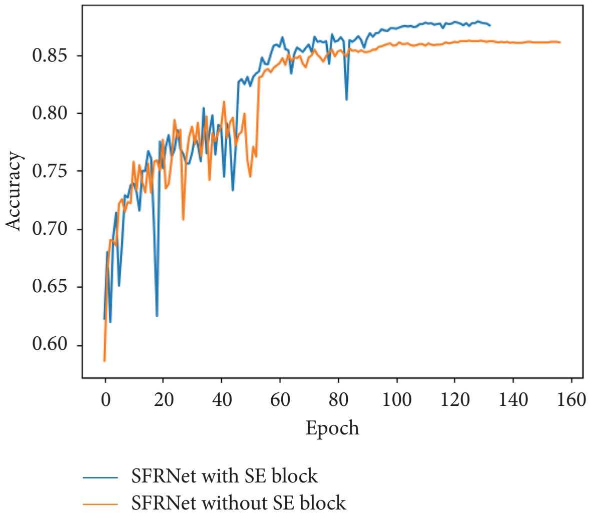


Figure 8: Comparing detection accuracy rate of SFRNet with the SE block and SFRNet without the SE block against the WOW algorithm at 0.4 bpp.

Table 1: The detection accuracy rate comparison of SFRNet against S-UNIWARD algorithm at 0.4bpp at different values of T.

T	1	2	3
Accuracy	0.8961	0.8869	0.8854

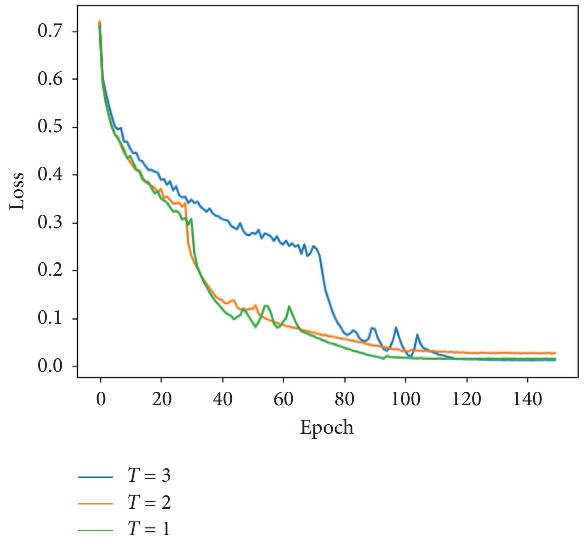


Figure 9: Comparing convergence performances of SFRNet with the different values of T against S-UNIWARD algorithm at 0.4 bpp.

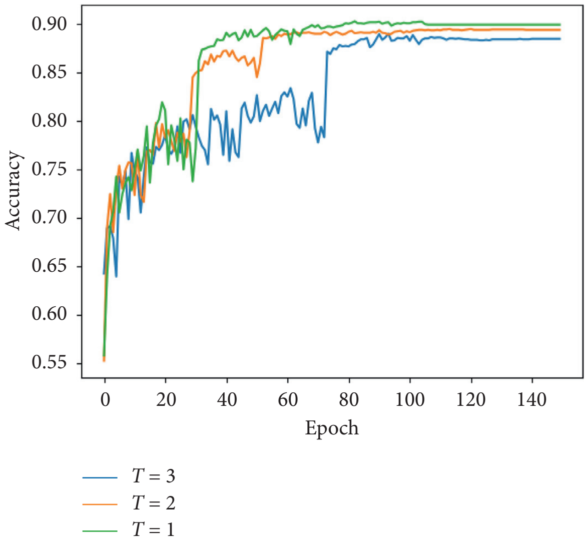


Figure 10: Comparing detection accuracy rate of SFRNet with the different values of T against S-UNIWARD algorithm at 0.4 bpp.

Table 2: Image and camera information.

Image number	Camera information
1-1354	Canon EOS 400D
1355-1415	Canon EOS 40D
1416-2769	Canon EOS 7D
2770-4811	Canon EOS DIGITAL REBEL XSi
4812-6209	PENTAX K2D
6210-7242	NIKON D70
7243-10000	M9 digital camera

Then, the feature extractor is employed to extract the feature for image steganalysis.

4.2. Hyperparameters. In SFRNet architecture, the Adam [32] optimizer is used to update the parameters of model in the learning phase since Adam can reach convergence faster than stochastic gradient descent (SGD). Due to GPU memory limitation, the minibatch size in training is set 64, containing 32 cover images and their 32 corresponding stego images. The training dataset was shuffled after each epoch. Dropout is used, which followed every fully connected layer. Based on the above settings, the networks are then trained to minimize the cross-entropy loss. The SFRNet training is up to 150 epochs. We often stop training before 150 epochs to prevent overfitting. When the cross-entropy loss on the training set keeps decreasing, detection accuracy rate on validation begins declining, and we stop the training. The performance was evaluated by the testing accuracy rate, where the best validation model obtained during training was selected.

5. Experimental Results and Analysis

5.1. Comparison with the State-of-the-Art Steganalysis. We report the detection accuracy rate obtained when detecting S-UNIWARD and WOW embedding algorithms at 0.2 bpp and 0.4 bpp, as shown in Table 3. The steganalysis methods are Yedroudj-Net, SRNet, DFSE-Net, and Zhu-Net. The detection accuracy of the Zhu-Net is 0.3% higher than the SFRNet when applied to the WOW algorithm with 0.4 bpp. In addition to this case, the SFRNet generally has better performance than the other four steganographic analysis networks against WOW and S-UNIWARD algorithm at 0.2 bpp and 0.4 bpp, as shown in Table 3 and Figure 11.

Table 3: Performance comparisons between proposed network and several state-of-the-art models on S-UNIWARD and WOW at two different payloads.

Steganography	CNN model	0.2 bbp	0.4 bpp
S-UNIWARD	Yedroudj-Net	63.0	78.1
	SRNet	67.4	81.6
	Zhu-Net	71.5	84.7
	DFSE-Net	65.9	78.5
	SFRNet	**72.5**	**89.6**
WOW	Yedroudj-Net	72.3	83.1
	SRNet	75.4	86.9
	Zhu-Net	76.7	**88.2**
	DFSE-Net	75.3	85.1
	SFRNet	**76.8**	87.9

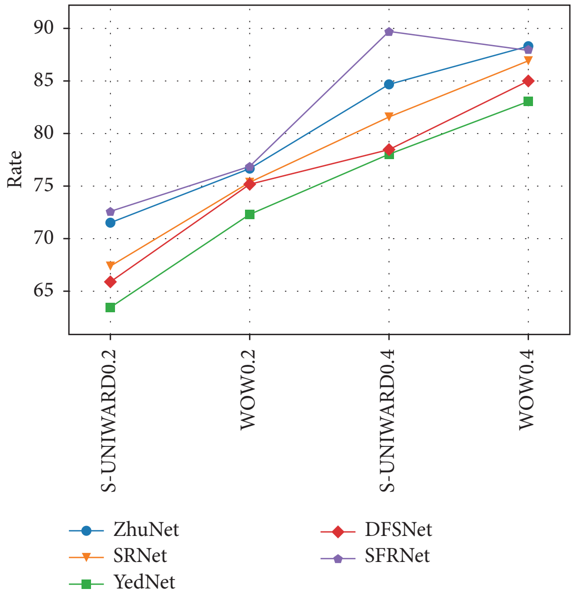


Figure 11: Comparison of the accuracy percentage of steganalysis algorithm among six steganalysis methods against two algorithms: S-UNIWARD and WOW at 0.2 bpp and 0.4 bpp.

The result shows that the proposed SFRNet outperforms Zhu-Net, SRNet against MiPOD, and HUGO embedding algorithms at 0.2 bpp and 0.4 bpp, as shown in Table 4 and Figure 12. The detection accuracy is increased by 8%–10% compared with the latest method Zhu-Net against MiPOD algorithm at 0.4 bpp and 0.2 bpp. Compared with Zhu-Net, the detection accuracy is increased by 4%–7% against HUGO algorithm at 0.2 bpp and 0.4 bpp. The good performance demonstrates the effectiveness of the network structure of the SFRNet in Figure 12.

5.2. The Time Consumption and Computational Complexity of SFRNet. We compare the number of parameters and times spent on network training and testing of the six types of steganalysis networks, as shown in Table 5. The SFRNet also reduces time consumption while improving accuracy. Although the SFRNet designed in this paper is a deeper network structure, the application of the RepVgg block reduces the computational complexity and time consumption than the Zhu-net and SRNet while still ensuring considerable accuracy. Compared to Yedroudj-Net, training time is reduced by about 35%.

Table 4: Performance comparisons between proposed network and several state-of-the-art models on MiPOD and HUGO at two different payloads.

Steganography	CNN model	0.2 bbp	0.4 bpp
MiPOD	SRNet	64.3	75.1
	Zhu-Net	65.2	76.1
	SFRNet	**75.2**	**84.1**
HUGO	SRNet	67.1	78.7
	Zhu-Net	68.1	79.3
	SFRNet	**75.4**	**83.6**

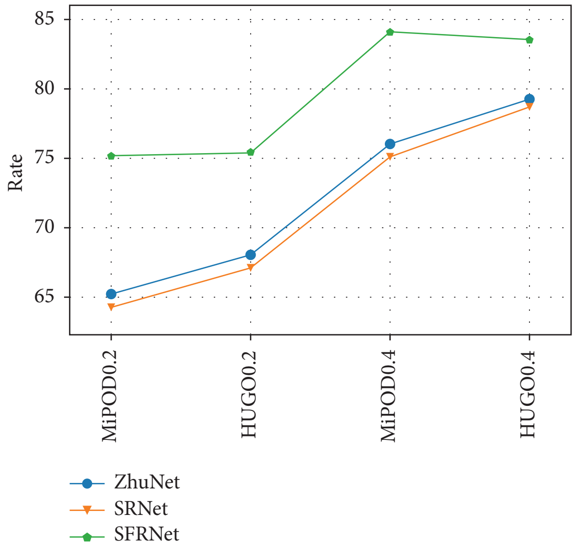


Figure 12: Comparison of the accuracy rate percentage of steganalysis algorithm among six steganalysis methods against two algorithms: MiPOD and HUGO at 0.2 bpp and 0.4 bpp.

Table 5: The computational complexity of the four networks.

Model	Parameter(10^4)	Train time(h)	Test time(s)
SFRNet	**150.5**	**2.91**	**9**
Zhu-net	287.1	8.65	33
SRNet	477.6	17.38	41
Yedroudj-Net	44.5	4.8	25

6. Conclusions

In this paper, a deep neural network with high accuracy and low time consumption is proposed for steganalysis. The feature extraction-fusion layers are used to extract features from original images and combine them into a feature matrix, which provides versatility for the steganographic analysis method. Furthermore, we use the SE block and the RepVgg block to construct the SFRNet, which significantly reduces the computational complexity while ensuring accuracy. At the same time, the SE block is used to extract channel correlation of the feature matrix. The experimental result show that the SFRNet has excellent steganalysis performance in the spatial domain. Especially, compared with the latest algorithm under low payload, the detection accuracy has been improved by 10%. In the future, we would extend our methods to the frequency domain.

Data Availability

The steganography algorithm code and BOSSBase ver. 1.01. data used to support the findings of this study are available at http://dde.binghamton.edu/download/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Science Foundation of China, under Grant no. U1831131, the Special Funds of Central Government of China for Guiding Local Science and Technology Development, under Grant no. [2018]4008, and the Science and Technology Planned Project of Guizhou Province, China, under Grant no. [2020]2Y013.

References

[1] M. Chaumont, "Deep learning in steganography and steganalysis," *Digital Media Steganography*, Academic Press, Cambridge, MA, US, pp. 321–349, 2020.

[2] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 920–935, 2011.

[3] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Proceedings of the International Workshop on Information Hiding*, June 2010.

[4] V. Holub and J. Fridrich, "Digital image steganography using universal distortion," in *Proceedings of the First ACM Workshop on Information Hiding and Multimedia Security*, Montpellier, France, June 2013.

[5] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS)*, December 2012.

[6] V. Sedighi, R Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 221–234, 2015.

[7] N. Jindal and B. Liu, *Review Spam Detection*, in *Proceedings of the 16th international conference on World Wide Web*, Alberta, Canada, 2007.

[8] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, 2012.

[9] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for steganalysis of digital images," in *Proceedings of the 2014 IEEE International Workshop on Information Forensics and Security (WIFS)*, December 2014.

[10] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, June 2015.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[12] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, "Xnor-net: imagenet classification using binary convolutional neural networks," in *Proceedings of the European Conference on Computer Vision*, September 2016.

[13] K. Huang, X. Liu, S. Fu, D. Guo, and M. Xu, "A lightweight privacy-preserving CNN feature extraction framework for mobile sensing," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, 2019.

[14] Z. Pan, F. Yuan, J. Lei, W. Li, N. Ling, and S. Kwong, "MIEGAN: mobile image enhancement via A multi-module cascade neural network," *IEEE Transactions on Multimedia*, p. 1, 2021.

[15] Z. Pan, W. Yu, J. Lei, N. Ling, and S. Kwong, "TSAN: synthesized view quality enhancement via two-stream attention network for 3D-HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, p. 1, 2021.

[16] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," in *Proceedings of SPIE–The International Society for Optical Engineering*, vol. 9409, San Francisco, CA, US, March 2015.

[17] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, Madison, WI, US, June 2010.

[18] G. Xu, H.-Z. Wu, and Y.-Q. Shi, "Structural design of convolutional neural networks for steganalysis," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 708–712, 2016.

[19] Y. Jian, J. Ni, and Y. Yang, "Deep learning hierarchical representations for image steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2545–2557, 2017.

[20] M. Boroumand, M Chen, and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 14, pp. 1181–1193, 2018.

[21] M. Yedroudj, F. Comby, and M. Chaumont, "Yedroudj-net: an efficient CNN for spatial steganalysis," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018.

[22] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," 2016, https://arxiv.org/abs/1602.07360.

[23] R. Zhang, F. Zhu, J. Liu, and G. Liu, "Depth-wise separable convolutions and multi-level pooling for an efficient spatial CNN-based steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1138–1150, 2019.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[25] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, "RepVGG: making VGG-style ConvNets great again," 2021, https://arxiv.org/abs/2101.03697.

[26] J. Hu, Li Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer*

Vision and Pattern Recognition, Salt Lake City, UT, USA, June 2018.

[27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556, Article ID 1556.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, June 2016.

[29] F. Liu, X. Zhou, X. Yan, Y. Lu, and S. Wang, "Image steganalysis via diverse filters and squeeze-and-excitation convolutional neural network," *Mathematics*, vol. 9, no. 2, p. 189, 2021.

[30] Z. Pan, X. Yi, Y. Zhang, B. Jeon, and S. Kwong, "Efficient in-loop filtering based on enhanced deep convolutional neural networks for HEVC," *IEEE Transactions on Image Processing*, vol. 29, pp. 5352–5366, 2020.

[31] L. Pibre, J. Pasquet, D. Ienco, and M. Chaumont, "Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source-mismatch," *Electronic Imaging*, vol. 2016, no. 8, pp. 1–11, 2016.

[32] D. P. Kingma and B. Jimmy, "Adam: a method for stochastic optimization," 2014, https://arxiv.org/abs/1412.6980.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 5542993, 9 pages
https://doi.org/10.1155/2021/5542993

Research Article

Research on Calibration Method of Binocular Vision System Based on Neural Network

Hao Zhu[1,2] **Mulan Wang**[2] **and Weiye Xu**[3]

[1]*School of Information and Communication, Nanjing Institute of Technology, Nanjing 211167, China*
[2]*Jiangsu Key Laboratory of Advanced Numerical Control Technology, Nanjing 211167, China*
[3]*Department of Informatics, University of Leicester, Leicester, LE1 7RH, UK*

Correspondence should be addressed to Hao Zhu; zhuhao@njit.edu.cn

Received 29 January 2021; Revised 15 April 2021; Accepted 14 June 2021; Published 21 June 2021

Academic Editor: Zhaoqing Pan

In binocular vision inspection system, the calibration of detection equipment is the basis to ensure the subsequent detection accuracy. The current calibration methods have the disadvantages of complex calculation, low precision, and poor operability. In order to solve the above problems, the calibration method of binocular camera, the correction method of lens distortion, and the calibration method of projector in the binocular vision system based on surface structured light are studied in this paper. For lens distortion correction, on the basis of analyzing the traditional correction methods, a distortion correction method based on radial basis function neural network is proposed. Using the excellent nonlinear mapping ability of RBF neural network, the distortion correction models of different lenses can be obtained quickly. It overcomes the defect that the traditional correction model cannot adjust adaptively with the type of lens. The experimental results show that the accuracy of the method can meet the requirements of system calibration.

1. Introduction

With the development of modern electronic technology, the application of 3D detection in the machining field is more and more mature. At present, the common 3D detection methods mainly include contact and noncontact. In the traditional reverse engineering, the common method of 3D object detection is the contact measurement technology represented by coordinate measuring machine (CMM). The advantage of this method is that it is easy to operate. However, it has a large error for the soft measured target. Besides, the cost of special large-scale CMM is also very high [1]. With the development of computer technology, the application of machine vision and noncontact measurement technology in the mechanical manufacturing system has gradually become a research hotspot.

Structured light detection is a representative method of noncontact measurement technology [2]. In the detection process, the projector projects structured light with specific rules to the target surface. The stripe of structured light changes with the depth of the target surface, resulting in distortion. Due to the different positions of the cameras on both sides of the projector, the distorted images captured by the cameras are also different. The distorted structured light image contains the depth information of the measured target surface and the relative position of the projector and camera. By analyzing and calculating the distortion characteristics, the target depth information can be obtained, and the target 3D coordinates can be achieved. In the process of calculation, in order to determine the relationship between the 3D geometric position of a point on the surface of a space object and its corresponding point in the image, it is necessary to establish the geometric model of the camera and projector. The parameters of these geometric models are the parameters of the camera and projector, including internal parameters, external parameters, and distortion parameters. In most cases, these parameters can only be obtained by experiment and calculation. This process of solving parameters is called system calibration. The accuracy of system calibration directly determines the accuracy of subsequent measurement and calculation [3–5]. Therefore, it is very important to study the high-precision and high-efficiency system calibration method for the 3D detection system.

2. Calibration Principle of Binocular Structured Light System

For a monocular vision system with only one projector and one camera, an equivalent camera can be created by rigid rotation and translation of the projector and camera. However, the premise of the transformation is that the internal parameters of the projector and camera are the same. In engineering practice, the above conditions are generally difficult to meet. Therefore, a binocular vision system composed of one projector and two cameras is usually used for 3D reconstruction [5, 6]. Generally speaking, the projector can be virtual as a pinhole imager, and the camera can be virtual as a linear camera. When the internal parameters of two cameras are the same and the optical center is in the same horizontal plane, the image height is the same. The corresponding points can be determined by searching for feature points at the same height. Therefore, when building a binocular vision system, two cameras of the same model can be selected and placed on a horizontal pan tilt. The projector is located between the two cameras. Two cameras are divided into two angles of view and simultaneously collect the projection image projected by the projector on the 3D target.

The calibration process of the structured light measurement system is the process of solving the functional relationship among the 3D coordinates of the measured point in space, the image information collected by the camera, and the structured light information. The parameters of this function include camera parameters, projector parameters, and the transformation relationship between camera coordinate system and world coordinate system. The calibration of the binocular structured light system includes camera calibration, relative position calculation of two cameras, camera distortion correction, projector calibration, and relative position calculation between the projector and camera [7].

3. Calibration of Camera and Projector

3.1. Principle of Camera Imaging. There are two cameras in the binocular vision detection system to obtain the target data. These cameras have their own positions and parameters. The final result of the reconstruction depends on the relationship between the spatial position of the target and the corresponding image points in the camera, that is, the set model and parameters of the camera. Therefore, it is necessary to model the camera and obtain the relevant parameters for 3D reconstruction of the target. Binocular vision detection is to calculate the camera coordinates corresponding to each point according to the coordinates of each point in the distorted structured light plane image obtained by the camera and then obtain the 3D world coordinates corresponding to each point on the target surface [8, 9].

As shown in Figure 1, let the upper left corner of the plane image coordinate be the coordinate origin O_0, and a known point in the image is $D(u, v)$. u and v are the number of pixels in the horizontal and vertical directions, respectively. The image coordinate system $(O_1 - xy)$ was established. Its origin position is (u_0, v_0). If the image coordinates corresponding to point $D(u, v)$ are (x, y), then the corresponding relationship between (u, v) and (x, y) is shown in the following equation:

$$\begin{cases} u = \left(\dfrac{x}{\mathrm{d}x}\right) + u_0, \\ v = \left(\dfrac{y}{\mathrm{d}y}\right) + v_0. \end{cases} \tag{1}$$

Its homogeneous form is shown in the following equation where dx and dy are the size of each pixel:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\mathrm{d}x} & 0 & u_0 \\ 0 & \dfrac{1}{\mathrm{d}y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{2}$$

The camera coordinate system $(O_2 - x_c y_c z_c)$ and world coordinate system $(O_3 - x_w y_w z_w)$ are established. According to the imaging principle of the pinhole camera, the relationship between the image coordinate system and the two is shown in Figure 2 where O_2 is the optical center of the camera and the line O_1O_2 is the focal length f of the camera.

As can be seen from Figure 2, the world coordinate system can be obtained by rotation and translation of the camera coordinate system. Let the rotation matrix be R and the translation matrix be t. The relationship between the world coordinate system and the camera coordinate system can be obtained as shown in the following equation:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ O^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{3}$$

According to the similar triangle principle, the relationship between the coordinates of point $D(x,\ y)$ on the plane image and its corresponding point $D(x_c, y_c, z_c)$ in the camera coordinate system is shown in the following equation:

$$\begin{cases} x = \dfrac{f x_c}{z_c}, \\ y = \dfrac{f y_c}{z_c}. \end{cases} \tag{4}$$

Equation (5) is the homogeneous form after sorting.

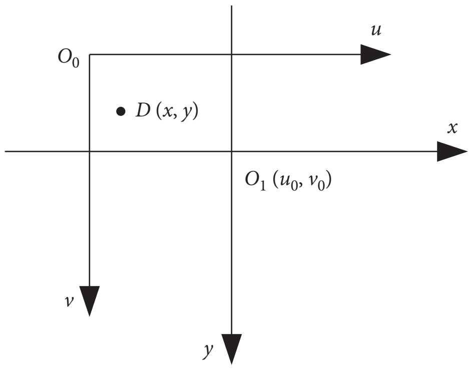


Figure 1: Definition of image coordinates.

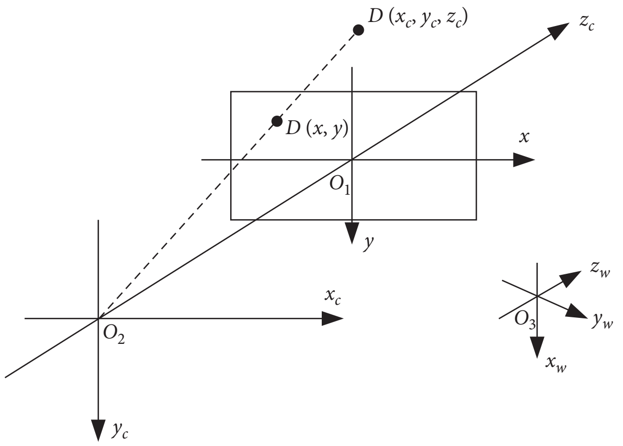


Figure 2: The relationship of three kinds of coordinates.

$$z_c = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}. \tag{5}$$

From equations (1), (3), and (5), the relationship between plane image coordinates and world coordinates can be obtained, as shown in the following equation:

$$\begin{aligned} z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= z_c \begin{bmatrix} \frac{1}{\mathrm{d}x} & 0 & u_0 \\ 0 & \frac{1}{\mathrm{d}y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{f}{\mathrm{d}x} & 0 & u_0 & 0 \\ 0 & \frac{f}{\mathrm{d}y} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ O^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_1 M_2 X_w = M X_w. \end{aligned} \tag{6}$$

Among them, $X = [x_w, y_w, z_w]^T$; M_1 is related to f, u_0, v_0 dx, and dy and determined by the internal structure of the camera, and is called internal parameter. M_2 is determined by the orientation of the camera relative to the world coordinate system, which is called external parameter. M is called the projection matrix.

3.2. Camera Calibration. Camera calibration is the process of obtaining the internal and external parameters of the camera. For the calibration plate, the 3D coordinates (x_w, y_w, z_w) of each feature point are known. The plane image coordinates of feature points are also known. Therefore, as long as there are enough characteristic points, the matrix M can be obtained and then M_1 and M_2 can be obtained. For each characteristic point on the calibration plate, the relationship is shown in the following equation:

$$z_{ci} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} = \begin{bmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{bmatrix}. \tag{7}$$

The system of equations can be obtained by eliminating z_{ci} as follows:

$$\begin{cases} u_i m_{34} = x_{mi} m_{11} + y_{wi} m_{12} + z_{wi} m_{13} + m_{14} - u_i x_{wi} m_{31} - u_i y_{wi} m_{22} - u_i z_{wi} m_{33}, \\ v_i m_{34} = x_{wi} m_{21} + y_{wi} m_{22} + z_{wi} m_{23} + m_{24} - v_i x_{wi} m_{31} - v_i y_{wi} m_{22} - v_i z_{wi} m_{33}. \end{cases} \tag{8}$$

It can be seen from equation (8) that each characteristic point can correspond to two independent equations. Therefore, the 12 unknowns in M matrix can be obtained from 12 equations obtained from 6 characteristic points. The solution method is the least square method. The more the number of feature points, the smaller the error. For n characteristic points, $2n$ equations are obtained as shown in the following equation:

$$\begin{bmatrix} x_{w1} & y_{w1} & z_{w1} & 1 & 0 & 0 & 0 & 0 & -u_1x_{w1} & -u_1y_{w1} & -u_1z_{w1} \\ 0 & 0 & 0 & 0 & x_{w1} & y_{w1} & z_{w1} & 1 & -v_1x_{w1} & -v_1y_{w1} & -v_1z_{w1} \\ & & & & & \cdots & \cdots & & & & \\ x_{wn} & y_{wn} & z_{wn} & 1 & 0 & 0 & 0 & 0 & -u_nx_{wn} & -u_ny_{wn} & -u_nz_{wn} \\ 0 & 0 & 0 & 0 & x_{wn} & y_{wn} & z_{wn} & 1 & -v_nx_{wn} & -v_ny_{wn} & -v_nz_{wn} \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \end{bmatrix} = \begin{bmatrix} u_1m_{34} \\ v_1m_{34} \\ \cdots \\ u_nm_{34} \\ v_nm_{34} \end{bmatrix}. \tag{9}$$

It can be seen from equation (6) that the multiplication of M matrix by any constant other than 0 does not affect the relationship between $[x_w, y_w, z_w]$ and $[u, v]$. Therefore, $m_{34} = 1$ can be specified in equation (9). At this time, the number of unknowns of M matrix is reduced to 11. Let these 11 unknowns be vector m, then equation (9) can be abbreviated to the following equation:

$$Km = u, \tag{10}$$

where K is a $2n \times 11$ matrix, m is an 11 dimensional unknown vector, and u is a $2n$-dimensional vector. When $2n > 11$, the solution of the equation obtained by the least square method is shown in the following equation:

$$m = \left(K^T K\right)^{-1} K^T u. \tag{11}$$

The larger the value of $2n$, the smaller the error.

Finding vector m is to get 11 unknowns in M matrix. The last unknown number m_{34} is solved as follows; equation (6) can be written as the following equation:

$$\begin{aligned} m_{34} \begin{bmatrix} m_1^T & m_{14} \\ m_2^T & m_{24} \\ m_3^T & 1 \end{bmatrix} &= M_1 M_2 = \begin{bmatrix} \alpha x & 0 & u_0 & 0 \\ 0 & \alpha y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_1^T & tx \\ r_2^T & ty \\ r_3^T & tz \\ 0^T & 1 \end{bmatrix} \\ &= \begin{bmatrix} \alpha x r_1^T + u_0 r_3^T & \alpha x tx + u_0 tz \\ \alpha y r_2^T + v_0 r_3^T & \alpha y ty + v_0 tz \\ r_3^T & tz \end{bmatrix}, \end{aligned} \tag{12}$$

where $\alpha = 1/\mathrm{d}x$. So, $m_{34}\ m_3^T = r_3^T$. Since r_3 is the third row of an orthogonal array of units, $|r_3| = 1$. From this, we can get

$$m_{34} = \frac{1}{|m_3|}. \tag{13}$$

After 12 unknowns of M matrix are obtained, each element in internal and external parameter matrix M_1 and M_2 can be obtained further.

3.3. Calculation of Relative Position between Two Cameras. In binocular vision camera calibration, in addition to calculating the internal and external parameters of each camera, it is also necessary to calculate the relative position between the two cameras. For the two cameras, there are

$$\begin{aligned} X_{c1} &= R_1 X_w + t_1, \\ X_{c2} &= R_2 X_w + t_2, \end{aligned} \tag{14}$$

where $X_w = [x_w, y_w, z_w 1]^T$.

After X_w is eliminated,

$$X_{c_1} = R_1 R_2^{-1} X_{c_2} + t_1 - R_2^{-1}. \tag{15}$$

Therefore, the relative position between two cameras can be represented by R and t as follows:

$$\begin{aligned} R &= R_1 R_2^{-1}, \\ t &= t_1 - R_2^{-1} t_2. \end{aligned} \tag{16}$$

3.4. Calibration of Projector. The projector can be regarded as a reverse working camera [10]. Therefore, the mathematical model of the projector can be represented by the pinhole camera model shown in equation (6). Although the mathematical model of the projector is the same as that of the camera, the projector cannot directly get the pixel coordinates of each feature point on the projector image plane. The solution is given in reference [11]; the projector projects the horizontal and vertical gray code fringes onto the calibration plate in the order of continuous subdivision. After the camera captures the image, the direct and indirect light components are calculated and the threshold segmentation is performed. Then, the gray code decoding algorithm is used to obtain the coordinates of each point on the image plane of the projector.

4. Lens Distortion Correction

4.1. Traditional Lens Distortion Correction Method. The ideal pinhole model is only an approximation of the real lens model. The actual camera and projector are different from the pinhole model because of the different lens structure and the processing error and assembly error in the production process. For ordinary lens, especially for wide-angle lens, lens distortion should be considered [12]. The most important influence on imaging is radial distortion. Let the radial distortion parameters be k_1 and k_2. Then, there are

$$\begin{cases} x\prime = x\left(1 + k_1 r^2\right), \\ y\prime = y\left(1 + k_2 r^2\right), \end{cases} \tag{17}$$

where $(x\prime, y\prime)$ is the image coordinate obtained from the single hole camera model and (x, y) is the actual image coordinate. Formula (17) only considers the radial distortion and ignores the high-order term. Eccentric distortion and thin prism distortion should be considered in real lens. Because of the difference of the optical model and the assembly error, they cannot be expressed by the same mathematical model. In reference [4], a lens distortion correction method based on BP neural network is proposed. However, BP network is slow in calculation and easy to fall into local optimum, so it cannot meet the real-time and accuracy requirements of 3D detection.

4.2. Lens Distortion Correction Based on RBF Network. RBF network is an efficient forward neural network. The relationship between input layer and hidden layer is nonlinear. There is a linear weighted relationship between the hidden layer and the output layer [13]. This structure avoids the tedious calculation of BP network. It has not only good nonlinear approximation ability but also fast computing ability. It is especially suitable for nonlinear prediction from n-dimensional space to m-dimensional space.

The RBF network structure of lens distortion correction is shown in Figure 3. The input signal (x, y) is the coordinates of the real shot image. The output signal $(x\prime, y\prime)$ is the actual image coordinates, which can be defined as the coordinates of feature points on the calibration board. The number of nodes in the hidden layer is the number of samples n.

The input vector of the system is $d = [x, y]^T$. The output vector is $d\prime = [x\prime, y\prime]^T$. The weight matrix w is $2 \times n$ matrix. The element w_{ij} is the weight between the i-th node in the hidden layer and the j-th node in the output layer. The radial basis function $\varphi\ (d^i, d^p)$ is Gaussian kernel function where d^i is the i-th input vector and d^p is the center point vector.

$$\Phi\left(d^i, d^p\right) = \exp\left(-\left(\frac{1}{2\sigma^2}\right)\left\|d^i - d^p\right\|^2\right). \tag{18}$$

The system provides n characteristic point samples on the calibration board. According to the network structure, the system output is

$$d_j' = \sum_{i=1}^{n} w_{ij} \exp\left(-\frac{1}{2\sigma^2}\left\|d^i - d^p\right\|^2\right) \quad , j = 1, 2. \tag{19}$$

In order to avoid each radial basis function being too sharp or too flat, the definition of the expansion constant of the radial basis function is shown in the following equation:

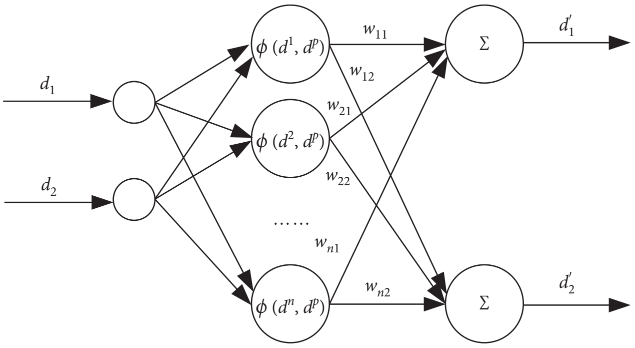


Figure 3: The RBF network structure of lens distortion correction.

$$\sigma = \frac{d_{\max}}{\sqrt{2n}}, \tag{20}$$

where $d_{\max}$ is the maximum distance among samples and n is the number of samples.

Learning of system is divided into two stages. The first stage is unsupervised learning. The concrete work is to solve the center and variance in the hidden layer. The second stage is supervised learning. The specific work is to solve the weight matrix from implicit layer to output layer. The adjustment of weight can be realized by the least mean square error. The weight adjustment formula is

$$\Delta w_{ij} = \eta\left(d_j' - w_j^T \Phi\right)\varphi_i. \tag{21}$$

Among them, d_j' is the j-th expected value.

$$\begin{aligned} w_j &= \left[\, w_{1j} \;\; w_{2j} \;\; \cdots \;\; w_{nj} \,\right]^T, \\ \Phi &= \left[\, \varphi_1 \;\; \varphi_2 \;\; \cdots \;\; \varphi_n \,\right]^T. \end{aligned} \tag{22}$$

5. Experiment and Analysis

5.1. Experimental Process. In this paper, the binocular detection system shown in Figure 4 is used to verify the above algorithm. The camera pixel is 1.3 million, and the measurement format is 200 mm × 150 mm. The nominal scanning accuracy is 0.01 mm. The calibration board used in the experiment is shown in Figure 5. There are $11 \times 13 = 169$ regularly arranged feature points on the calibration board, including 17 locating points.

In the calibration experiment, the position of the calibration plate is fixed first. The projector projects the gray code structured light to the calibration plate as shown in Figure 6. The camera takes pictures. The attitude of calibration board is changed, and the above work is repeated. The calibration board is placed in 8 different positions in the measurement space, as shown in Figure 7. Each camera obtains 8 images for system calibration. In calibration, firstly, the edge of the image is extracted at pixel level to identify the mark points and fit the center point and number the marker points according to the locating points. Then, the 3D coordinates of the positioning point are reconstructed by using the location of the positioning point in the first two pictures. Next, the 3D coordinates of the landmarks are reconstructed with the rest of the images.

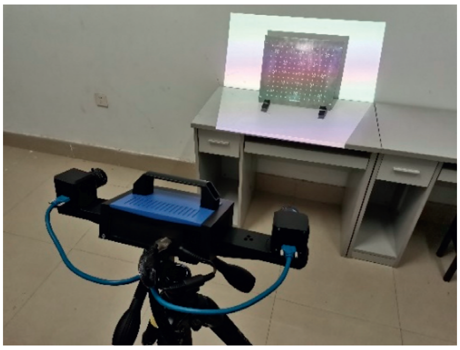


Figure 4: Binocular vision system.

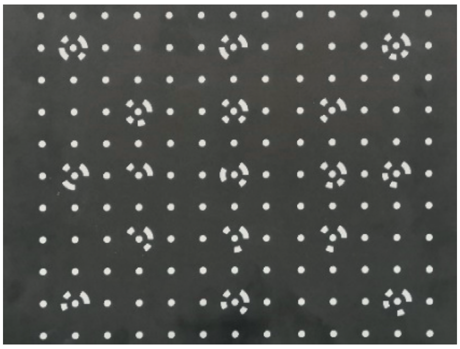
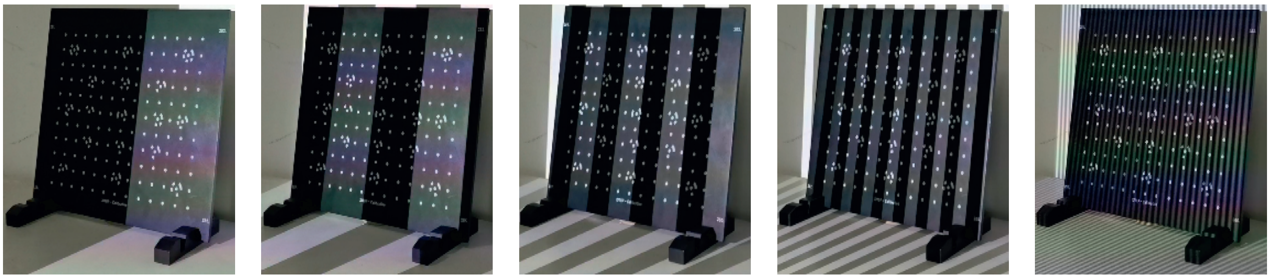
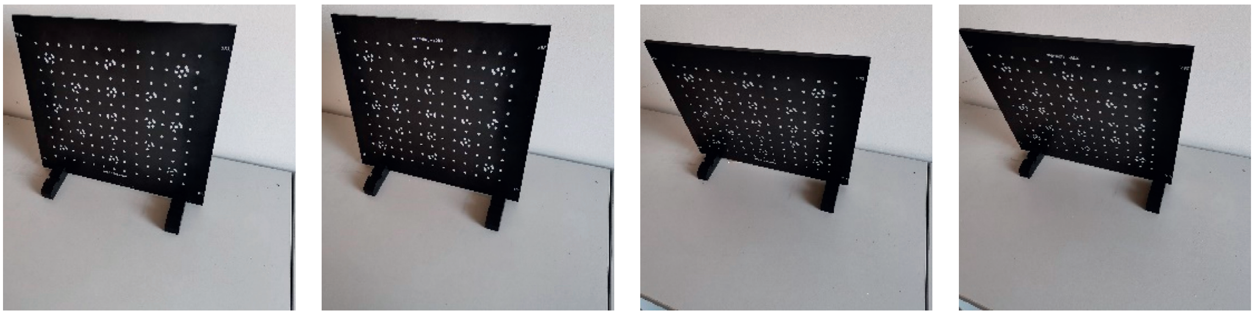


Figure 5: Calibration board.



(a) (b) (c) (d) (e)

Figure 6: Projected gray code structured light.



(a) (b) (c) (d)

Figure 7: Continued.

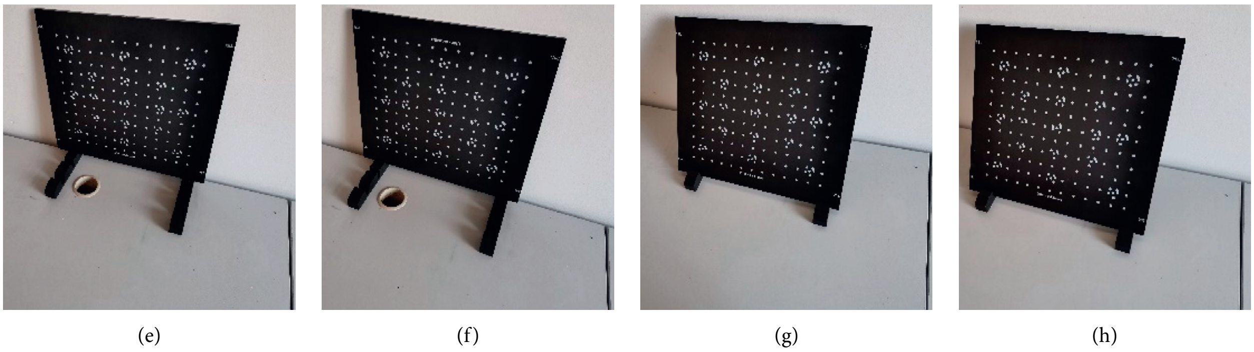

(e) (f) (g) (h)

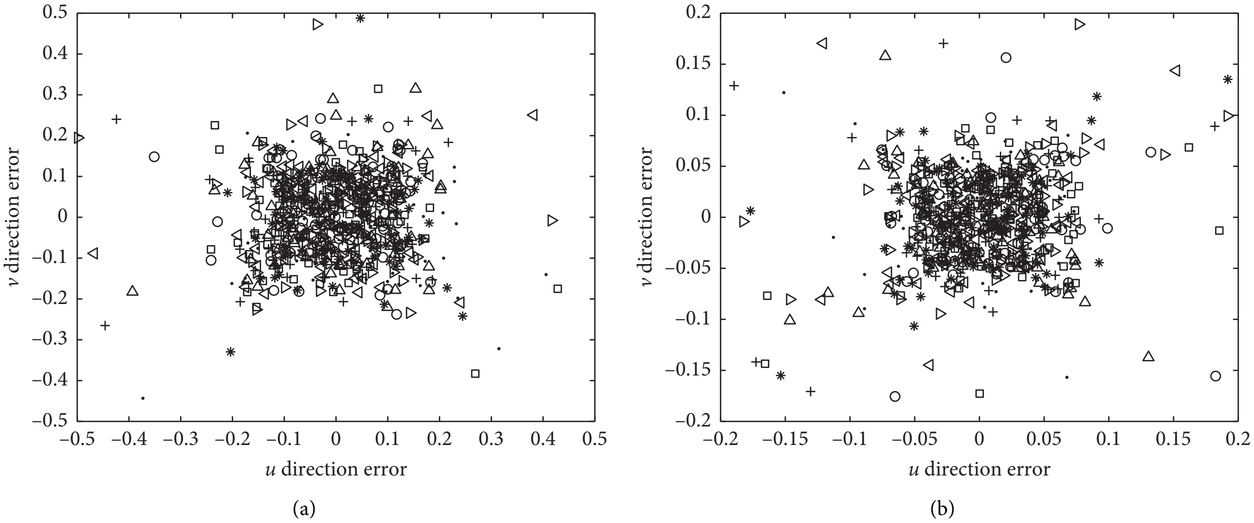
Figure 7: Eight attitudes of calibration board.

Table 1: Internal parameters of left and right cameras.

Camera	Internal parameters			
	ax	αy	u_0	v_0
Left	2785.35	2678.67	229.21	145.54
Right	2820.91	2732.28	257.82	107.85

Table 2: External parameters of left and right cameras.

Camera	External parameters	
	R	t
Left	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$
Right	$\begin{bmatrix} 0.9006 & -0.02775 & -0.4338 \\ 0.01377 & 0.9992 & -0.3533 \\ 0.4345 & 0.02584 & 0.9003 \end{bmatrix}$	$\begin{bmatrix} -279.1 \\ -15.82 \\ -127.8 \end{bmatrix}$



(a) (b)

Figure 8: Reprojection error.

5.2. Data Analysis. After obtaining the 3D coordinates of the marker points, the internal and external parameters of the camera can be calculated according to the method mentioned above, as shown in Tables 1 and 2.

In the experiment, the reprojection method is used to verify the accuracy of the calibration data. According to the parameters obtained from the calibration, the positioning points are reprojected to the image plane of the camera and compared with the actual image points. The traditional method and the method proposed in this paper are used to calibrate, respectively. The calculated error is shown in Figure 8. On the left is the result of the traditional calibration method. On the right is the calibration result of the method.

Table 3: Residuals corresponding to eight postures of the traditional algorithm.

Attitude	Average error u	Average error v	Standard error u	Standard error v
1	0.0129	0.0146	0.0992	0.1187
2	−0.0107	0.0090	0.1085	0.1007
3	−0.0029	0.0063	0.1118	0.1191
4	0.0145	−0.0072	0.1059	0.1109
5	−0.0148	0.0168	0.1104	0.1133
6	0.0081	−0.0034	0.1162	0.1109
7	0.0151	0.0018	0.1218	0.1030
8	0.0013	−0.0198	0.1175	0.1228

Table 4: Residuals corresponding to eight postures of the algorithm in this paper.

Attitude	Average error u	Average error v	Standard error u	Standard error v
1	0.0083	0.0027	0.0462	0.0485
2	−0.0087	−0.0010	0.0426	0.0494
3	0.0067	0.0076	0.0463	0.0495
4	0.0022	0.0035	0.0488	0.0456
5	−0.0030	0.0012	0.0443	0.0516
6	−0.0031	−0.0011	0.0492	0.0510
7	0.0021	−0.0097	0.0501	0.0478
8	−0.0096	−0.0089	0.0501	0.0495

When the two algorithms are used, the residual data corresponding to the eight attitudes of the calibration board are shown in Tables 3 and 4.

6. Conclusion

In this paper, the calibration algorithm of the camera and projector in the binocular vision structured light detection system is introduced. The ideal image and the actual image coordinates are taken as the input and output system, and the image distortion correction system based on RBF neural network is constructed. The actual camera distortion correction calculation is completed by using the good nonlinear fitting ability of neural network. Experimental results show that the algorithm can overcome the shortcomings of traditional methods, and the detection results can meet the actual needs.

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 regarding the publication of this paper.

Acknowledgments

This work was supported by the Creation Foundation Project of Nanjing Institute of Technology (Grant No. CXY201933).

References

[1] L. Y. Fu, D. Zhang, and Q. L. Ye, "Recurrent thrifty attention network for remote sensing scene recognition," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–12, 2020.

[2] C. F. Jiang, L. Beatrice, and Z. Song, "Three-dimensional shape measurement using a structured light system with dual projectors," *Applied Optics*, vol. 57, no. 14, p. 3983, 2018.

[3] Z. Wang, Z. Wu, X. Zhen, R. Yang, J. Xi, and X. Chen, "A two-step calibration method of a large FOV binocular stereovision sensor for onsite measurement," *Measurement*, vol. 62, pp. 15–24, 2015.

[4] D. Zhang, G. Zhang, and L. Li, "Calibration of a six-axis parallel manipulator based on BP neural network," *Industrial Robot: The International Journal of Robotics Research and Application*, vol. 46, no. 5, pp. 692–698, 2019.

[5] W. G. Li, H. Li, and H. Zhang, "Light plane calibration and accuracy analysis for multi-line structured light vision measurement system," *Optik*, vol. 207, Article ID 163882, 2019.

[6] X. Chen, F. Zhou, and T. Xue, "Omnidirectional field of view structured light calibration method for catadioptric vision system," *Measurement*, vol. 148, Article ID 106914, 2019.

[7] J. E. Ha, "Calibration of structured light vision system using multiple vertical planes," *Journal of Electrical Engineering and Technology*, vol. 13, no. 1, pp. 438–444, 2018.

[8] X. Chen, R. Fan, J. Wu et al., "Fourier-transform-based two-stage camera calibration method with simple periodical pattern," *Optics and Lasers in Engineering*, vol. 133, Article ID 106121, 2020.

[9] J. Jiang, L. Zeng, B. Chen, Y. Lu, and W. Xiong, "An accurate and flexible technique for camera calibration," *Computing*, vol. 101, no. 4, pp. 1971–1988, 2019.

[10] S. Yang, M. Liu, J. Song et al., "Projector calibration method based on stereo vision system," *Optical Review*, vol. 24, no. 5, pp. 1–7, 2017.

[11] S. K. Nayer, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 935–944, 2006.

[12] M. Zhang, F. L. Wu, L. X. Jin, G. N. Li, S. Han, and Y. Zhang, "Correction optimization of lens radial distortion with

bending measurement function," *Transactions of Tianjin University*, vol. 22, no. 98, pp. 94–99, 2016.

[13] Q. Ye, H. Zhao, Z. Li et al., "L1-norm distance minimization-based fast robust twin support vector *k*-plane clustering," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4494–4503, 2018.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 9968652, 13 pages
https://doi.org/10.1155/2021/9968652

Research Article

Quantitative Weighted Visual Cryptographic (k, m, n) Method

Yewen Wu,[1] **Shi Zeng**,[2,3] **Bin Wu**,[4] **Bin Yang**,[5] **and Xianyi Chen**[3]

[1] *Institute of Space Weather, Nanjing University of Information Science and Technology, Nanjing, China*
[2] *College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China*
[3] *School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China*
[4] *School of Computer and Big Data Science, Jiujiang University, Jiujiang, Jiangxi 332005, China*
[5] *School of Design, Jiangnan University, Wuxi, China*

Correspondence should be addressed to Xianyi Chen; 0204622@163.com

Received 19 March 2021; Accepted 11 May 2021; Published 28 May 2021

Academic Editor: Md. Hasanul Kabir

The weighted visual cryptographic scheme (WVCS) is a secret sharing technology, where weights are assigned to each shadow (participant) according to its importance. Among WVCS, the random grid-based WVCS (RGWVCS) is a frequently visited subject. It considers the premise of equality of all participants, without taking into account the existence of privileged people in reality. To address this problem of RGWVCS, this paper designs a new model, named as (k, m, n)-RGWVCS (where $m < k < n$), in which the secret is encrypted into n shares and sent to k participants. In the recovery end, the secret could be reconstructed by minimum m shares when the privileged join in; otherwise, k shares are needed. The experimental results show that our method has the advantage of no pixel expansion and no codebook design by means of random grid. Moreover, the contrast of our model increased by 32.85% on average compared with that of other WVCS.

1. Introduction

Visual cryptography scheme (VCS) sometimes also called visual secret sharing was formally proposed by Naor and Shamir at Eurocrypt'94 [1], in which they encrypt a binary secret image into multiple shares, and any qualified share combination can recover the hidden secret in the secret image, while any unqualified share combination cannot obtain any content related to the secret information. Here, the set of combinations that meet the conditions and combinations that do not meet the conditions are called a qualified set and a forbidden set, respectively, and the access structure consists of the pair of the qualified and forbidden sets. The (k, n)-threshold scheme [2, 3] is a typical case of VCS, in which a binary secret image is encrypted into n shares. Any k or more shares, when superimposed, can reveal the secret image to the naked eye without any computations, while any $k-1$ or less shares leak no information about the secret in an information-theoretic sense even with high-performance computing devices and technologies.

As can be seen from the (k, n)-VCS, it encrypts a secret image into visual secret shares so that humans can reconstruct the secret directly with the naked eye without any cryptographic expertise or numerical computation by superposing a qualified visual sharing combination printed on a transparency.

Due to the high security and concealment of VCS, it is essential in the transmission of highly confidential information in a completely hostile channel. More precisely, the problem here is to transmit highly confidential information or authentication information through one or more insecure channels which are under full control of the adversary. Although the emergence of identity authentication technology has made a momentous contribution to keep the identity security, the filch or illegal decoding of smartcard and certificate will pose a high safety hazard. Therefore, the visual cryptography technology is combined with the identity authentication to store the shares generated when the identity information is encrypted in the authentication device and the database, respectively. The authentication of the user needs to be decrypted through the superposition of the stored shares, while the attacker cannot obtain any valuable information of the secret image from the individual share through existing analytical methods, and it is impossible to decrypt by forging user shares. In uncontrolled channels, this problem is difficult to solve, and visual authentication [4] has been the main appropriate security solution up to now.

The loss of contrast and resolution of the secret image enables the VC to achieve security confidentiality, so in general, the quality of the secret recovered image is lower than that of the original secret image. With the gradual maturity of VC technology, efficiency and security have become the focus [5–7]. Scholars have attempted to enhance the contrast and resolution (visibility) of the reconstructed images. To alleviate the pixel expansion in the generated shares, several VCSs without pixel expansion were proposed previously. The VCS based on random grids (RGs) [8, 9] are characterized by no codebook design and no pixel expansion. The follow-up research work of RGVCS mainly focuses on a (k, n) threshold [10] and improves image recovered quality [11]. Many other schemes were proposed to realize diversified visual sharing such as halftone secret images [12], grayscale secret images [13], and color secret images [14]. Moreover, some VC extension methods have also been proposed [11, 15, 16]. In terms of the functionality and security, how to share multiple secrets at one time [17–20] and prevent participants from cheating [21, 22] has become particularly significant in the past few years.

1.1. Related Works. Traditional VCSs mostly devote to optimizing the quality of secret recovered images [23], multilayer secret image encryption [17, 19], application of grayscale images and color images [8], and so on. Kannojia and Kumar [24] proposed a XOR-based $(n, n)\text{-}VCS_{\mathrm{XOR}}$ visual secret sharing scheme using pixel vectorization. The proposed scheme uses implicit codebook and pixel vectorization technology to directly encode the grayscale secret image without converting to a halftone image, which solves the problems of pixel expansion, contrast loss, clear codebook requirements, and restrictions on the number of participants. Shivani et al. [25–27] proposed a novel PVSS scheme based on an effective preprocessing method and a basic matrix creation algorithm, which has four or more amounts of space efficiency and meaningful shares. Many avoidable encryption restrictions are solved, and the human visual system can easily decrypt without any cryptographic calculations. Besides, Shivani S also proposed a novel multiple secret sharing scheme with unexpanded shares and meaningful shares, which can protect two secret images at once. The experimental results show that all meaningful shares meet the contrast and safety conditions. The secret image can be easily decoded by the human eye without any calculation [28]. However, traditional VCSs can only handle shares of the same weight and ignore the difference in the importance of shares. Therefore, this work considers a weighted VCS with privileged based on random grids, in which secret sharing participant acquires a share with different weights, respectively, according to the importance, and all the eligible participant combinations can recover the secret through the overlay of secret shares. So far, the following VCSs considering the weight difference have been proposed: the privilege-based visual secret sharing model (PVSSM) [29], (2, n)-PVCS [30], random grid-based progressive visual secret sharing scheme with adaptive priority (RGPVSS) [31], and weighted (k, n)-threshold visual cryptography method [32]. In most existing VCSs, each share has the same secret recovery capability. However, in the reality, some participants may enjoy privileges because of their importance or status, e.g., president of company, government official, and so on. Secret image recovery using traditional VCSs cannot highlight their privileges. Therefore, Hou et al. [29] proposed a novel secret sharing mechanism PVSSM that combines the concepts of progressiveness and privilege. In PVSSM, n participants with different privileges share a secret image. Set a privilege level (PL) $PL_i\,(1 \le i \le n)$ for each participant and create corresponding shares to participants based on the PL. The quality of the recovered secret image depends on the importance of all participants. Therefore, the higher the privilege level, the more secret information contained in the sharing, and the higher ability to recover the secret image. In this scheme, each share assigned to the participant has the same size as the secret image. PVSSM is capable of encrypting binary and halftone images and can clearly recover the secret information hidden in secret images. Compared with other relevant VCSs, the recovery image in this scheme has a superior contrast $(n-2)/(n-1)$. Although the method proposed by Hou et al. assigns different weights to different shares, it is only for the (2, n) threshold. Different weights of shares lead to different average optical transmissions, resulting in the ability to visually distinguish shares, i.e., weight leakage.

To address the issue of weight leakage, Yang et al. [30] proposed a new (2, n)-PVCS to solve the critical problems. This solution solves the problem that Hou et al.'s scheme is not a general solution to implement all PVCSs with arbitrary privilege level, i.e., it is the general solution of all (2, n)-PVCS. The scheme also overcomes the problem of shadow whiteness inequality and provides the equal whiteness for each shadow. Unfortunately, the (2, n)-PVCS proposed by Yang et al. requires codebook design. Then, Fan et al. [31] described a random grid-based progressive visual secret sharing scheme (RGPVSS). In this scheme, the priority weighting of each share can be adjusted, and the secret will be recovered gradually with the increase in the number of shares involved in secret recovery. Therefore, as the number of shares superimposed increases, the hidden information in the secret image will be recovered more and vice versa. As a result, each share with different priorities will be generated in this scheme. During secret recovery, the priority weight of the share superimposed determines the extent to which the secret image can be recovered. Each share generated in this scheme has no pixel extension and has the same size as the original secret image. The average light transmission of each share is 1/2, thus preventing the sharing of different priority weights from being discernible. Crucially, for this scheme, codebook design is not indispensable. Tu et al. [32] proposed a weighted (k, n)-threshold nonexpanded visual secret sharing scheme. The secret image can be recovered when the sum of the participant's weights is not less than the threshold k. Unfortunately, the quality of the secret images recovered by the above scheme is not satisfactory. Yan et al. [33] proposed a weighted visual secret sharing scheme to improve the quality of secret restored images. Compared with other schemes, this scheme is mainly different from the generating mode of last n-k bits. Yan et al. [34] also proposed a VCS which the value of the last n-k bits is equal to the first k bits. The contrast of the secret image recovered by this scheme is significantly improved. To the best of our knowledge, this scheme achieves the most satisfactory contrast in RG-based VCSs. However, although the contrast of the secret restored image is improved, the original secret image cannot be restored losslessly. In the military and medical fields, lossless recovery is essential for image transmission and storage [36]. The study of VCSs with lossless recovery capability has very profound implications. Therefore, this paper proposes a weighted VCS with privileged based on random grids, which not only overcomes a series of problems mentioned above but also achieves a lossless recovery of the secret image.

1.2. Our Contributions

1.2. Our Contributions. In this paper, we propose a novel RGWVC method for a (k, m, n) ($m < k$) threshold. Compared with the basic WVCSs, the proposed RGWVC considers the possible social stratum differences. In combination with the actual situation, the proposed scheme empowers the secret distributor to designate a privileged role among all the participants, in which the privileged has a greater power for secret recovery than the ordinary participant. Assuming that the privileged joins in the secret recovery, only m shares are required, otherwise at least k shares. More precisely, when privileged share participates in secret recovery, a minimum of m participants is required. The innovation of this scheme compared with the preliminary scheme is that it not only assigns different weights to the participants based on their importance but also gives privilege to the significant shareholder in the secret sharing process.

Therefore, the proposed scheme can be better applied to practical models such as management pyramid, and the experimental results and the comparison with the existing schemes are shown that the scheme we designed has several highlights as follows: (1) no codebook design, (2) no pixel expansion, (3) no weight leakage, (4) a weighted VCS, (5) (k, m, n) threshold, considering the existence of privileged role, is more applicable to the reality, and (6) improved image quality compared with relevant WVCSs.

1.3. Organization. The rest of this paper is organized as follows. Section 2 reviews preliminaries on the basic WRGVCS. Section 3 presents a generic construction of the designed (k, m, n)-RGWVCS and discusses details and feasibility of the scheme. In Section 4, the experimental results are presented and analysed in the form of pictures and data. Besides, the scheme we designed is compared with several relevant VCSs from the aspects of contrast, pixel expansion, codebook design, and so on. Finally, conclusions and research issues are given in Section 5.

2. Preliminaries

In this section, we provide some preliminaries including background, definitions, notations, and conditions that will be used later. For more details about information theory and the definition of secret sharing, see, e.g., [36–38].

2.1. Weighted Random Grid-Based VCS (WRGVCS) for a (k, n) Threshold. The weighted random grid-based visual cryptographic scheme (WRGVCS) is the basis of the proposed scheme. This method can assign weights to each participant of different importance so that each share has different ability to restore the secret image. Our visual cryptography scheme inherits the advantages of WRGVCS and considers the possible privileged role in the reality, which increases the practicality of the scheme and makes it more comprehensive.

Suppose a binary secret image S with a size of $H \times W$, the pixel value denotes $S(s,t)\,(1 \leq s \leq H, 1 \leq t \leq W)$, and the weighted values are $\omega = \{\omega_1, \omega_2, \omega_3 \cdots \omega_n\}$ corresponding to participants $\{1, 2, \ldots, n\}$, where $\omega_1 + \omega_2 + \cdots + \omega_n = 1$. So, for each $s = 1, 2, 3 \ldots H; t = 1, 2, 3 \ldots W$, the WRGVCS can be summarized as follows:

(1) Randomly select k order numbers from $\{1, 2, \ldots, n\}$ according to $\omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, denoted as $\{i_1, i_2, \ldots, i_k\}$, where the probability of selecting the i-th order number is ω_i, for $i = 1, 2, \ldots, n$.

(2) Generate $b_1, b_2, \cdots, b_n \in \{0, 1\}$ randomly. If $S(s,t) \neq b_{i_1} \oplus b_{i_2} \oplus \cdots \oplus b_{i_k}$, select $\underline{p} \in \{1, 2, \ldots, k\}$ randomly and update b_{i_p} with $\overline{b_{i_p}}$, such that $S(s,t) = b_{i_1} \oplus b_{i_2} \oplus \cdots \oplus b_{i_k}$.

(3) Set $b_{j_1} = b_{i_{x_1}}, b_{j_2} = b_{i_{x_2}}, \cdots, b_{j_{n-k}} = b_{i_{x_{n-k}}}$, where $\{j_1, j_2, \cdots, j_{n-k}\}$ is the difference set of $\{1, 2, \cdots, n\}$ and $\{i_1, i_2, \cdots, i_k\}$, and x_q is randomly picked up from $\{1, 2, \cdots, k\}$, for $q = 1, 2, \cdots, n-k$.

(4) Arrange $b_1, b_2, \cdots, b_n$ in order in the position (s, t) of participant's matrix $M_1, M_2, \cdots, M_n$.

(5) Repeat the Steps 1–4 to fill in the above n matrix $M_1, M_2, \cdots, M_n$, which is the n shares.

2.2. Monte Carlo Method. Monte Carlo method is a numerical simulation method. It takes probabilistic phenomena as its research objective. In equipment effectiveness evaluation, it is often used to determine the efficiency index with random factors, such as the probability of discovery, the probability of hit, and the average number of damaged targets.

The simulation processes are described as follows. (1) Construct a simple and feasible stochastic or probabilistic model to describe the problem. The solution of the proposed problem is bound to some features of the random variables in the model (such as probability, mean value, and variance); (2) Generate a sufficient amount of random numbers according to the different distribution of each random variable in the stochastic or probabilistic model; (3) Design a sampling method suitable for the probabilistic model and random variable distribution; and (4)

According to the established model, the simulation test and calculation are carried out to obtain the random solution of the problem.

2.3. Indispensable Notations and Definitions. In this section, we will introduce some notations and definitions to prepare for the further work. Later, the symbols ⊗ and ⊕ represent the Boolean XOR and OR, while $\overline{b}$ represents a bitwise complementary operation of any binary bit b. In the visual secret sharing scheme, we generate n secret shares to hide the binary secret image S with size $H \times W$, defined as $\{M_1, M_2, \cdots, M_n\}$. The corresponding secret recovery image S′ is reconstructed from any $p\,(k \leq p \leq n, p \in Z^+)$ shadows $\{M_{i1}, M_{i2}, \cdots, M_{ip}\}$ by a superposing operation. The white and black areas of the secret image S are denoted as CS0 and CS1, respectively, where $\mathrm{CS0} = \{(\mathrm{s},\ t)\ |\mathrm{S}(\mathrm{s},\ t) = 0,\ 1 \leq s \leq H, 1 \leq t \leq W\}$ and $\mathrm{CS1} = \{(\mathrm{s},\ t)\ |\mathrm{S}(\mathrm{s},\ t) = 1,\ 1 \leq s \leq H,\ 1 \leq t \leq W\}$. For any pixel in S, $P(s = 0)$ represents that the probability of pixel color is transparent or white (0), whereas $P(s = 1)$ represents that the probability of pixel color is opaque or black (1). In addition, $P(S = 0) = 1 - P(S = 1) = 1 - 1/HW \sum_{i=1}^{H} \sum_{j=1}^{W} S(s,t)$. Based on above notations, we will give the definition of contrast.

Definition 1 (contrast). The contrast can describe the quality of the recovered secret image, denoted as α, as follows [34]:

$$\alpha = \frac{P_0 - P_1}{1 + P_1} = \frac{P(S'[CS0] = 0) - P(S'[CS1] = 0)}{1 + P(S'[CS1] = 0)}, \tag{1}$$

where P_0 and P_1 denote the correctly and incorrectly revealed probabilities for the white and the black areas of S′, respectively, $\alpha \in [-0.5, 1]$, which is proportional to the quality of the secret restored image after superposition. In secret recovery, $\alpha = 0$ means S′ has nothing to do with S while $\alpha = 1$ indicates S′ is the same as S, a.k.a., S is lossless.

Figure 1 shows the change of contrast under different values of P_0 and P_1 in formula (1) by using the heatmap. It can be seen from the figure that the contrast is positively correlated with P_0 and negatively correlated with P_1. When $P_0 = 1$ and $P_1 = 0$, the contrast value is 1, indicating that the image is lossless. Therefore, contrast is a very convincing indicator used to evaluate the quality of S′ compared with S, and it is adopted to determine whether the naked eye can recognize the secret in the recovery image in this paper. For the different contrast ranges for different clarity, see [39].

2.4. Feasibility and Safety Conditions. In this part, we will introduce two necessary conditions to verify the feasibility and security of the proposed scheme, in terms of secure encryption and visual decryption.

In the visual cryptography scheme, the ultimate goal is to realize the visualization of secret under the condition of meeting the requirements of the scheme. Therefore, a qualified visual cryptography scheme must meet the following condition in visual recognition:

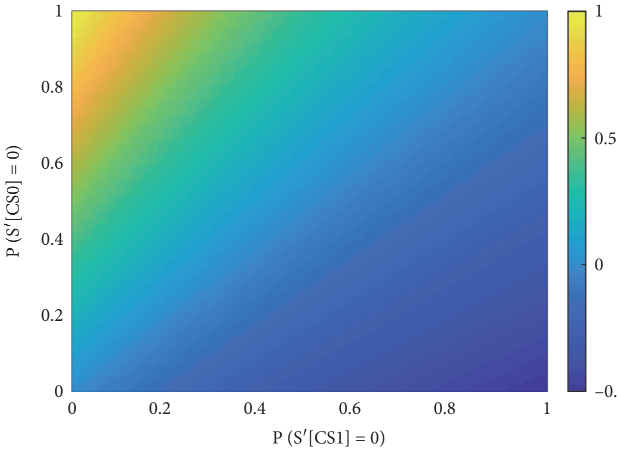


Figure 1: The heatmap of the contrast α under different P_0 and P_1.

Condition 1 (visually recognizable). The recovery image S′ is visually recognizable ($\alpha > 0$) when $p \geq k$, a.k.a., S′ can visually recognize the same secret as S with sufficient shadows [1, 32].

In the visual cryptography scheme, the security of encrypted secret is of capital importance. Therefore, a secure visual cryptography scheme must meet the following condition in secret recovery.

Condition 2 (security). The secret recovery image S′ is visually unrecognizable ($\alpha = 0$) when $p < k$, a.k.a., S′ will not give away any information about the secret if there is not enough shadow overlay during the recovery process.

3. The Proposed (*k*, *m*, *n*)-RGWVCS

Based on the basic WRGVCS described in the previous section, we designed a new (k, m, n)-RGWVCS $(m < k)$ that takes privileged role into account, in which the privileged people have a better decryption ability than nonprivileged. In the following section, we will first introduce the construction of (k, m, n)-RGWVCS, then analyse the advancement of the scheme, demonstrate its safety and feasibility, and finally analyse the application of the scheme in practice.

3.1. Construction of (k, m, n)-RGWVCS. The proposed (k, m, n)-RGWVCS $(1 \leq m \leq k < n)$, considering the privileged in secret recovery, can be implemented by the combination of binary matrices $\{M_1, \ \cdots, M_n\}$ and weight of n participants $\{\omega_1, \omega_2, \cdots, \omega_n\}$ (where $\omega_1 + \omega_2 + \cdots + \omega_n = 1$). There are n participants involved in the VCS in all, and the addition of the privileged person allows the secret recovery to be completed with only m participants, rather than at least k participants in the conventional scheme. In this section, we will give a detailed description of the proposed algorithm and analyse its feasibility. Construction 1 shows the proposed (k, m, n)-RGWVCS, which satisfies Condition 1 and Condition 2.

Construction 1. For every pixel, assign the n sharing matrixes using the Monte Carlo method with $\{0, 1\}$ randomly. If the privileged is not considered, select k or more matrixes as the candidate of the secret recovery; otherwise, select m or more shares matrixes. Then, flip an element randomly if their superposition of these shares is not equal the secret pixel value. After the above operations, we will get the n secret sharing matrices. The detailed steps are listed as Algorithm 1.

As can be seen from Figure 2, it provides a visual overview of the encryption and decryption secret image into n subpixels and generates n secret sharing matrices, while the right half discusses the different cases of pixel stacking and recovery. The secret sharing matrix and the secret image have the same attributes except the different distribution of pixel values. Different quality secret images can be recovered when the secret sharing matrix meets the quantity requirement.

3.2. Encryption of Single Pixel.

3.2. Encryption of Single Pixel. After the overall grasp of the shadow encryption architecture in the previous section, this section will analyse the details of the right half of Figure 2. This part is mainly responsible for superimposing n subpixels generated by each single pixel of S. In the ordinary VCS schemes, both the original secret and the generated share are data in numerical form, which are finally calculated and decrypted by the computer. In VCS schemes, however, both the original secret and the secret recovered by superposition are visible, i.e., the naked eye can decode directly. All black and white pixels in S are encrypted into corresponding sets of black and white subpixels in secret shares. Therefore, each subpixel set is able to be denoted as a set of matrices $C^b = (c_{ij}^b)$ with $b \in \{0, 1\}$, where $b = 1$ represents black pixel and $b = 0$ otherwise and $c_{ij}^b = 1$ represents a black j-th strategy in the i-th share and $c_{ij}^b = 0$ otherwise.

For the sake of clarity, $(2, 2)$-threshold VCS is taken as an illustrative example where S is encrypted into two shares. All the shares generated are noisy images, and so no secret information could be obtained. Nevertheless, when the generated two shares are superimposed, S can be restored, and its construction method is as follows.

One black-white pixel e is splitted from original image into two black-white subpixels. If e is black (resp., white), then Strategy 1 or Strategy 2 in the lower (resp., upper) row of Table 1 is randomly selected, and the overlay of all subpixels is entirely black (resp., entirely white). As a result, the initial secret image can be capable of recovering lossless assuming that all shares are superimposed. This structure can be indicated as the sets $\mathscr{C}^0$ and $\mathscr{C}^1$ of matrices in Table 1; more specifically, the above encoding and decoding procedures can be indicated as the function Enc: $\{0, 1\} \longrightarrow \{0, 1\}^{2\times 1}$ and Dec: $\{0, 1\}^{2\times 1} \longrightarrow \{0, 1\}$ given by

$$Enc(b) = \mathscr{C}_U^b \text{ and } Dec(M) = (m_{11} \oplus m_{21}), \tag{3}$$

for $b \in \{0, 1\}$ and $M = (m_{ij}) \in \{0, 1\}^{2\times 2}$, respectively. From the function operation, we can see that when the XOR operation result is different from the original secret pixel, we only need to invert the pixel value of any secret sharing matrix. Therefore, in the initial generation of subpixels, we can randomly select the values without considering whether the results of all pixel XOR are correct. If the result of the pixel set XOR is not correct, take the reverse correction.

Pixel expansion refers to the number of subpixels in shares encrypted from a secret pixel. In the example above, each pixel in the secret image is encrypted as a subpixel in each share. Therefore, there is no pixel extension. The degree of pixel expansion is negatively correlated with the actual resolution of the share image. A VCS with the lowest pixel extension indicates that its encryption is optimal.

3.3. Instance of (k, m, n)-RGWVCS. Through the above analysis of the shadow encryption system, we can have a sense of the system. In this section, we will demonstrate the detailed process of encryption and decryption using an instance (3, 2, 4)-RGWVCS. Thus, we will have a clearer and more specific understanding of the shadow encrypting architecture. In the (3, 2, 4)-RGWVCS, the secret image needs to be encrypted into 4 secret sharing matrices with the same size. During the decryption side, it requires at least 2 shares to complete the secret recovery when the privileged participate in, whereas at least 3 shares are required.

For each pixel of S, the same position of the 4 secret sharing matrices, $\{M_1, M_2, M_3, M_4\}$, is randomly initialized to 0 or 1 by the Monte Carlo method, then $m_1 \oplus m_2 \oplus m_3 \oplus m_4$ is calculated, and it is compared with the pixel of S. If the comparison result is inconsistent, the random one of the four pixels is reversed. When all pixels of S are handled, 4 secret sharing matrices in the final form of noise are generated. Next, we will prove the feasibility of the Monte Carlo method (Analysis 1).

Analysis 1. In the proposed RGWVCS scheme, step 1 (initialize n zero secret sharing matrices with weight ω_i) satisfies the completely random fairness principle.

Proof. In the proposed scheme, we use the Monte Carlo method to achieve secret layering according to the weight, in which 1000 groups of random cast experiments are conducted on four targets with equal probability by the Monte Carlo method. The experimental results are listed in Figure 3.

The unit square plane in Figure 3 is divided into four areas with the same size: S1, S2, S3, and S4, and then 1000 points are scattered on the plane with equal probability. The dots are randomly placed in different regions, and the dots in different regions are assigned different colors. Through experiments and statistics, we can see the projection of 1000 points in the four regions in Figure 3(a) and its statistical results in Figure 3(b). The probability of hitting S1, S2, S3, and S4 is 25.9%, 24.7%, 25.7%, and 23.7%, respectively. When subtracted from the average, it was +0.9%, −0.3%, +0.7%, and −1.3%, respectively. From the above data analysis, we can draw a conclusion that the hit ratio of these four regions is nearly the same, indicating that in the designed RGWVCS scheme, the Monte Carlo method satisfies the completely random fairness principle.

```
Input: secret image $Q$ of size $H \times W$, $(k, m, n)$ and $\{\omega_1, \omega_2, \cdots, \omega_n\}$;
Output: the $n$ secret sharing matrices $\{M_1, M_2, \cdots, M_n\}$.
(1) Initialize $n$ zero secret sharing matrices $M_i = \{m^i(s,t) = 0 | 1 \le s \le H, 1 \le t \le W\}$ with weight $\omega_i$, $(i = 1, 2, \cdots, n)$, i.e., the probability
    of $M_i$ participates in the secret recovery is $\omega_i$.
(2) For any pixel $m^i(s,t)$ in $M_i$, assign $\{0, 1\}$ randomly to it using the Monte Carlo method.
(3) Assume the $p$-th participant is a privileged person. There are two cases as follows:
    (a) If $M_p$ participates the secret recovery, select $j$ secret sharing matrices randomly, i.e., $\{M_{i1}, M_{i2}, \cdots, M_{ij}\} (1 \le i \le n, m \le j \le n)$;
    (b) If $M_p$ is not involved the secret recovery, select $j$ secret sharing matrices randomly, i.e., $\{M_{i1}, M_{i2}, \cdots, M_{ij}\} (1 \le i \le n, k \le j \le n)$.
(4) Select $c \in \{1, 2, \cdots, j\}$ randomly and execute the following formula:
        $m_c^i(s,t) = \{ m_c^i(s,t), Q(s,t) = m_1^i(s,t) \oplus m_2^i(s,t) \oplus \cdots \oplus \overline{m_j^i(s,t) m_c^i(s,t)}$, otherwise.
(5) Generate $j$ secret sharing matrices $\{M_{i1}, M_{i2}, \cdots, M_{ij}\}$.
(6) Repeat steps 2–5 $(s * t)$ times to generate $n$ secret sharing matrices $\{M_1, M_2, \cdots, M_n\}$.
```

Algorithm 1: (k, m, n)—RGWVCS.

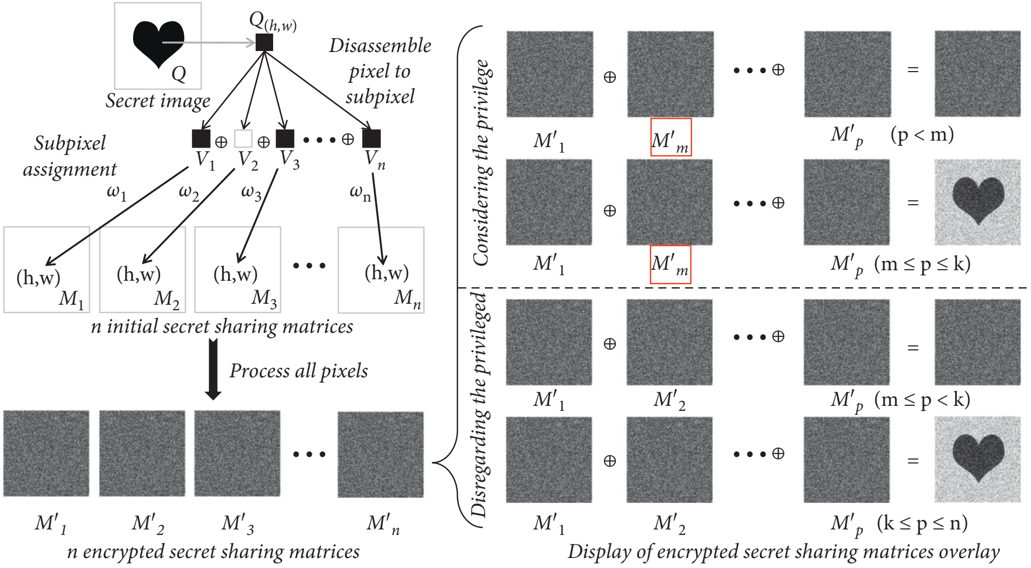

Q
Secret image
Q(h,w)
Disassemble pixel to subpixel
Subpixel assignment
V1 ⊕ V2 ⊕ V3 • • • ⊕ Vn
ω1
ω2
ω3
ωn
(h,w) M1
(h,w) M2
(h,w) M3
(h,w) Mn
n initial secret sharing matrices
Process all pixels
M'1
M'2
M'3
M'n
n encrypted secret sharing matrices
Considering the privilege
Disregarding the privileged
M'1 ⊕ M'm • • • ⊕ M'p = (p < m)
M'1 ⊕ M'm • • • ⊕ M'p = (m ≤ p ≤ k)
M'1 ⊕ M'2 • • • ⊕ M'p = (m ≤ p < k)
M'1 ⊕ M'2 • • • ⊕ M'p = (k ≤ p ≤ n)
Display of encrypted secret sharing matrices overlay

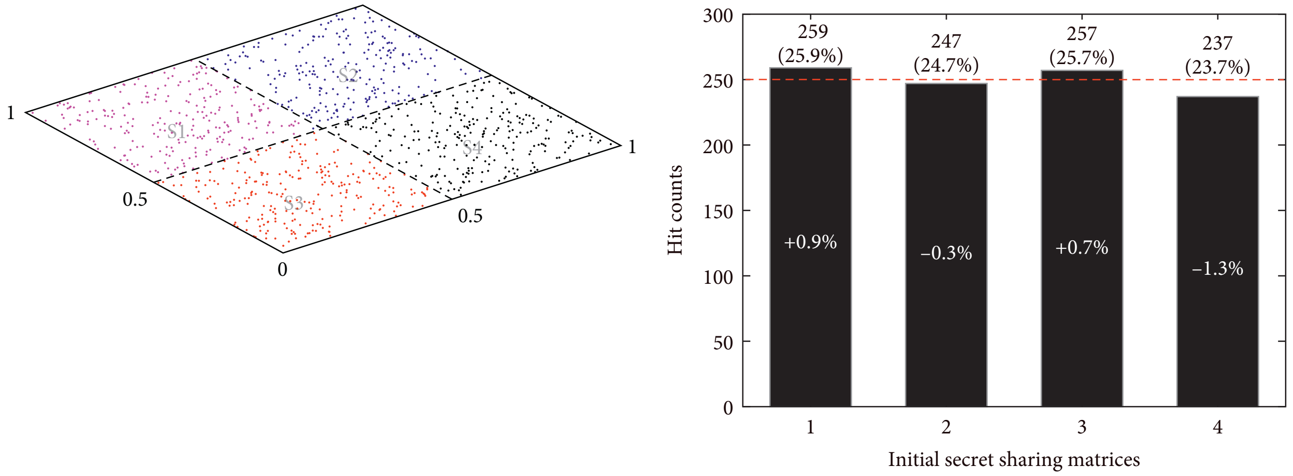
Figure 2: Shadow encrypting architecture of the designed RGWVCS. The left half of the encryption architecture describes the generation process of the subpixel and encrypted secret sharing matrices, while the right half is the decryption process of secret sharing matrices.

Table 1: Encryption structure for a single black-white pixel, and the sets $\mathscr{C}^0$ and $\mathscr{C}^1$ represent matrices (row: share, column: strategy, 0: white, 1: black).

Pixel		Strategy 1	Strategy 2	Subpixel matrix
□	Share 1	■	□	$\mathscr{C}^0 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$
	Share 2	■	□	
■	Share 1	■	□	$\mathscr{C}^1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$
	Share 2	□	■	

3.4. Feasibility and Safety Analysis. After the detailed introduction and analysis of the relevant algorithms, we will carry out theoretical analysis on the security and visual recognition of the designed scheme from Conditions 1 and 2.

Lemma 1. *In the proposed (k, m, n)-RGWVCS from Construction 1,* $\{0, 1\}$ *are assigned to* $\{m^1(s,t), m^2(s,t), \cdots, m^n(s,t)\}$ *randomly in the third step. If* $S'(s,t) = m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)$, *we can deduce* $P(S'[CS0] = 0) = P(S'[CS1] = 0) = (1/2)^{k-1}$.



(a)

(b)

Figure 3: Distribution and statistical histograms of 1000 random points in four regions of equal area indicate that the Monte Carlo method is reliable: (a) random point density graph; (b) statistical histogram of casting point probability.

Proof. In the proposed (k, m, n)-RGWVCS from Construction 1, $\{0, 1\}$ are assigned to $\{m^1(s,t), m^2(s,t), \cdots, m^n(s,t)\}$ randomly using the Monte Carlo method in the third step. That means that the probability of m being equal to 0 and 1 is equivalent, i.e., $P(m^i(s,t)=0) = P(m^i(s,t)=1) = 1/2$. The white areas of the secret image S are denoted as CS0 = {(s, t) |S(s, t) = 0, $1 \leq s \leq H$, $1 \leq t \leq W$}. Therefore, $S'[CS0] = S'(s,t) = m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)$. Given that $P(m^i(s,t)=0) = P(m^i(s,t)=1) = 1/2$, so $P(S'[CS0]=0) = (1/2)^{k-1}$. By the same token, $P(S'[CS1]=0) = (1/2)^{k-1}$.

Lemma 2. *In the proposed (k, m, n)-RGWVCS from Construction 1, if* $S(s,t)=0$, $P(m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)=0) = P(m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^{k-1}(s,t)=0) = (1/2)^{k-1}$. *On the contrary,* $S(s,t)=1$, $P(m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)=0)=0$.

Proof

(1) Firstly, when $S(s,t)=0$, we will prove the following:

$$P\left(m^1(s,t) \otimes \cdots \otimes \mathrm{m}^k(s,t)=0\right) = P\left(m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)=0\right) = (1/2)^{k-1}. \tag{4}$$

If $m^k(s,t)=0$, then $P(m^1(s,t) \otimes \cdots \otimes m^k(s,t)=0) = P(m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)=0)$ is set up because $m^k(s,t)$ is transparent and $m^1(s,t) \otimes \cdots \otimes m^k(s,t) = m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)$.

If $m^k(s,t)=1$, $m^k(s,t)$ is equal to one of $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$. If not, $m^k(s,t)$ is complementary to each bit of $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$, thus $m^1(s,t) = m^2(s,t) = \cdots = m^{k-1}(s,t) = 0$. According to the third step, we have $S(s,t) = m^1(s,t) \oplus m^2(s,t) \oplus \cdots \oplus m^k(s,t)$ so that $S(s,t) = 0 \oplus 0 \oplus \cdots \oplus 0 \oplus 1 = 1$ with contradiction to $S(s,t)=0$. Therefore, $m^k(s,t)$ is equal to one of $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$, $m^1(s,t) \otimes \cdots \otimes m^k(s,t) = m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)$; therefore, it can be proved that $P(m^1(s,t) \otimes \cdots \otimes \mathrm{m}^k(s,t)=0) = P(m^1(s,t) \otimes \cdots \otimes m^{k-1}(s,t)=0) = (1/2)^{k-1}$.

(2) Secondly, when $S(s,t)=1$, we will prove the following:

$$P\left(m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)=0\right) = 0. \tag{5}$$

If $m^k(s,t)=0$, then $m^k(s,t)$ is complementary to one of $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$. Otherwise, $m^k(s,t)$ is equal to each bit of $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$, thus $m^1(s,t) = m^2(s,t) = \cdots = m^{k-1}(s,t) = 0$. According to the third step, we have $S(s,t) = m^1(s,t) \oplus m^2(s,t) \oplus \cdots \oplus m^k(s,t)$ so that $S(s,t) = 0 \oplus 0 \oplus \cdots \oplus 0 \oplus 0 = 0$ with contradiction to $S(s,t)=1$. Therefore, $m^k(s,t)$ is complementary to one of $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$, $m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t) = 1$; therefore, it can be proved that $P(m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)=0)=0$.

If $m^k(s,t)=1$, $P(m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)=0)=0$ is satisfied because $m^k(s,t)$ is opaque and $m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)=1$.

In conclusion, if we superpose $m^1(s,t), m^2(s,t), \cdots, m^k(s,t)$, contrast will be visible to the naked eye, i.e., the secret image will be recovered, which is the core to secret recovery in (k, m, n)-RGWVCS. Otherwise, if we superpose any $k-1$ shares $m^1(s,t), m^2(s,t), \cdots, m^{k-1}(s,t)$, $\alpha = 0$ due to Lemma 1.

Theorem 1. *The designed scheme is an effective (k, m, n)-RGWVCS structure.*

Proof. Firstly, based on Lemma 2, when $j < k$, the generated j bits cannot cover $m^1(s,t), m^2(s,t), \cdots, m^k(s,t)$. $S'(s,t) = m^1(s,t) \otimes \cdots \otimes m^j(s,t)$, $P(S'[CS0] = 0) = (1/2)^j$, and $P(S'[CS1] = 0) = (1/2)^j$. Hence, $P_0 = P_1$ when $j < k$ so that Condition 2 is satisfied. Secondly, when $j \geq k$, the generated j bits can cover $m^1(s,t), m^2(s,t), \cdots, m^k(s,t)$ with a certain probability. $S'(s,t) = m^1(s,t) \otimes m^2(s,t) \otimes \cdots \otimes m^k(s,t)$, on account of $P(S'[CS0] = 0) = (1/2)^{k-1} > P(S'[CS1] = 0) = 0$ and $P_0 > P_1$, the designed scheme satisfies Condition 1. Finally, based on the above justification of Conditions 1 and 2, the proposed scheme is an effective (k, m, n)-RGWVCS structure.

Proposition 1. *The scheme is designed to take the possible existence of the privileged into account.*

Proof. Through the analysis of the designed method, a larger weight leads to a larger probability of covering $m^1(s,t), m^2(s,t), \cdots, m^k(s,t)$, because the k initial secret sharing matrices $\{M_1, M_2, \cdot \cdot \cdot, M_n\}$ are constructed corresponding to $\{\omega_1, \omega_2, \cdot \cdot \cdot, \omega_n\}$ in the first step of Algorithm 1. This indicates that $\alpha_{i'_1,i'_2,\cdots,i'_j} > \alpha_{i_1,i_2,\cdots,i_j}$, when $\sum_{t=1}^{j} \omega_{i'_t} > \sum_{t=1}^{j} \omega_{i_t}$ and $\alpha_{i_1,i_2,\cdots,i_{j+1}} > \alpha_{i_1,i_2,\cdots,i_j}$ where $j \geq k$.

3.5. Practical Application. There is a lot of class division in real life, and being high class also means having certain privileged rights. The traditional visual secret sharing scheme ignores the existence of real privileges and becomes unsuitable for all real situations. The weighted visual secret sharing scheme with privileges can solve this problem well. The following company is taken as an example.

In this case of the company above, there are two classes, including manager and clerk. Managers are superior to their clerks and thus have a greater capacity of secret recovery. It can be seen from Figure 4 that when $m \leq n \leq k$, the combination of n clerks cannot recover S, while the combination of a manager and $n - 1$ clerks can recover S, highlighting the effectiveness of the privilege. When $n \geq k$, although the shares of n clerks can recovery the secret image, the quality of S$'$ is lower than that recovered by one manager and $n - 1$ clerks. It also highlights the superiority of the privilege enjoyed by the privileged. Most of the existing weighted visual secret sharing schemes give high-weighted participants the privilege of recovering higher-quality secret images. However, the proposed visual secret sharing scheme considering the privileged not only gives the privileged that right but also gives the privileged the privilege to be capable of recovering S while the same number of ordinary participants cannot recover S. This makes the privileged more prominent and more relevant to the actual situation.

4. Experimental Results and Discussion

In this section, we will show the performance of RGWVCS through experimental results, in which the validity of the scheme is verified from the quality and characteristics of the encrypted and decrypted images, and the superiority of the scheme is illustrated by comparing with other schemes.

4.1. Display of Image Recovery. In the designed RGWVCS for (3, 2, 4) threshold, secret image S is divided into four encrypted images M_1, M_2, M_3, and M_4. When the privileged people are considered in the secret recovery, at least two encrypted images are required for restoring the secret. Conversely, when no privileged people participates in the secret recovery, a minimum of three encrypted images are required.

It can be seen from Figure 5 that when $S_q = 1$, what is displayed is the encrypted images M_1, M_2, M_3, and M_4. When $S_q = 2$, the result of any two encrypted images superimposed is shown in the figure. The figure shows that half of the six possible combinations can complete the secret recovery. The commonality of the combinations that can complete the secret recovery is that they all contain the privileged M_4. This is in line with the designed scheme that only $m(m = 2)$ participants are needed to complete the secret recovery when the privileged participates in the secret recovery. In terms of the recognition of recovered secrets, it can be concluded that the superimposed weight is proportional to the secret recognition. When $S_q = 3$, any combination of three encrypted images can accomplish the secret recovery. As with $S_q = 2$, the greater the total weight of participants, the higher the degree of secret identification. When $S_q = 4$, i.e., all participants M_1, M_2, M_3, and M_4 are superposed together, it can recover S losslessly. The recovered S$'$ is identical to the binary secret image S$(\alpha = 1)$.

So, according to the above analysis, we can draw the following conclusions:

(1) The experimental results of $S_q = 2$ and $S_q = 3$ agree with the (3, 2, 4) threshold, and the design idea of considering the privileged is at least three participants are needed to complete the secret recovery when the privileged is not considered; otherwise, only two participants are needed.

(2) The quality of secret recovery is proportional to the total weight of participants.

4.2. Contrast and Weight Analysis. In this section, we will first use the contrast as the main indicator to accurately describe the quality of the recovered secret images from different combinations. Then, we will analyse the relation between contrast and weight. Finally, we will compare the contrast between our scheme and relevant schemes.

From Table 2, we can see that the contrast obtained by the superposition of any two encrypted images in M_1, M_2, and M_3 is always 0. Once M_4 participates in the secret recovery, it only needs 2 encrypted images to restore S that can be recognized by the naked eye. With the increase in the total weight of the participants, the recognition degree of the secret image is also higher; that is, the clarity of the secret is directly proportional to the total weight of the participants. The analysis shows that the contrast is positively correlated

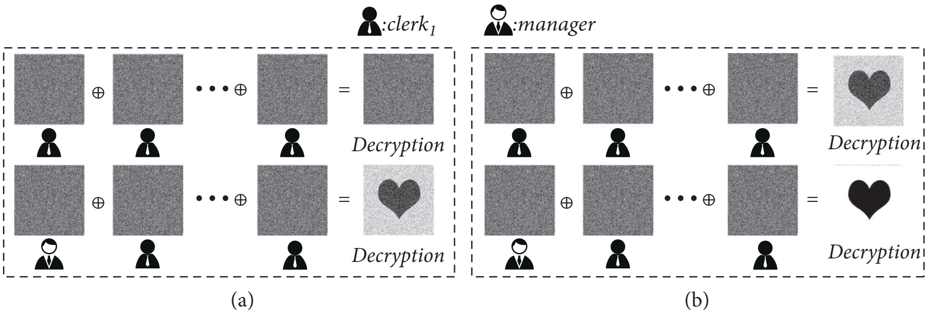


Figure 4: Schematic of an instance of privilege in a company situation, consisting of two distinct classes of manager and clerk. The two classes have different secret image recovery ability: (a) $m \leq n \leq k$ and (b) $n \geq k$.

with the weight, and the image quality after recovery is better than relevant schemes.

In Figure 6, we analyse the relation between the contrast and total weight of the decrypted image with the different combinations, where Q_1, $Q2$, Q_3, and Q_4 are the four abnormal points. For Q_1, the total weight of the participants $\omega(\omega = 0.5)$ is the same in combination $M_2 \oplus M_3$ and $M_1 \oplus M_4$. However, $M_2 \oplus M_3$ cannot recover the secret while $M_1 \oplus M_4$ can make the secret identifiable. This phenomenon of the same weight but different decoding ability shows the superiority of the privileged in the designed scheme. For $Q2$ and Q_3, as can be seen from the figure, the trend around $Q2$ and Q_3 is basically the same. Both Q_3 and Q_3 are maximum points, which indicates that the contrast is positively correlated with the total weight. For Q_4, the total weight at this point is as high as 0.9 while the contrast is only 0.4189. This phenomenon of high weight and low contrast is in urgent need of solution. In terms of the rest of the features, the designed scheme has the following characteristics: (1) no pixel expansion; (2) no codebook design; and (3) no weight leakage.

For fear of the influence of randomness of the secret image in above part, we carried out much more identical tests with the designed scheme and other related schemes, in which we selected randomly 100 binary images with different sizes and patterns for encryption and decryption. Finally, we calculated the average contrast for different combinations in 100 tests to prove the superiority of the designed scheme. In the experiment, we set that $k = 2$, $n = 4$, $\omega_1 = 0.1$, $\omega_2 = 0.2$, $\omega_3 = 0.3$, and $\omega_4 = 0.4$.

Table 3 shows the contrast values of the secret restored images after superimposed with the proposed privileged scheme and the relevant representative WVCS schemes. The results show that the image restoration quality of the proposed scheme is superior to that of other schemes in general except for {3, 4}. Since in the proposed scheme, the $n - k$ bits in the tail must have the same value as a random bit in the front k bits, and thus in the secret recovery phase, the proposed scheme improves the coverage of the front k bits and enhances the clarity of the secret image recovered after superposition.

4.3. Feature Comparison. In this section, we will compare the quality of the recovered image and a series of characteristics of the proposed scheme with some typical WVCS schemes with admirable features, such as Yang et al. [30], Fan [31], and Tu et al. [32] schemes. While we are concerned about the quality of the recovery image, we are also concerned about other features of the designed scheme, such as pixel expansion, codebook design, and weight leakage.

In Table 4, our design scheme is compared with other related schemes in a series of main features where the threshold value represents the threshold parameter supported by the scheme; recovery measure refers to the recovery methods that may be used by the solution; no pixel expansion indicates that the share distributed to the participants has the same size as S; no codebook design means that there is no codebook design in the generation phase of shares; and the scheme will be more secure and practical without weight leakage after weighted treatment.

From Table 4, we can see that the proposed RGWVCS has the following superiorities: (1) there is no pixel expansion; (2) no codebook design; (3) (k, m, n) threshold, the privileged are considered in the scheme; (4) the generated shares have no weight leakage; and (5) the image quality is better than that of the correlation weighted scheme, and the secret image can be recovered losslessly. To sum up, in a series of representative characteristic indicators, the scheme we designed is superior to other relevant schemes.

After comparing the characteristics of different schemes, we use the objective evaluation index recall to quantitatively compare our scheme with others. The formula of recall is as follows, where TP is the number of samples with the restored pixel value of 1 and the judgment is correct, and FN is the number of samples with the restored pixel value of 0 and the judgment is wrong. It can be seen from Figure 7 that the program proposed in this paper has an average recall of 87.6% due to the addition of the privilege mechanism, which is higher than all the schemes that participated in the comparison. The schemes proposed by Tan et al. and Liu et al. have relatively moderate results due to weight leakage. Yang et al. and Shamir et al. proposed the schemes earlier, and they have varying degrees of feature defects, so the recall is low. From this, we can see the advanced nature of the scheme proposed in this article.

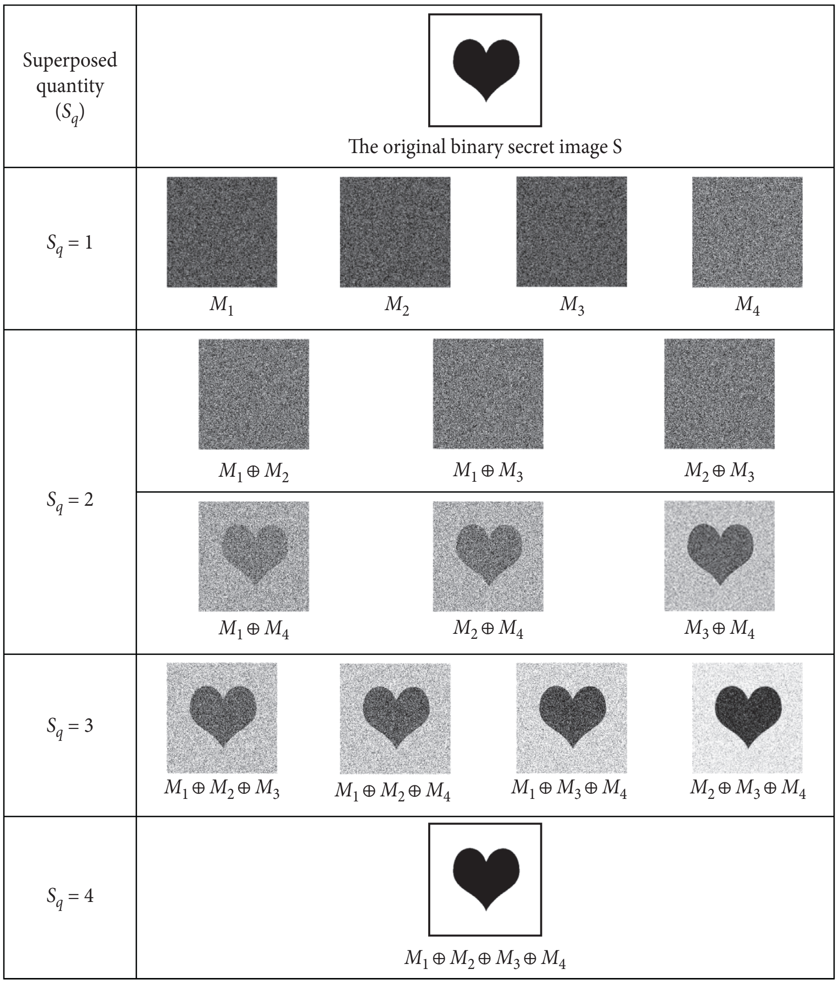

Superposed quantity (S_q)
The original binary secret image S
$S_q = 1$
M_1
M_2
M_3
M_4
$S_q = 2$
$M_1 \oplus M_2$
$M_1 \oplus M_3$
$M_2 \oplus M_3$
$M_1 \oplus M_4$
$M_2 \oplus M_4$
$M_3 \oplus M_4$
$S_q = 3$
$M_1 \oplus M_2 \oplus M_3$
$M_1 \oplus M_2 \oplus M_4$
$M_1 \oplus M_3 \oplus M_4$
$M_2 \oplus M_3 \oplus M_4$
$S_q = 4$
$M_1 \oplus M_2 \oplus M_3 \oplus M_4$

Figure 5: The outcomes of the proposed RGWVCS for (3, 2, 4) threshold where S_q refers to the number of shares stacked when recovering the secret. The image S is the selected original heart-shaped binary secret image of size 454 × 454. When $S_q = 1$, M_1, M_2, M_3, and M_4 are four encrypted images of the same size as S. The weights of M_1, M_2, M_3, and M_4 are $\omega_1 = 0.1$, $\omega_2 = 0.2$, $\omega_3 = 0.3$, and $\omega_4 = 0.4$, respectively. M_4 has the maximum weight so that he is a privileged.

Table 2: The contrast and weight of the decrypted image formed by superimposing shares.

Superimposed encrypted image	Contrast	Weight
M_1	—	0.1
M_2	—	0.2
M_3	—	0.3
M_4	—	0.4
$M_1 \oplus M_2$	—	0.3
$M_1 \oplus M_3$	—	0.4
$M_2 \oplus M_3$	—	0.5
$M_1 \oplus M_4$	0.1266	0.5
$M_2 \oplus M_4$	0.1732	0.6
$M_3 \oplus M_4$	0.2581	0.7
$M_1 \oplus M_2 \oplus M_3$	0.2525	0.6
$M_1 \oplus M_2 \oplus M_4$	0.2804	0.7
$M_1 \oplus M_3 \oplus M_4$	0.3399	0.8
$M_2 \oplus M_3 \oplus M_4$	0.4189	0.9
$M_1 \oplus M_2 \oplus M_3 \oplus M_4$	1	1

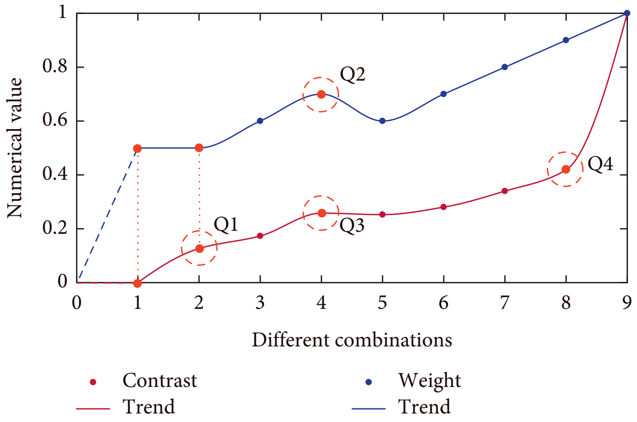


Figure 6: The relation between the contrast and total weight of the decrypted image with different combinations.

Table 3: The contrast of the proposed scheme under the (3, 2, 4) threshold, compared with Tu et al. [32], Fan et al. [31], and Yang et al. [30].

Collected shadows	Ours	Tu et al. [32]	Fan et al. [31]	Yang et al. [30]
{1, 2}	—	0.1035	0.1105	0.1096
{1, 3}	—	0.1226	0.1526	0.1531
{1, 4}	**0.2432**	0.1505	0.1993	0.1985
{2, 3}	—	0.1681	0.1981	0.1986
{2, 4}	**0.2759**	0.1872	0.2486	0.2491
{3, 4}	0.2962	0.2176	0.3021	**0.3045**
{1, 2, 3}	**0.2925**	0.2249	0.2488	0.2485
{1, 2, 4}	**0.3486**	0.2523	0.3038	0.3040
{1, 3, 4}	**0.3821**	0.2841	0.3623	0.3642
{2, 3, 4}	**0.4414**	0.3192	0.4263	0.4278
{1, 2, 3, 4}	**1**	0.3873	0.4988	0.4996

Table 4: A comparison of series of representative features between the designed scheme and relevant VCSs.

Proposer	Threshold	Recovery efficiency (complexity)	No pixel expansion	No codebook design	No weight leakage	Weighted
Shamir et al. [1]	(k, n)	OR (O (1))	✗	✗	✓	✓
Yang et al. [30]	$(2, n)$	OR (O (1))	✓	✗	✓	✓
Tan et al. [40]	(k, n)	Modular ($O(k)$)	✓	✓	✗	✓
Liu et al. [41]	(k, n)	OR/XOR (O (1)/O (k))	✓	✓	✗	✓
Fan et al. [31]	(k, n)	OR (O (1))	✓	✓	✓	✓
Tu et al. [32]	(k, n)	OR (O (1))	✓	✓	✓	✓
Ours	(k, m, n)	OR (O (1))	✓	✓	✓	✓

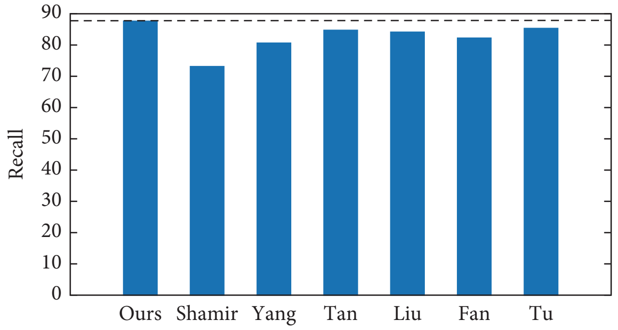


Figure 7: A comparison of recall between the designed scheme and relevant VCSs.

$$R = \frac{TP}{TP + FN}. \quad (6)$$

5. Conclusions

The traditional (k, n)-threshold visual secret sharing scheme encrypts a secret image into n noise shares, and the hidden secret in the secret image can be recovered by superimposing k or more shares. Nevertheless, those methods did not consider differences in the importance of participants, so each participant in the scheme has the same level of privilege, that is, the same secret recovery capability. This paper designs a novel random grid-based weighted visual cryptography scheme for a (k, m, n) ($m < k < n$) threshold (RGWVCS). This scheme takes the differences between privileged participants and ordinary participants into account and gives the privileged people a better decryption ability than ordinary people. When privileged people participate in secret recovery, only m participants are needed, and when no privileged person participates, k participants are needed. Through the analysis of the experimental results and the comparison with the relevant schemes, the designed scheme has several features as follows: (1) there is no pixel expansion; (2) no codebook design; (3) (k, m, n) threshold, the privileged are considered in the scheme, (4) the generated shares have no weight leakage, and (5) the contrast of the designed RGWVCS is, on average, increased by 32.85% compared with the relevant schemes which shows that the image quality of the revealed secret image is greatly enhanced. In the future, we will further modify the algorithm to improve the resolution of secret images and the security of encryption. Furthermore, the current visual secret sharing scheme can only be used for binary image encryption, and grayscale and color images cannot be well encrypted and restored. We will work to make the designed scheme also applicable to the encryption of grayscale and color images.

Data Availability

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Key R&D Program of China under grant no. 2018YFB1003205; by the National Natural Science Foundation of China under grant nos. U1836208, U1536206, U1836110, 61972207, 61602253, and 61672294; by the Engineering Research Center of Digital Forensics, Ministry of Education; by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) Fund; and by the Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET) Fund, China.

References

[1] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology—EUROCRYPT'94*, pp. 1–12, Springer-Verlag, Berlin, Germany, 1995.

[2] R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of the National Computer Conference*, pp. 313–317, Monval, NJ, USA, 1979.

[3] Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612-613, 1979.

[4] M. Naor and B. Pinkas, "Visual authentication and identification," *Advances in Cryptolog—CRYPTO'97*, Springer-Verlag, Berlin, Germany, pp. 322–336, 1997.

[5] M. Naor and A. Shamir, "Visual cryptography ii: improving the contrast via the cover base," in *Security Protocols*, pp. 197–202, Springer, Berlin, Germany, 1996.

[6] S. J. Shyu and M. C. Chen, "Optimum pixel expansions for threshold visual secret sharing schemes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 960–969, Sep. 2011.

[7] S. J. Shyu and M. C. Chen, "Minimizing pixel expansion in visual cryptographic scheme for general access structures," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 9, pp. 1557–1561, 2015.

[8] C.-C. Chang and J.-C. Chuang, "An image intellectual property protection scheme for gray-level images using visual secret sharing strategy," *Pattern Recognition Letters*, vol. 23, no. 8, pp. 931–941, 2002.

[9] R. Lukac and K. N. Plataniotis, "Bit-level based secret sharing for image encryption," *Pattern Recognition*, vol. 38, no. 5, pp. 767–772, 2005.

[10] G. Ateniese, C. Blundo, A. D. Santis, and D. R. Stinson, "Extended capabilities for visual cryptography," *Theoretical Computer Science*, vol. 250, no. 1-2, pp. 143–161, 2001.

[11] C. Blundo, A. De Santis, and D. R. Stinson, "On the contrast in visual cryptography schemes," *Journal of Cryptology*, vol. 12, no. 4, pp. 261–289, 1999.

[12] X. Yan, S. Wang, X. Niu, and C.-N. Yang, "Halftone visual cryptography with minimum auxiliary black pixels and uniform image quality," *Digital Signal Processing*, vol. 38, pp. 53–65, 2015.

[13] C.-C. Lin and W.-H. Tsai, "Visual cryptography for gray-level images by dithering techniques," *Pattern Recognition Letters*, vol. 24, no. 1-3, pp. 349–358, 2003.

[14] C.-N. Yang, L.-Z. Sun, and S.-R. Cai, "Extended color visual cryptography for black and white secret image," *Theoretical Computer Science*, vol. 609, pp. 143–161, 2016.

[15] S. Cimato, R. De Prisco, and A. De Santis, "Probabilistic visual cryptography schemes," *The Computer Journal*, vol. 49, no. 1, pp. 97–107, 2006.

[16] I. Kang, G. R. Arce, and H. K Lee, "Color extended visual cryptography using error diffusion," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 20, no. 1, pp. 132–145, 2011.

[17] J. Chen, T.-S. Chen, H.-C. Hsu, and Y.-H. Lin, "Using multi-ringed shadow image of visual cryptography to hide more secret messages," *The Imaging Science Journal*, vol. 57, no. 2, pp. 101–108, 2009.

[18] S. Droste, "New results on visual cryptography," in *Advances in Cryptology—CRYPTO*, pp. 401–415, Springer, Berlin, Germany, 1996.

[19] S. J. Shyu and H. W. Jiang, "General constructions for threshold multiple-secret visual cryptographic schemes," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 5, pp. 733–743, 2013.

[20] C.-N. Yang, H.-W. Shih, C.-C. Wu, and L. Harn, “k out of n region incrementing scheme in visual cryptography,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 5, pp. 799–810, 2012.

[21] Y. C. Chen, G. Horng, and D. S. Tsai, “Comment on “Cheating prevention in visual cryptography”” *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3319–3323, 2012.

[22] M. Hu and W. G. Tzeng, “Cheating prevention in visual cryptography,” *IEEE Transactions on Image Processing*, vol. 16, no. 6, pp. 36–45, 2007.

[23] X. Liu, S. Wang, X. Yan, and W. Zhang, “Random grid-based threshold visual secret sharing with improved visual quality and lossless recovery ability,” *Multimedia Tools and Applications*, vol. 77, no. 16, pp. 20673–20696, 2018.

[24] S. P. Kannojia and J. Kumar, “XOR-based visual secret sharing scheme using pixel vectorization,” *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 14609–14635, 2021.

[25] S. Shivani and S. Agarwal, “Novel basis matrix creation and preprocessing algorithms for friendly progressive visual secret sharing with space-efficient shares,” *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 8711–8744, 2017.

[26] S. Shivani and S. Agarwal., “Progressive visual cryptography with unexpanded meaningful shares,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 12, no. 4, pp. 1–24, 2016.

[27] S. Shivani and S. Agarwal, “VPVC: verifiable progressive visual cryptography,” *Pattern Analysis & Applications*, vol. 21, no. 1, pp. 139–166, 2016.

[28] S. Shivani, “Multi secret sharing with unexpanded meaningful shares,” *Multimedia Tools & Applications*, vol. 77, no. 5, pp. 6287–6310, 2017.

[29] Y.-C. Hou, Z.-Y. Quan, and C.-F. Tsai, *A Privilege-Based Visual Secret Sharing Model*, Academic Press, Cambridge, MA, USA, 2015.

[30] C.-N. Yang, J.-K. Liao, and D.-S. Wang, “New privilege-based visual cryptography with arbitrary privilege levels,” *Journal of Visual Communication and Image Representation*, vol. 42, pp. 121–131, 2017.

[31] H. C. Chao, “Random-grid based progressive visual secret sharing scheme with adaptive priority,” *Digital Signal Processing*, vol. 68, pp. 69–80, 2017.

[32] T. Y. Tu, T. H. Chen, J. Yang, and C. H. Wang, “A weighted threshold visual cryptography,” *Communications in Computer and Information Science*, vol. 1013, 2019.

[33] X. Yan, S. Wang, A. A. A. El-Latif, and X. Niu, “Random grids-based visual secret sharing with improved visual quality via error diffusion,” *Multimedia Tools and Applications*, vol. 74, no. 21, pp. 9279–9296, 2015.

[34] X. Yan, X. Liu, and C.-N. Yang, “An enhanced threshold visual secret sharing based on random grids,” *Journal of Real-Time Image Processing*, vol. 14, no. 1, pp. 61–73, 2018.

[35] G. Gao, X. Wan, S. Yao, Z. Cui, C. Zhou, and X. Sun, “Reversible data hiding with contrast enhancement and tamper localization for medical images,” *Information Sciences*, vol. 385-386, pp. 250–265, 2017.

[36] M. Alloghani, M. M. Alani, D. Al-Jumeily et al., “A systematic review on the status and progress of homomorphic encryption technologies,” *Journal of Information Security and Applications*, vol. 48, Article ID 102362, 2019.

[37] X. Jia, D. Wang, D. Nie, X. Luo, and J. Z. Sun, “A new threshold changeable secret sharing scheme based on the Chinese remainder theorem,” *Information Sciences*, vol. 473, no. 13–30, 2019.

[38] G. Gao, Z. Cui, and C. Zhou, “Blind reversible authentication based on PEE and CS reconstruction,” *IEEE Signal Processing Letters*, vol. 25, no. 7, pp. 1099–1103, 2018.

[39] X. Yan, Y. Lu, H. Huang, L. Liu, and S. Wan, “Clarity corresponding to contrast in visual cryptography,” in *Proceedings of the Social Computing: Second International Conference of Young Computer Scientists, Engineers and Educators, ICYCSEE 2016*, pp. 249–257, Harbin, China, August 2016.

[40] L. Tan, Y. Lu, X. Yan, L. Liu, and L. Li, “Weighted secret image sharing for a (k, n) threshold based on the Chinese remainder theorem,” *IEEE Access*, vol. 7, pp. 59278–59286, 2019.

[41] F. Liu, X. Yan, L. Liu, Y. Lu, and L. Tan, “Weighted visual secret sharing with multiple decryptions and lossless recovery,” *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 5750–5764, 2019.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 5595243, 11 pages
https://doi.org/10.1155/2021/5595243

Research Article

Delegated Key-Policy Attribute-Based Set Intersection over Outsourced Encrypted Data Sets for CloudIoT

Yanfeng Shi[1] **and Shuo Qiu**[2]

[1]*School of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China*
[2]*School of Software Engineering, Jinling Institute of Technology, Nanjing 211169, China*

Correspondence should be addressed to Shuo Qiu; shuoqiu@jit.edu.cn

Received 8 January 2021; Revised 11 February 2021; Accepted 24 February 2021; Published 18 March 2021

Academic Editor: Zhaoqing Pan

Private set intersection (PSI) is a fundamental cryptographic primitive, allowing two parties to calculate the intersection of their data sets without exposing additional private information. In cloud-based IoT system, IoT-enabled devices would like to outsource their data sets in their encrypted form to the cloud. In this scenario, how to delegate the set intersection computation over outsourced encrypted data sets to the cloud and how to achieve the fine-grained access control for PSI without divulging any additional information to the cloud are still open problems. With that in mind, in this work, we combine key-policy attribute-based encryption (KP-ABE) and PSI to introduce such a novel concept, called *delegated key-policy attribute-based set intersection over outsourced encrypted data sets (KP-ABSI)*, to solve this problem. Then we propose a first concrete KP-ABSI scheme and analyze its efficiency.

1. Introduction

Internet of Things (IoT) is enabling Smart City initiatives all over the world. Recently, IoT-based applications have been widely developed, such as smart grid and smart healthcare [1, 2]. With the growth of IoT, enormous amount of data is generated by IoT-enabled devices. They need to be stored, processed, and accessed. Thus, Alessio et al. firstly merged cloud and IoT to introduce a new paradigm named CloudIoT to solve the issues [3]. For cloud-based IoT, the research on the security and privacy for the big data of IoT is a hot spot.

Private set intersection (PSI), firstly proposed by [4], is a special case of secure multiparty computation. It enables two parties to calculate the set intersection of their data sets under the condition of privacy preservation. It is applied to many practical scenarios, such as IoT and internet-based personal health record (PHR) systems. In traditional PSI solutions, the data users hold their own data sets. However, in CloudIoT computing, data users (i.e., IoT-enabled devices) with limited computing power and storage resources would like to outsource their data sets to the cloud. For confidentiality and privacy, data sets should be encrypted before outsourcing. Cloud service providers provide flexible services to fulfill cloud users' demand.

Based on this, we research on PSI over outsourced encrypted data sets in the CloudIoT system. In this scenario, the data users (i.e., the IoT-enabled devices) will encrypt and outsource their data sets to the cloud and then delegate the cloud to perform the set intersection. It has been studied by some works [5–7]. But it indeed raises a concern on how to enforce fine-grained access control for limiting the cloud's capability on computing set intersection. For this, Mohammad Ali et. al combined ciphertext-policy attribute-based encryption and private set intersection to propose an attribute-based private set intersection [8]. However, in their solution, the data user, who requests the set intersection operation, should hold the data sets in plaintext form. It does not really focus on outsourced encrypted data sets. Besides, there is still no solution for key-policy setting.

In this paper, we firstly combined key-policy attribute-based encryption and private set intersection to introduce a novel concept called *delegated key-policy attribute-based set intersection over outsourced encrypted data sets* (KP-ABSI). KP-ABSI focuses on the problem of set intersection over outsourced data sets in the cloud paradigm. It allows data owners to specify some attributes set on his/her data set and encrypt it before outsourcing, respectively. A data user with proper access control policy (satisfied by the attribute set specified by the data owner and himself/herself) can generate a token to delegate the cloud sever to perform the set intersection over his/her and the data owner's outsourced encrypted data sets. We formally give the definition and security notion for KP-ABSI and propose a concrete construction.

Our KP-ABSI scheme has three distinctive properties: (1) Our solution realizes fine-grained authorization for set intersection over encrypted outsourced data sets by combining KP-ABE and PSI. (2) The cloud server cannot obtain any information about the plaintexts beyond the result of set intersection, which is also with the form of ciphertexts. (3) Compared with existing PSI schemes, our schemes do not require interaction with the data owner or the trusted authority.

2. Related Work

Although the scholars have carried out extensive research on PSI, the existing solutions cannot solve the problems considered in this paper. In the following part, we will briefly introduce the related works. In general, they can be divided into three categories as follows.

> *Two-Party Private Set Intersection*. The traditional PSI has two participants, a data owner and a data user. Both of them hold their own data sets and interactively compute the set intersection [9, 10]. However, two-party PSI does not apply to cloud computing because two parties must hold their data sets by themselves.
>
> *Three-Party Private Set Intersection*. Typically, three-party PSI involves three participants: a data owner, a data user, and the cloud server. The data user and the data owner would like to outsource their data sets to the cloud and delegate set intersection computation to the cloud. [5, 7, 11]. Moreover, public key encryption with equality test [12–15] can also be used to attain this goal. However, in these solutions, there is not any authorization mechanism and the data owner online is required to authorize the data user. So, they are not practical in the cloud computing.
>
> *Attribute-Based Encryption*. ABE, which is introduced by Sahai and Waters, achieves fine-grained access control for outsourced data [16]. There are two variants of ABE: key-policy attribute-based encryption (KP-ABE) where the decryption key is associated with the access control policy (e.g., [17–19]) and ciphertext-policy attribute-based encryption (CP-ABE) where the ciphertext is associated with the access control policy (e.g., [20–22]). In 2017, Zhu et al. presented a key-policy attribute-based encryption with equality test, which can be utilized to do the set intersection over outsourced encrypted data set of one element. After that, Wang et al. proposed the first ciphertext-policy attribute-based encryption with equality test scheme [23, 24]. Later, Cui et al. improved its efficiency [25]. However, attribute-based encryption with equality test is only for one element. For this, in 2020, Mohammad Ali et. al firstly combined CP-ABE and PSI to propose an attribute-based set intersection scheme [8]. It achieves fine-grained access control for set intersection computation. Unfortunately, their solution requires the data user to hold his/her data set in the plaintext form. It did not really focus on outsourced encrypted data sets in the cloud computing. Moreover, there are no key-policy setting solutions for attribute-based set intersection.

Thus, in this paper, we combine KP-ABE with PSI to introduce a novel primitive-delegated key-policy attribute-based set intersection over outsourced encrypted data sets (KP-ABSI). For fairness, we summarize the properties of KP-ABSI scheme in Table 1.

3. Problem Formulation

3.1. System Model. The system model for KP-ABSI is shown in Figure 1. There are three participants: the trusted attribute authority, the cloud users (e.g., data owner Alice, authorized data user Bob, and unauthorized data user Carlos), and the cloud server. The trusted attribute authority primarily initiates the public parameters and issues private keys for data users according to their access control polices. Cloud server provides powerful storage and computing services for cloud users. The cloud users outsource their private data sets to the cloud server. Specifically, a cloud user, Alice, outsources her data set to the cloud in encrypted form, where the encryption is conducted according to some attribute set UAtt_A. An authorized user, Bob, whose access control policy is satisfied by the attribute set UAtt_A, can delegate to the cloud the computation of set intersection between Alice's outsourced encrypted sets and his own outsourced encrypted sets (Bob naturally has the private key to decrypt his own outsourced encrypted data). Meanwhile, any unauthorized user, Carlos, is neither able to decrypt Alice's outsourced encrypted data sets nor able to delegate the cloud to perform the set intersection operation.

In this model, we assume that the cloud is semitrusted (i.e., honest-but-curious), which means that the cloud honestly executes the protocol for two honest users, but tries to learn useful information beyond the ciphertexts through set intersection operations. Cloud users may be malicious and may collude with each other. We even allow a malicious user, say Bob, to collude with the semitrusted cloud. However, in this case, we cannot require that the cloud be not able to decrypt the honest user's, say Alice, ciphertext data set when the malicious and colluding user, Bob, has the private key for decrypting Alice's data set (e.g., Bob can simply give his private key to the cloud).

3.2. Functional Definition. In this part, we introduce the formal definition for delegated key-policy attribute-based set intersection over outsourced encrypted data sets (KP-ABSI),

Table 1: Property summary for PSI solutions in the literature and KP-ABSI in this paper. PSI delegation means that the data owners can delegate set intersection operations to the cloud. Outsourced encrypted data set means it can do PSI over outsourced encrypted data sets in the cloud. Fine-grained authorization means that it supports attribute-based access control policy.

Schemes	PSI delegation	Outsourced encrypted data sets	Fine-grained authorization
Two-party PSI [9, 10]	×	×	×
Three-party PSI [5, 12]	✓	✓	×
AB-PSI [8]	✓	×	✓
KP-ABSI	✓	✓	✓

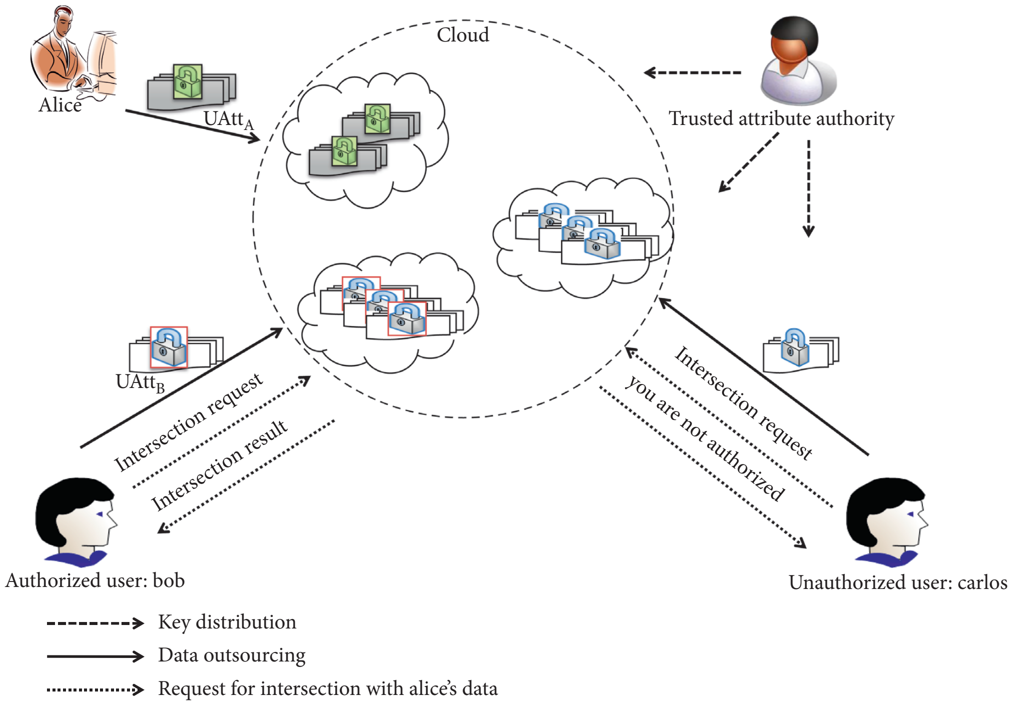


Figure 1: System model of delegated key-policy attribute-based set intersection over outsourced encrypted data sets.

where private keys are associated with access control policies. For convenience, we denote by UAtt an attribute set and by T an access policy in KP-ABSI. Let $F(\text{UAtt}, T) = 1$ if and only if UAtt satisfies T in KP-ABSI.

Definition 1. Delegated key-policy attribute-based set intersection over outsourced encrypted data sets (KP-ABSI) includes five algorithms as follows:

$(\text{pk}, \text{mk}) \longleftarrow \text{Setup}(1^{\ell})$: Given a security parameter ℓ as input, the trusted attribute authority initializes the system public parameters pk and the master secret key mk.

$\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$: Given the master secret key mk and an access control policy T, the trusted attribute authority issues private keys sk for a data user.

$\text{cph} \longleftarrow \text{Enc}(D, \text{UAtt})$: Given an attribute set UAtt, a data user encrypts his/her private data set D to the ciphertext cph. The resulting ciphertexts will be outsourced to the cloud.

$\text{tkn} \longleftarrow \text{TokenGen}(\text{sk})$: With his/her private key sk, the data user generates a token tkn and delegates the set intersection computation to the cloud.

$\text{rslt} \longleftarrow \text{SI}(\text{tkn}, \text{cph}, \text{cph}')$: The cloud utilizes tkn to compute, on behalf of two data users, the set intersection rslt only if the access control policy T corresponding to tkn satisfies both $F(\text{UAtt}, T) = 1$ and $F(\text{UAtt}', T) = 1$, where the attribute sets UAtt and UAtt′ are, respectively, specified by cph and cph′. We say that a KP-ABSI scheme is correct if the following holds: Given $(\text{pk}, \text{mk}) \longleftarrow \text{Setup}(1^{\ell})$, $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$, $\text{tkn} \longleftarrow \text{TokenGen}(\text{sk})$, and $\text{cph} \longleftarrow \text{Enc}(D, \text{UAtt})$ for set D and $\text{cph}' \longleftarrow \text{Enc}(D', I_{\text{Enc}}')$ for set D', if $F(\text{UAtt}, T) = 1$ and $(\text{UAtt}', T) = 1$, then rslt is the encrypted form of set intersection $D \cap D'$, where $\text{rslt} \longleftarrow \text{SI}(\text{cph}, \text{cph}', \text{tkn})$.

3.3. Security Definitions. The security for KP-ABSI can be expressed by the three properties as follows.

3.3.1. Selective Security against Chosen-Plaintext Attack. It indicates that a probabilistic polynomial-time (PPT) adversary $\mathscr{A}$, without being given the corresponding tokens, is not able to obtain any useful information about the encrypted data sets. Notice that "selective" means that adversary $\mathscr{A}$ should choose a target attribute set UAtt^* which it wants to challenge before the public parameters are generated. The security definition for selective security against chosen-plaintext attack can be formalized via the following game between an adversary $\mathscr{A}$ and a challenger.

Setup: $\mathscr{A}$ selects a target attribute set UAtt^* and sends it to the challenger. The challenger runs Setup algorithm to initialize pk and mk, sends pk to $\mathscr{A}$, and sets mk as the master private key.

Phase 1: The adversary $\mathscr{A}$ can make polynomial queries for the following oracles:

$\mathcal{O}_{\text{KeyGen}}(T)$: If $F(\text{UAtt}^*, T) = 1$, the challenger aborts; otherwise, the challenger returns $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, \text{pk}, T)$ to $\mathscr{A}$.

$\mathcal{O}_{\text{TokenGen}}(T)$: If $F(\text{UAtt}^*, T) = 1$, the challenger aborts; otherwise, the challenger calculates $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$ and returns $\text{tkn} \longleftarrow \text{TokenGen}(\text{sk})$ to $\mathscr{A}$.

Challenge: The adversary $\mathscr{A}$ randomly gives two data sets D_0 and D_1, where $|D_0| = |D_1|$ but $D_0 \neq D_1$, to the challenger. Then, the challenger picks $\sigma \xleftarrow{R} \{0, 1\}$ at random, builds the challenge ciphertext $\text{cph}^* = \text{Enc}(D_\sigma, \text{UAtt}^*)$, and sends cph^* to $\mathscr{A}$.

Phase 2: Same as Phase 1.

Guess: $\mathscr{A}$ eventually outputs a guess σ' of σ. If $\sigma = \sigma'$, we say that $\mathscr{A}$ wins the game.

Definition 2. We say that a KP-ABSI scheme is selective secure against chosen-plaintext attack, if any PPT adversary $\mathscr{A}$ wins the above game with a negligible advantage, where the advantage can be described as $|\Pr[\sigma' = \sigma] - 1/2|$.

3.3.2. One-Way Security against Chosen-Plaintext Attack. It says that a PPT adversary $\mathscr{A}$, even given an appropriate token, cannot obtain the plaintexts corresponding to the ciphertexts. Note that the term "appropriate" means that the access control policy that generates the token is satisfied by the attribute set associated with the target ciphertext. Of course, $\mathscr{A}$ can choose a plaintext data set of its choice, encrypt it with public keys, and then utilize the token to check whether or not the target ciphertext is equal to the ciphertext of his choice. In other words, this type of brute-force attack is inherent to the set intersection problem and we can only demand that $\mathscr{A}$ cannot have any attack strategy significantly better than the brute-force attack, as captured by this property via the following game between an adversary $\mathscr{A}$ and a challenger.

Setup: The challenger runs Setup to initialize (pk, mk), sends pk to $\mathscr{A}$, and sets mk as the master private key.

Phase 1: The adversary $\mathscr{A}$ can make polynomial queries for the following oracles. Meanwhile, the challenger maintains a list L_T, which is initially empty.

$\mathcal{O}_{\text{KeyGen}}(T)$: The challenger returns $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$ to $\mathscr{A}$ and records T to L_T.

$\mathcal{O}_{\text{TokenGen}}(T)$: The challenger calculates $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$ and returns $\text{tkn} \longleftarrow \text{TokenGen}(\text{sk})$ to $\mathscr{A}$.

Challenge: $\mathscr{A}$ gives a target attribute set UAtt^* to the challenger, where, $\forall T \in L_T$, $F(\text{UAtt}^*, T) = 0$. The challenger selects an access control T^* such that $F(\text{UAtt}^*, T^*) = 1$, picks D^* uniformly at random, runs $\text{cph}^* \longleftarrow \text{Enc}(D^*, \text{UAtt}^*)$ and $\text{tkn}^* \longleftarrow \text{TokenGen}(\text{KeyGen}(T^*))$, and returns $\text{cph}^*, \text{tkn}^*$ to $\mathscr{A}$.

Phase 2: $\mathscr{A}$ executes the same as in Phase 1, except that $F(\text{UAtt}^*, T) = 0$ when querying $\mathcal{O}_{\text{KeyGen}}(T)$.

Guess: $\mathscr{A}$ outputs a guess d. If $d \in D^*$, we say that $\mathscr{A}$ wins the game.

Definition 3. We say that a KP-ABSI scheme achieves one-way security against chosen-plaintext attack if, for any PPT adversary $\mathscr{A}$, the advantage of $\mathscr{A}$ winning the game is negligible, where the advantage is defined as $|\Pr[d \in D^*] - (m|D^*|/|\text{Msg}|)|$, where m is the number of guess/brute-force attacks $\mathscr{A}$ makes, and Msg is the message space of set elements.

3.3.3. Fine-Grained Authorization Security. This property says that the cloud is unable to utilize the given tokens to conduct set intersection over ciphertexts if no access control policy (associated with the data user's private key) that generates the tokens is satisfied by both of the attribute sets associated with the two ciphertexts. More specifically, consider the token tkn_1 that can be used to conduct set intersection over ciphertexts cph_1 and cph_2 and the token tkn_2 that can be used to conduct set intersection over ciphertexts cph_2 and cph_3. If the access control policy that is used to generate tkn_1 is not satisfied by the attribute set associated with ciphertext cph_3, and the access control policy that is used to generate tkn_2 is not satisfied by the attribute set associated with ciphertext cph_1; then the cloud cannot do the set intersection computation over ciphertexts cph_1 and cph_3 by using tkn_1 and/or tkn_2. The definition for fine-grained authorization security can be described via a game between an adversary $\mathscr{A}$ and a challenger.

Setup: The challenger runs Setup to initialize (pk, mk), sends pk to $\mathscr{A}$, and sets mk as the master secret key.

Phase 1: The adversary $\mathscr{A}$ makes polynomial queries for the following oracles. Meanwhile, the challenger maintains two lists L_T and L_{tkn}, which are empty initially.

$\mathcal{O}_{\text{KeyGen}}(T)$: The challenger returns $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$ to $\mathscr{A}$ and records T to L_T.

$\mathcal{O}_{\text{TokenGen}}(T)$: The challenger runs $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, T)$ and $\text{tkn} \longleftarrow \text{TokenGen}(\text{sk})$, returns tkn back to $\mathscr{A}$, and records T to L_{tkn}.

Challenge: $\mathscr{A}$ gives two target attribute sets UAtt_1^* and UAtt_2^* to the challenger. Then the challenger chooses two data sets D_0, D_1, picks a bit $\sigma \xleftarrow{R} \{0, 1\}$ randomly, runs $\mathrm{cph}_1^* \longleftarrow \mathrm{Enc}(D_0, \mathrm{UAtt}_1^*)$ and $\mathrm{cph}_2^* \longleftarrow \mathrm{Enc}(D_\sigma, \mathrm{UAtt}_2^*)$, and returns $\mathrm{cph}_1^*, \mathrm{cph}_2^*$ to $\mathscr{A}$. Here, we require that

$\forall T \in L_T$, $F(\mathrm{UAtt}_1^*, T)$ and $F(\mathrm{UAtt}_2^*, T)$ do not output 1 simultaneously;
$\forall T \in L_{\mathrm{tkn}}$, $F(\mathrm{UAtt}_1^*, T)$ and $F(\mathrm{UAtt}_2^*, T)$ do not output 1 simultaneously.

Phase 2: $\mathscr{A}$ executes the same as in Phase 1, except for the following:

When querying $\mathscr{O}_{\mathrm{KeyGen}}(T)$, $F(\mathrm{UAtt}_1^*, T)$ and $F(\mathrm{UAtt}_2^*, T)$ do not output 1 simultaneously.
When querying $\mathscr{O}_{\mathrm{TokenGen}}(T)$, $F(\mathrm{UAtt}_1^*, T)$ and $F(\mathrm{UAtt}_2^*, T)$ do not output 1 simultaneously.

Guess: The adversary $\mathscr{A}$ eventually outputs a guess σ'. If $\sigma = \sigma'$, we say that $\mathscr{A}$ wins the game.

Definition 4. If, for any PPT adversary $\mathscr{A}$, the advantage of $\mathscr{A}$ winning the game is negligible, where the advantage can be expressed as $|\Pr[\sigma' = \sigma] - (1/2)|$, we say that a KP-ABSI scheme achieves fine-grained authorization security.

4. Scheme Construction

4.1. Basic Idea. To illustrate the idea, let a data user's private key be (g^{at}, g^{bt}), which can be generated by running $\{q_v(0) \mid v \in \mathrm{lvs}(T)\} \longleftarrow \mathrm{Share}(T, \mathrm{abt})$ for some random t and setting $(g^{q_v(0)} H_1(\mathrm{att}(v))^{t_v}, g^{t_v})$, where t_v is a random number with respect to leaf $v \in \mathrm{lvs}(T)$. A set element d is encrypted into two parts:

The first part is related to d; namely,

$$\left(g^{br_1}, g^{a(r_1+r_2)} H_2(d), g^{r_2}\right), \qquad (1)$$

where g is a generator of G, r_1 and r_2 are two random numbers, H_1, H_2 are two hash functions, and a, b are private keys.

The second part is related to the attribute set corresponding to the access control policy in question, namely, $H_1(at_i)^{r_2}$ for $at_i \in \mathrm{UAtt}$.

A data user can generate the token as $\{g^{atk}, g^{btk}, (g^{q_v(0)k} H_1(\mathrm{att}(v))^{kt_v}, g^{kt_v})\}$, by which the cloud is able to translate the ciphertext into an intermediate form $e(H_2(d), g)^{btk}$ once the attribute set UAtt satisfies the access control policy T.

4.2. KP-ABSI Construction

Setup(1^ℓ): Given the security parameter ℓ as input, the public parameters and the master secret key can be generated as follows:

Let $(e, q, g, g_T, G, G_T) \longleftarrow \mathrm{BMapGen}(1^\ell)$.

Let $H_1\colon \{0,1\}^* \longrightarrow G$ and $H_2\colon \{0,1\}^* \longrightarrow G$ be two secure hash functions that are modeled as random oracles.

Select $a, b \xleftarrow{R} \mathbb{Z}_p$ and set the public parameters and the master secret key as

$$\begin{aligned} \mathrm{pk} &= \left(e, G, G_T, g, g^a, g^b, H_1, H_2\right), \\ \mathrm{mk} &= (a, b). \end{aligned} \qquad (2)$$

KeyGen(mk, T): Given access tree T, this algorithm selects $t \xleftarrow{R} \mathbb{Z}_p$, computes $X_1 = g^{at}$ and $X_2 = g^{bt}$, and runs $\{q_v(0) \mid v \in \mathrm{lvs}(T)\} \longleftarrow \mathrm{Share}(T, abt)$. Then, for each leaf $v \in \mathrm{lvs}(T)$, the algorithm selects $t_v \xleftarrow{R} \mathbb{Z}_p$ and sets $Y_v = g^{q_v(0)} H_1(\mathrm{att}(v))^{t_v}$ and $Z_v = g^{t_v}$. The secret key is

$$\mathrm{sk} = \left(T, X_1, X_2, \{(Y_v, Z_v) \mid v \in \mathrm{lvs}(T)\}\right). \qquad (3)$$

Enc(D, UAtt): Given set for outsourcing, $D = \{d_0, \ldots, d_n\}$, this algorithm encrypts the set as follows: for each d_j, it selects $r_1, r_2 \xleftarrow{R} \mathbb{Z}_p$, sets $A_1 = g^{br_1}$, $A_2 = g^{a(r_1+r_2)} H_2(d_j)$, and $A_3 = g^{r_2}$, and computes $B_i = H_1(at_i)^{r_2}$ for each $at_i \in \mathrm{UAtt}$. The ciphertext of d_j is

$$\mathrm{cph}_j = \left(\mathrm{UAtt}, A_1, A_2, A_3, \{B_i \mid at_i \in \mathrm{UAtt}\}\right). \qquad (4)$$

The set of ciphertexts is $\mathrm{cph} = \{\mathrm{cph}_0, \ldots, \mathrm{cph}_n\}$.

TokenGen(sk): Given secret key sk, this algorithm selects $k \xleftarrow{R} \mathbb{Z}_p$, sets $\widehat{X}_1 = X_1^k = g^{atk}$ and $\widehat{X}_2 = X_2^k = g^{btk}$, and computes $\widehat{Y}_v = Y_v^k$ and $\widehat{Z}_v = Z_v^k$ for leaf $v \in \mathrm{lvs}(T)$. The token is

$$\mathrm{tkn} = \left(T, \widehat{X}_1, \widehat{X}_2, \left\{\left(\widehat{Y}_v, \widehat{Z}_v\right) \mid v \in \mathrm{lvs}(T)\right\}\right). \qquad (5)$$

SI$(\mathrm{tkn}, \mathrm{cph}, \mathrm{cph}')$: Given $\mathrm{cph} = \{\mathrm{cph}_0, \ldots, \mathrm{cph}_n\}$, $\mathrm{cph}' = \{\mathrm{cph}_0', \ldots, \mathrm{cph}_m'\}$, and tkn, this algorithm is executed as follows:

Given $\mathrm{cph}_j \in \mathrm{cph}$, it selects an attribute set $S \in \mathrm{UAtt}$ satisfying T. If S does not exist, it returns 0. Otherwise, it computes

$$E_v = \frac{e\left(\widehat{Y}_v, A_3\right)}{e\left(\widehat{Z}_v, B_i\right)} = e(g,g)^{q_v(0)kr_2}, \qquad (6)$$

for all $v \in \mathrm{lvs}(T)$ with $\mathrm{att}(v) = at_i$, computes

$$\begin{aligned} E_{\mathrm{root}} &= e(g,g)^{abtkr_2} \longleftarrow \mathrm{Combine}(T, E_v \mid \mathrm{att}(v) \in S), \\ E_1 &= e\left(A_1, \widehat{X}_1\right) = e(g,g)^{abtkr_1}, \\ E_{2j} &= e\frac{\left(A_2, \widehat{X}_2\right)}{\left(E_{\mathrm{root}} E_1\right)} = e\left(H_2\left(d_j\right), g\right)^{btk}, \end{aligned} \qquad (7)$$

and sets $E = \{E_{20}, \ldots, E_{2n}\}$.

Given $\mathrm{cph}'_j \in \mathrm{cph}'$, it selects an attribute set $S' \in \mathrm{UAtt}'$ satisfying T. If S' does not exist, it returns 0. Otherwise, it computes $E'_v = e(\widehat{Y}_v, A'_3)/e(\widehat{Z}_v, B_i) = e(g,g)^{q_v(0)kr_2\prime}$ for all $v \in \mathrm{lvs}(T)$ with $\mathrm{att}(v) = at_i$, computes

$$E'_{\mathrm{root}} = e(g,g)^{abtkr'_2} \longleftarrow \mathrm{Combine}(T, E'_v \mid \mathrm{att}(v) \in S'),$$
$$E'_1 = e\left(A'_1, \widehat{X}_1\right) = e(g,g)^{abtkr'_1},$$
$$E_{2j}' = \frac{e\left(A_2, \widehat{X}_2\right)}{\left(E'_{\mathrm{root}}E'_1\right)} = e\left(H_2\left(d'_j\right), g\right)^{btk}, \tag{8}$$

and sets $E' = \{E_{20}', \ldots, E_{2m}'\}$.

Output the set intersection $\mathrm{rslt} = \left\{\mathrm{cph}_j \mid \mathrm{cph}_j \in \mathrm{cph}, E_{2j} \in E \cap E'\right\}$.

The correctness of the above KP-ABSI scheme can be verified by following the protocol. In what follows, we analyze its security.

4.3. Security Analysis

Theorem 1. *Under the DLN assumption, the above KP-ABSI scheme achieves the selective security against chosen-plaintext attacks in the random oracle as specified in Definition 2.*

Firstly, we prove that our scheme achieves the security goal when $|D| = 1$ (i.e., the message space has a single element) and then extends the proof to the case $|D| > 1$.

Proof. We show that if there is a PPT adversary $\mathscr{A}$ that wins the selective security game with a nonnegligible advantage μ, then a challenger that can solve the DLN problem with the advantage at least $\mu/2$ can be constructed. Specifically, given a DLN instance $(g, h, f, f^{r_1}, g^{r_2}, Q)$, where $g, f, h, Q \xleftarrow{R} G$ and $r_1, r_2 \xleftarrow{R} \mathbb{Z}_p$ are unknown, the game is simulated by the challenger as follows.

Setup: The adversary $\mathscr{A}$ gives an attribute set UAtt^* to the challenger. Then the challenger produces the bilinear map $e\colon G \times G \longrightarrow G_T$, constructs $g^b = f$ and $g^a = h$ with a and b unknown, and sets $\mathrm{pk} = (e, g, f, h, G, G_T)$. The challenger sends pk to $\mathscr{A}$. The challenger maintains two lists $\mathrm{List}_{H_1}(at_i, \alpha_i, \beta_i)$ and $\mathrm{List}_{H_2}(d, \gamma)$, which are initially empty. $\mathscr{A}$ can query $\mathscr{O}_{H_1}$ and $\mathscr{O}_{H_2}$ polynomially many times as follows:

$\mathscr{O}_{H_1}(at_i)$: Given attribute at_i, the challenger responds as follows:

The case at_i was queried before: it retrieves α_i, β_i from List_{H_1} and returns $f^{\alpha_i}g^{\beta_i}$ to $\mathscr{A}$.

The case at_i was not queried before: if $at_i \in \mathrm{UAtt}^*$, it selects $\beta_i \xleftarrow{R} \mathbb{Z}_p$, adds $(at_i, \alpha_i = 0, \beta_i)$ to List_{H_1}, and returns g^{β_i} to $\mathscr{A}$; otherwise, it selects $\alpha_i, \beta_i \xleftarrow{R} \mathbb{Z}_p$, adds $(at_i, \alpha_i, \beta_i)$ to List_{H_1}, and returns $f^{\alpha_i}g^{\beta_i}$ to $\mathscr{A}$.

$\mathscr{O}_{H_2}(d)$: Given a message d as input, if d was queried before, it retrieves γ from List_{H_2} and returns g^{γ} to $\mathscr{A}$; otherwise, it picks $\gamma \xleftarrow{R} \mathbb{Z}_p$ randomly, records (d, γ) to List_{H_2}, and returns g^{γ} to $\mathscr{A}$.

Phase 1: The adversary $\mathscr{A}$ makes polynomial queries for the following oracles as follows:

$\mathscr{O}_{\mathrm{KeyGen}}(T)$: If $F(\mathrm{UAtt}^*, T) = 1$, the simulation is aborted; otherwise, the challenger produces sk according to the two following procedures:

$\mathrm{PolySat}(T_v, \mathrm{UAtt}^*, \lambda_v)$: Given a secret value λ_v, this procedure builds the polynomial for each node of subtree T_v, where $F(\mathrm{UAtt}^*, T_v) = 1$. Suppose that the threshold value of node v is k_v; it lets $q_v(0) = \lambda_v$ and chooses $k_v - 1$ coefficients uniformly at random to uniquely determine the polynomial q_v. Then it recursively runs $\mathrm{PolySat}(T_{v'}, \mathrm{UAtt}^*, \lambda_{v'})$ to build the polynomial for each child node v' of v, by letting $\lambda_{v'} = q_v(\mathrm{index}(v'))$.

$\mathrm{PolyUnsat}(T_v, \mathrm{UAtt}^*, g^{\lambda_v})$: Given an element $g^{\lambda_v} \in G$, where λ_v is unknown to the challenger, this procedure is to build the polynomial for each node of subtree T_v, where $F(\mathrm{UAtt}^*, T_v) = 0$. Assume that the threshold value of node v is k_v and V is the set of children of node v such that, $\forall v' \in V, F(\mathrm{UAtt}^*, T_{v'}) = 1$. Since $F(\mathrm{UAtt}^*, T_v) = 0$, we have $|V| < k_v$. For each $v' \in V$, it selects $\lambda_{v'} \xleftarrow{R} \mathbb{Z}_p$ and sets $\lambda_{v'} = q_v(\mathrm{index}(v'))$. It then determines the other $k_v - |V|$ points of polynomial q_v such that $g^{q_v(0)} = g^{\lambda_v}$. For each child node v' of node v, it executes the following:

(i) If v' is a node with $F(at^*, T_{v'}) = 1$, it runs $\mathrm{PolySat}(T_{v'}, \mathrm{UAtt}^*, q_v(\mathrm{index}(v')))$, where $q_v(\mathrm{index}(v'))$ is known.

(ii) If v' is a node with $F(at^*, T_{v'}) = 0$, it runs $\mathrm{PolySat}(T_{v'}, \mathrm{UAtt}^*, g^{\lambda_{v'}})$, where $g^{\lambda_{v'}} = g^{q_v(\mathrm{index}(v'))}$ is known.

Based on the two procedures above, the challenger executes $\mathrm{PolyUnsat}(T, \mathrm{UAtt}^*, g^a)$ by implicitly defining $q_{\mathrm{root}}(0) = a$. Note that, for each $v \in \mathrm{lvs}(T)$, the challenger knows $q_v(0)$ if $\mathrm{att}(v) \in \mathrm{UAtt}^*$ and knows $g^{q_v(0)}$ otherwise. Therefore, it generates credentials as follows by selecting $t \xleftarrow{R} \mathbb{Z}_p$ and setting $X_1 = h^t$ and $X_2 = f^t$ (while noting that $btq_v(0)$ is the secret share of abt):

If $\text{att}(v) = at_j$ for some $at_j \in \text{UAtt}^*$, it selects $t_v \xleftarrow{R} \mathbb{Z}_p$ and sets $Y_v = f^{tq_v(0)} g^{\beta_j t_v} = g^{btq_v(0)} H_1(\text{att}(v))^{t_v}$ and $Z_v = g^{t_v}$. If $\text{att}(v) \notin \text{UAtt}^*$, where $\text{att}(v) = at_j$ for some j, it selects $t_v' \xleftarrow{R} \mathbb{Z}_p$ and sets $Y_v = f^{tt_v{\prime}\alpha_j} (g^{q_v(0)})^{-(\beta_j/\alpha_j)} g^{t_v{\prime}\beta_j}$ and $Z_v = g^{-(q_v(0)/\alpha_j)} g^{t_v{\prime}}$. Note that (Y_v, Z_v) is valid because the challenger implicitly sets $t_v = -(q_v(0)/\alpha_j) + t_v'$ and the following:

$$\begin{aligned} Y_v &= f^{tt_v'\alpha_j} \left(g^{q_v(0)}\right)^{-(\beta_j/\alpha_j)} g^{t_v'\beta_j} = f^{tq_v(0)} \left(f^{\alpha_j} g^{\beta_j}\right)^{-(q_v(0)/\alpha_j)+t_v'} \\ &= f^{tq_v(0)} H_1\left(at_j\right)^{t_v}, \\ Z_v &= g^{-(q_v(0)/\alpha_j)} g^{t_v'} = g^{t_v}. \end{aligned} \tag{9}$$

$\mathcal{O}_{\text{tkn}}(T)$: The challenger queries $\mathcal{O}_{\text{KeyGen}}(T)$ to obtain sk and returns $\text{tkn} \longleftarrow \text{TokenGen}(\text{sk})$ to $\mathcal{A}$.

Challenge: The adversary $\mathcal{A}$ gives two messages $D_0 = \{d_0\}$ and $D_1 = \{d_1\}$ of equal length to the challenger. Then the challenger randomly picks $\sigma \xleftarrow{R} \{0,1\}$, encrypts d_σ to $\text{cph}^* = (\text{UAtt}^*, f^{r_1}, QH_2(d_\sigma), g^{r_2}, \{(g^{r_2})^{\beta_j} \mid at_j \in \text{UAtt}^*\})$, and sends cph^* to $\mathcal{A}$.

Phase 2: $\mathcal{A}$ executes the same as in the above Phase 1.

Guess: The adversary $\mathcal{A}$ will eventually output a guess σ' of σ. If $\sigma' = \sigma$, the challenger outputs $Q = h^{r_1+r_2}$; otherwise, it outputs $Q \neq h^{r_1+r_2}$.

The simulation is completed. In Challenge phase, if $Q = h^{r_1+r_2}$, then cph^* is indeed a valid ciphertext of D_σ and the probability that $\mathcal{A}$ outputs $\sigma' = \sigma$ is $(1/2) + \mu$. Otherwise, if Q is a random element from G, then cph^* is a random group element and the probability that $\mathcal{A}$ outputs $\sigma' = \sigma$ is $(1/2)$. In conclusion, the probability that the challenger correctly guesses $Q \overset{?}{=} h^{r_1+r_2}$ is $(1/2)((1/2) + (1/2) + \mu) = (1/2) + (\mu/2)$. That is, if $\mathcal{A}$ wins the game with the advantage μ, then the challenger solves the DLN problem with the advantage $(\mu/2)$.

So far, we have shown that our KP-ABSI scheme is selective secure against chosen-plaintext attack when $|D| = 1$. In what follows, we prove that our KP-ABSI scheme achieves the selective security against chosen-plaintext attack for the general case of $|D| > 1$.

Suppose that $\mathcal{A}$ gives two sets $D_0 = (d_0, \ldots, d_n)$ and $D_1 = (d_0', \ldots, d_n')$. Denote by $\text{cph}^{(i)}$ the encryption of $(d_0, \ldots, d_i, d_{i+1}', \ldots, d_n')$, meaning that $\text{cph}^{(n)}$ is the encryption of D_0 and $\text{cph}^{(-1)}$ is the encryption of D_1. The Challenge phase is extended to accommodate an additional adversary $\mathcal{A}'$ as follows:

$\mathcal{A}'$ picks a random index $i \xleftarrow{R} [0, n]$ and presents (d_i, d_i') to the challenger. Then the challenger sends back cph_i to $\mathcal{A}'$ by encrypting d_i if $\sigma = 0$ or returns d_i' if $\sigma = 1$.

$\mathcal{A}'$ encrypts $(d_0, \ldots, d_{i-1})$ and $(d_{i+1}', \ldots, d_n')$ and returns $(\text{cph}_0, \ldots, \text{cph}_n)$ to $\mathcal{A}$. $\mathcal{A}'$ outputs $\mathcal{A}$'s output σ'.

Note that $\mathcal{A}'$ sends to $\mathcal{A}$ the ciphertext $\text{cph}^{(i)}$ if $\sigma = 0$ and the ciphertext $\text{cph}^{(i-1)}$ if $\sigma = 1$. Denote by $\mathcal{A}(\text{cph}^{(i)})$ the guess of $\mathcal{A}$ with ciphertexts $\text{cph}^{(i)}$. Then we show the probability that $\mathcal{A}'$ wins the game. Note that

$$\begin{aligned} \Pr\left[\mathcal{A}' \text{ outputs } 0 \mid \sigma = 0\right] &= \sum_{j=0}^{n} \frac{1}{n+1} \Pr\left[\mathcal{A}\left(\text{cph}^{(j)}\right) = 0\right], \\ \Pr\left[\mathcal{A}' \text{ outputs } 0 \mid \sigma = 1\right] &= \sum_{j=0}^{n} \frac{1}{n+1} \Pr\left[\mathcal{A}\left(\text{cph}^{(j-1)}\right) = 1\right]. \end{aligned} \tag{10}$$

Therefore, the probability that $\mathcal{A}'$ wins the game is

$$\begin{aligned} &\frac{1}{2} \Pr\left[\mathcal{A}'(\text{cph}) = 0 \mid \sigma = 0\right] + \frac{1}{2} \Pr\left[\mathcal{A}'(\text{cph}) = 1 \mid \sigma = 1\right] \\ &= \sum_{j=0}^{n} \frac{1}{2(n+1)} \Pr\left[\mathcal{A}\left(\text{cph}^{(j)}\right) = 0\right] \\ &\quad + \sum_{j=0}^{n} \frac{1}{2(n+1)} \Pr\left[\mathcal{A}\left(\text{cph}^{(j-1)}\right) = 1\right] \\ &= \frac{n}{2(n+1)} + \frac{1}{2(n+1)} \Pr\left[\mathcal{A}\left(\text{cph}^{(n)}\right) = 0\right] \\ &\quad + \frac{1}{2(n+1)} \Pr\left[\mathcal{A}\left(\text{cph}^{(-1)}\right) = 1\right] \\ &= \frac{n}{2(n+1)} + \frac{1}{(n+1)} \left(\frac{1}{2} \Pr\left[\mathcal{A}\left(\text{cph}^{(n)}\right) = 0\right]\right. \\ &\quad \left. + \frac{1}{2} \Pr\left[\mathcal{A}\left(\text{cph}^{(-1)}\right) = 1\right]\right) \\ &\leq \frac{1}{2} + \varepsilon, \end{aligned} \tag{11}$$

where ε is negligible because the advantage that $\mathcal{A}'$ wins the game is negligible. Thus, the probability that $\mathcal{A}$ distinguishes cph^{-1} from cph^{n} is

$$\frac{1}{2} \Pr\left[\mathcal{A}\left(\text{cph}^{(n)}\right) = 0\right] + \frac{1}{2} \Pr\left[\mathcal{A}\left(\text{cph}^{(-1)}\right) = 1\right] \leq \frac{1}{2} + (n+1)\varepsilon. \tag{12}$$

That is, the advantage of $\mathcal{A}$ distinguishing cph^{-1} from cph^{n} is at most $(n+1)\varepsilon$. Therefore, the scheme achieves the selective security against chosen-plaintext attack when $|D| \geq 1$. □

Theorem 2. *Given one-way hash function H_2, the above KP-ABSI scheme achieves the one-way security against chosen-plaintext attack as specified in Definition 3.*

Proof. We prove this theorem by showing that if there is a PPT adversary $\mathcal{A}$ winning the one-way security game against chosen-plaintext attack with a nonnegligible advantage μ, then a challenger breaking the one-way hash function H_2 can be simulated.

Given $H_2(d^*) = y^*$, the challenger can simulate the one-way security game as follows:

Setup: The challenger randomly picks $a, b \xleftarrow{R} \mathbb{Z}_p$, produces $\text{pk} = (e, G, G_T, g, p, g^a, g^b)$, $\text{mk} = (a, b)$, and sends pk to $\mathscr{A}$.

Phase 1: The challenger maintains a list L_T, which is empty initially. $\mathscr{A}$ makes polynomial queries for the following oracles:

$\mathscr{O}_{\text{KeyGen}}(T)$: Given an access control policy T, it returns $\text{sk} \longleftarrow \text{KeyGen}(\text{mk}, \text{pk}, T)$ to $\mathscr{A}$ and adds T to L_T.

$\mathscr{O}_{\text{tkn}}(T)$: Given an access control policy T, it runs $\text{tkn} \longleftarrow \text{TokenGen}(\text{KeyGen}(\text{mk}, \text{pk}, T), \text{pk})$ and returns tkn to $\mathscr{A}$.

Challenge: $\mathscr{A}$ gives an attribute set UAtt^* to the challenger, where, $\forall T \in L_T$, $F(\text{UAtt}^*, T) = 0$. The challenger picks $j \xleftarrow{R} \mathbb{Z}_p$ and sets a data set $D^* = (d_0, d_1, \ldots, d_{|D^*|-1})$, where $d_0, \ldots, d_{j-1}, d_{j+1}, \ldots, d_{|D^*|-1}$ are randomly chosen from the message space Msg and d_j is implicitly set as $d_j = d^*$ and generates the ciphertext $\text{cph} = \{\text{cph}_k\}_{k \in [0, |D^*|-1]}$ as follows:

If $k \neq j$, cph_k is generated the same as in the real construction.
If $k = j$,

$$\begin{aligned}\text{cph}_k = \Big(&\text{UAtt}^*, A_1 = g^{br_1}, A_2 = g^{a(r_1+r_2)} y^*, A_3 \\ &= g^{r_2}, B_i = \left\{H_1(at_i)^{r_2} \mid at_i \in \text{UAtt}^*\right\}\Big),\end{aligned} \tag{13}$$

by randomly choosing $r_1, r_2 \xleftarrow{R} \mathbb{Z}_p$ and implicitly setting $d_j = d^*$.

The challenger chooses an access tree T^* satisfying $F(\text{UAtt}^*, T^*) = 1$, runs $\text{tkn}^* \longleftarrow \text{TokenGen}(\text{KeyGen}(\text{mk}, \text{pk}, T^*), \text{pk})$, and returns $(\text{cph}^* = (\text{cph}), \text{tkn}^*)$ to the adversary $\mathscr{A}$.

Phase 2: The adversary $\mathscr{A}$ executes the same as in Phase 1 while complying with the necessary requirements defined by the game.

Guess: $\mathscr{A}$ will eventually output a guess d to the challenger. The challenger wins if $d = d^*$.

The simulation is completed. If the probability that $\mathscr{A}$ outputs $d \in D^*$ is $|\Pr[d \in D^*] - (m|D^*|/|\text{Msg}|)| = \mu$, then $\Pr[d \in D^*] = (m|D^*|/|\text{Msg}|) + \mu$. Since the data set size is $|D^*|$, $\Pr[d\colon d = d^*] \geq (\Pr[d \in D^*]/|D^*|) = (1/|D^*|)((m|D^*|/|\text{Msg}|) + \mu)$. Therefore, if $\mathscr{A}$ wins the one-way security game against chosen-plaintext attack with a nonnegligible advantage μ, the one-way hash function H_2 can be broken by the challenger with a nonnegligible probability at least $(1/|D^*|)((m|D^*|/|\text{Msg}|) + \mu)$. □ □

Theorem 3. *The KP-ABSI achieves fine-grained authorization security in the generic bilinear group model as specified in Definition 4.*

Proof. Similar to the proof for Theorem 1, firstly, we prove that our KP-ABSI scheme achieves fine-grained authorization when the challenge size $|D_0| = |D_1| = 1$ and then extend the proof to the case of challenge size $|D_0| = |D_1| > 1$.

Setup: The challenger randomly picks $a, b \xleftarrow{R} \mathbb{Z}_p$, produces $\text{pk} = (e, G, G_T, g, p, g^a, g^b)$, and sends pk to $\mathscr{A}$. The challenger maintains two lists $\text{List}_{H_1}(at_j, \alpha_j)$ and $\text{List}_{H_2}(d, \beta)$, which are empty initially. $\mathscr{A}$ can make polynomial queries for the following $\mathscr{O}_{H_1}$ and $\mathscr{O}_{H_2}$.

$\mathscr{O}_{H_1}(at_j)$: Given an attribute at_j, if at_j was queried before, the challenger returns g^{α_j} by retrieving α_j from List_{H_1}; otherwise, the challenger picks $\alpha_j \xleftarrow{R} \mathbb{Z}_p$, records (at_j, α_j) to List_{H_1}, and returns g^{α_j} to $\mathscr{A}$.

$\mathscr{O}_{H_2}(d)$: Given a message d, if d was queried before, the challenger returns g^{β}; otherwise, the challenger picks $\beta \xleftarrow{R} \mathbb{Z}_p$, records (d, β) to List_{H_2}, and returns g^{β} to $\mathscr{A}$.

Phase 1: The challenger keeps the two lists, L_T and L_{tkn}, which are initially empty. $\mathscr{A}$ can make polynomial queries for the following oracles.

$\mathscr{O}_{\text{KeyGen}}(T)$: The challenger selects $t^{(u)} \xleftarrow{R} \mathbb{Z}_p$ and runs $\left\{q_v(0)^{(u)} \mid v \in \text{lvs}(T)\right\} \longleftarrow \text{Share}(T, abt^{(u)})$. For each node $v \in \text{lvs}(T)$, the challenger chooses $t_v^{(u)} \xleftarrow{R} \mathbb{Z}_p$ and sets

$$\text{sk} = \left(T, X_1 = g^{at^{(u)}}, X_2 = g^{bt^{(u)}}, \left\{Y_v = g^{q_v(0)^{(u)} + \alpha_j t_v^{(u)}}, Z_v = g^{t_v^{(u)}} \mid v \in \text{lvs}(T)\right\}\right), \tag{14}$$

where $\text{att}(v) = at_j$. The challenger sends sk to $\mathscr{A}$ and records T to L_T.

$\mathscr{O}_{\text{tkn}}(T)$: The challenger runs $\mathscr{O}_{\text{KeyGen}}(T)$, selects $k^{(u)} \xleftarrow{R} \mathbb{Z}_p$, and sets

Table 2: tkn_1 and tkn_2 that can be obtained by $\mathscr{A}$, where $F(\text{UAtt}_1^*, T_1) = 1$ and $F(\text{UAtt}_2^*, T_2) = 1$.

$\text{tkn}_1 = \left(T_1, X_1 = g^{atk}, X_2 = g^{btk}, \left\{Y_v = g^{k(q_v(0)+\alpha_j t_v)}, Z_v = g^{kt_v} \mid v \in \text{lvs}(T_1)\right\}\right)$
$\text{tkn}_2 = \left(T_2, X_1' = g^{at'k}, X_2' = g^{bt'k}, \left\{Y_v' = g^{k(q_v'(0)+\alpha_j t_v')}, Z_v = g^{kt_v{}'} \mid v \in \text{lvs}(T_2)\right\}\right)$

$$\begin{aligned}\text{tkn} = \Bigl(T, \Bigl\{\widehat{Y}_v = Y_v^{k^{(u)}} = g^{q_v(0)^{(u)}k^{(u)}+\alpha_j t_v^{(u)}k^{(u)}}, \widehat{Z}_v = Z_v^{k^{(u)}} = g^{t_v^{(u)}k^{(u)}} \mid v \in \text{lvs}(T)\Bigr\},\\ \widehat{X}_1 = X_1^{k^{(u)}} = g^{at^{(u)}k^{(u)}}, \widehat{X}_2 = X_2^{k^{(u)}} = g^{bt^{(u)}k^{(u)}}\Bigr),\end{aligned} \tag{15}$$

where $\text{att}(v) = at_j$. It returns tkn to $\mathscr{A}$ and adds T to L_{tkn}.

Challenge: $\mathscr{A}$ chooses two attribute sets UAtt_1^* and UAtt_2^* with the following restrictions: (1) $\forall T \in L_T$, $F(\text{UAtt}_1^*, T)$ and $F(\text{UAtt}_2^*, T)$ cannot output 1 at the same time, and, (2) $\forall T \in L_{\text{tkn}}$, $F(\text{UAtt}_1^*, T)$ and $F(\text{UAtt}_2^*, T)$ cannot output 1 at the same time. Then, $\mathscr{A}$ sends UAtt_1^* and UAtt_2^* to the challenger. The challenger chooses two sets $(D_0 = \{d_0\}, D_1 = \{d_1\})$ of equal length. For d_0, the challenger selects $r_1, r_2 \xleftarrow{R} \mathbb{Z}_p$ and computes

$$\begin{aligned}\text{cph}_1^* = \Bigl(A_1 = g^{br_1}, A_2 = g^{a(r_1+r_2)+\beta_0}, A_3\\ = g^{r_2}, \left\{B_j = g^{\alpha_j r_2} \mid at_j \in \text{UAtt}_1^*\right\}\Bigr).\end{aligned} \tag{16}$$

The challenger randomly picks $\sigma \xleftarrow{R} \{0,1\}$ and $r_3, r_4 \xleftarrow{R} \mathbb{Z}_p$ for d_σ and sets

$$\begin{aligned}\text{cph}_2^* = \Bigl(A_1' = g^{br_3}, A_2' = g^{a(r_3+r_4)+\beta_\sigma}, A_3'\\ = g^{r_4}, \left\{B_j' = g^{\alpha_j r_4} \mid at_j \in \text{UAtt}_2^*\right\}\Bigr).\end{aligned} \tag{17}$$

Phase 2: Same as Phase 1.

Guess: Finally, $\mathscr{A}$ will eventually output a guess σ' of σ.

If $\mathscr{A}$ can determine whether d_0 is equal to d_σ or not, $\mathscr{A}$ also can determine whether g^{β_σ} is equal to g^{β_0} or not. The only way for $\mathscr{A}$ to achieve this is to construct a query $\Gamma_2(\beta_0 - \beta_\sigma)$ for some Γ_2. To prove Theorem 3, we will show that $\mathscr{A}$ can never construct a query for $\Gamma_2(\beta_0 - \beta_\sigma)$.

Table 2 shows all the possible queries of G by means of the bilinear map and group elements given to the adversary. Note that β_0, β_σ only incurs in terms $a(r_1 + r_2) + \beta_0$ and $a(r_3 + r_4) + \beta_\sigma$, respectively. Thus, $\mathscr{A}$ must construct $g^{\Gamma_2 a(r_1+r_2-r_3-r_4)}$ for obtaining $g^{\Gamma_2(\beta_0-\beta_\sigma)}$. Moreover, since (r_1, r_2) and (r_3, r_4) are independent, $\mathscr{A}$ must construct $g^{\Gamma_2 a(r_1+r_2)}$ and $g^{\Gamma_2 a(r_3+r_4)}$ for the same Γ_2. Then we show that adversary $\mathscr{A}$ can never build $g^{\Gamma_2 a(r_1+r_2)}$ and $g^{\Gamma_2 a(r_3+r_4)}$ for the same Γ_2.

To construct $\Gamma_2 a(r_1 + r_2)$ for $\mathscr{A}$, as r_1 only appears in the term br_1, we let $\Gamma_2 = \Gamma_2' b$ for some Γ_2'. That is, $\mathscr{A}$ needs to construct the term $\Gamma_2' abr_2$. In order to get that, the only way of constructing $\Gamma_2' abr_2$ is to apply tkn_1 in Table 2 with $\left\{B_j = g^{\alpha_j r_2} \mid at_j \in \text{UAtt}_1^*\right\}$ of cph_1^*, which will result in $e(g,g)^{abtkr_2}$, meaning that $\mathscr{A}$ can construct the query $abtkr_2$. That is, Γ_2 can be written as $\Gamma_2 = \Gamma_2' b = \Gamma_2'' tkb$ for a known constant Γ_2''. Similarly, we can show that Γ_2 can be written as $\Gamma_2 = \Gamma_2' b = \Gamma_2'' t'kb$ for a known constant Γ_2'' to build $\Gamma_2 a(r_3 + r_4)$. Since t and t' are unknown to $\mathscr{A}$, then Γ_2' cannot be constructed, since $\mathscr{A}$ cannot find a known constant Γ_2'' that is the product of t and t'.

In conclusion, $\mathscr{A}$ is able to construct $g^{\Gamma_2 a(r_1+r_2)}$ and $g^{\Gamma_2 a(r_3+r_4)}$ for the same Γ_2 with a negligible probability and get a negligible advantage in the fine-grained authorization game.

Similar to the proof in Theorem 1, an adversary $\mathscr{A}'$ can be simulated and it can be proved that if $\mathscr{A}$ can break the fine-grained authorization security for $|D_0| = |D_1| > 1$, then $\mathscr{A}'$ can break the fined-grained authorization security for $|D_0| = |D_1| = 1$. This completes the proof. □

4.4. Efficiency Analysis.

Now we evaluate the efficiency of the schemes in terms of the asymptotic computational complexity. The asymptotic complexity is measured in terms of operations: H denotes the operation of mapping a bit-string to an element of G, E denotes the group exponentiation operation in G, E_T denotes the group exponentiation operation in G_T, and P denotes the pairing operation. We ignore the multiplication operations because they are much more efficient than the operations mentioned above (Table 3).

We can see that TokenGen incurs small cost when compared with SI. This implies that the data user should use the token to outsource the set intersection operations to the cloud.

Table 3: N is the number of attributes involved in the access tree in question, S is the number of a data user's attributes, and n is the size of sets (here we assume that both sets have the same size n).

Scheme	KP-ABSI
KeyGen	$(3N+2)E+NH$
Enc	$(S+3)nE+(S+1)nH$
TokenGen	$(2N+2)E$
SI	$(2S+2)\cdot 2nP+S\cdot 2nE_T$

5. Conclusions

In this paper, we present a novel cryptographic primitive: delegated key-policy attribute-based set intersection over outsourced encrypted data sets (KP-ABSI). It simultaneously achieves the following: (1) Each data owner outsources his/her data set in encrypted form to a cloud, where the outsourced data set is associated with an attribute set. (2) A data user is associated with an access control policy that is satisfied by the attribute sets of two encrypted data sets (owned by two data owners, respectively) and can delegate to the cloud the set intersection computation over the two data owners' outsourced encrypted data sets. (3) The cloud can conduct the set intersection operation on behalf of the data user without being able to obtain any useful information about the data owners' plaintext data set.

Thus, our scheme can solve the PSI problem in CloudIoT system. Of course, in our solution, the cloud is semihonest. How to build a construction in the malicious model is still an open problem.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Program for the Scientific Research Foundation of Nanjing Institute of Technology (YKJ201980), Program for Natural Science Research Projects of Universities (19KJB520033), and Program for Scientific Research Foundation for Talented Scholars of Jinling Institute of Technology (JIT-B-201726).

References

[1] A. Kobusińska, C. Leung, C. H. Hsu, S. Raghavendra, and V. Chang, "Emerging trends, issues and challenges in internet of things, big data and cloud computing," *Future Generation Computer Systems*, vol. 87, pp. 416–419, 2018.

[2] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A survey on security and privacy issues in internet-of-things," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1250–1258, 2017.

[3] A. Botta, W. De Donato, V. Persico, and A. Pescapé, "Integration of cloud computing and internet of things: a survey," *Future Generation Computer Systems*, vol. 56, pp. 684–700, 2016.

[4] M. J. Freedman, K. Nissim, and B. Pinkas, "Efficient private matching and set intersection," in *Advances in Cryptology-EUROCRYPT 2004*, pp. 1–19, Springer, Berlin, Germany, 2004.

[5] Q. Wang, F. Zhou, J. Xu, and S. Peng, "Tag-based verifiable delegated set intersection over outsourced private datasets," *IEEE Transactions on Cloud Computing*, 2020.

[6] A. Abadi, S. Terzis, R. Metere, and C. Dong, "Efficient delegated private set intersection on outsourced private datasets," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 4, pp. 608–624, 2017.

[7] B. Pinkas, T. Schneider, and M. Zohner, "Scalable private set intersection based on ot extension," *ACM Transactions on Privacy and Security*, vol. 21, no. 2, pp. 1–35, 2018.

[8] M. Ali, J. Mohajeri, M. R. Sadeghi, and X. Liu, "Attribute-based fine-grained access control for outscored private set intersection computation," *Information Sciences*, vol. 536, 2020.

[9] B. Pinkas, M. Rosulek, N. Trieu, and A. Yanai, "PSI from PaXos: fast, malicious private set intersection," in *Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 739–767, Springer, Zagreb, Croatia, October 2020.

[10] O. Ruan and H. Mao, "Efficient private set intersection using point-value polynomial representation," *Security and Communication Networks*, vol. 2020, Article ID 8890677, 12 pages, 2020.

[11] X. Yang, X. Luo, X. A. Wang, and S. Zhang, "Improved outsourced private set intersection protocol based on polynomial interpolation," *Concurrency and Computation: Practice and Experience*, vol. 30, no. 1, p. e4329, 2018.

[12] K. Zhang, J. Chen, H. T. Lee, H. Qian, and H. Wang, "Efficient public key encryption with equality test in the standard model," *Theoretical Computer Science*, vol. 755, pp. 65–80, 2019.

[13] M. Zeng, J. Chen, K. Zhang, and H. Qian, "Public key encryption with equality test via hash proof system," *Theoretical Computer Science*, vol. 795, pp. 20–35, 2019.

[14] H. T. Lee, S. Ling, J. H. Seo, and H. Wang, "Public key encryption with equality test from generic assumptions in the random oracle model," *Information Sciences*, vol. 500, pp. 15–33, 2019.

[15] D. H. Duong, K. Fukushima, S. Kiyomoto, P. S. Roy, and W. Susilo, "A lattice-based public key encryption with equality test in standard model," in *Proceedings of the Australasian Conference on Information Security and Privacy*, pp. 138–155, Springer, Christchurch, New Zealand, July 2019.

[16] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology–EUROCRYPT 2005*, pp. 457–473, Springer, Berlin, Germany, 2005.

[17] J. Li, Q. Yu, Y. Zhang, and J. Shen, "Key-policy attribute-based encryption against continual auxiliary input leakage," *Information Sciences*, vol. 470, pp. 175–188, 2019.

[18] Y. Liu, L. Wang, X. Shen, L. Li, and D. An, "Space-efficient key-policy attribute-based encryption from lattices and two-dimensional attributes," *Security and Communication Networks*, vol. 2020, Article ID 2345369, 11 pages, 2020.

[19] J. Zhang and H. Gao, "A compact construction for non-monotonic key-policy attribute-based encryption," *International Journal of High Performance Computing and Networking*, vol. 13, no. 3, pp. 321–330, 2019.

[20] Z. Liu, S. Duan, P. Zhou, and B. Wang, "Traceable-then-revocable ciphertext-policy attribute-based encryption

scheme," *Future Generation Computer Systems*, vol. 93, pp. 903–913, 2019.

[21] Q. M. Malluhi, A. Shikfa, V. D. Tran, and V. C. Trinh, "Decentralized ciphertext-policy attribute-based encryption schemes for lightweight devices," *Computer Communications*, vol. 145, pp. 113–125, 2019.

[22] H. Ma, Z. Wang, and Z. Guan, "Efficient ciphertext-policy attribute-based online/offline encryption with user revocation," *Security and Communication Networks*, vol. 2019, Article ID 8093578, 11 pages, 2019.

[23] H. Zhu, L. Wang, H. Ahmad, and X. Niu, "Key-policy attribute-based encryption with equality test in cloud computing," *IEEE Access*, vol. 5, pp. 20428–20439, 2017.

[24] Q. Wang, L. Peng, H. Xiong, J. Sun, and Z. Qin, "Ciphertext-policy attribute-based encryption with delegated equality test in cloud computing," *IEEE Access*, vol. 6, pp. 760–771, 2017.

[25] Y. Cui, Q. Huang, J. Huang, H. Li, and G. Yang, "Ciphertext-policy attribute-based encrypted data equality test and classification," *The Computer Journal*, vol. 62, no. 8, pp. 1166–1177, 2019.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 5556634, 8 pages
https://doi.org/10.1155/2021/5556634

Research Article

Improved CNN-Based Hashing for Encrypted Image Retrieval

Wenyan Pan,[1] **Meimin Wang**,[1] **Jiaohua Qin**,[2] **and Zhili Zhou**[1]

[1] *Jiangsu Engineering Center of Network Monitoring & School of Computer and Software,*
Nanjing University of Information Science and Technology, Nanjing 210044, China
[2] *College of Computer Science and Information Technology, Central South University of Forestry & Technology,*
Changsha 410004, China

Correspondence should be addressed to Jiaohua Qin; qinjiaohua@163.com and Zhili Zhou; zhou_zhili@163.com

Received 10 January 2021; Revised 5 February 2021; Accepted 17 February 2021; Published 26 February 2021

Academic Editor: Yuan Tian

As more and more image data are stored in the encrypted form in the cloud computing environment, it has become an urgent problem that how to efficiently retrieve images on the encryption domain. Recently, Convolutional Neural Network (CNN) features have achieved promising performance in the field of image retrieval, but the high dimension of CNN features will cause low retrieval efficiency. Also, it is not suitable to directly apply them for image retrieval on the encryption domain. To solve the above issues, this paper proposes an improved CNN-based hashing method for encrypted image retrieval. First, the image size is increased and inputted into the CNN to improve the representation ability. Then, a lightweight module is introduced to replace a part of modules in the CNN to reduce the parameters and computational cost. Finally, a hash layer is added to generate a compact binary hash code. In the retrieval process, the hash code is used for encrypted image retrieval, which greatly improves the retrieval efficiency. The experimental results show that the scheme allows an effective and efficient retrieval of encrypted images.

1. Introduction

With the development of cloud computing, more and more companies and individuals store image data on the cloud server. Therefore, how to efficiently retrieve images in the cloud becomes an urgent problem. Cloud computing [1] is an emerging new computing paradigm with efficient image storage, which makes it an attractive choice for image retrieval. Despite the benefits, image information privacy becomes the main concern with image retrieval in cloud computing.

In order to protect the image information, it is necessary to encrypt the image before it is submitted to the cloud. The widely used encryption methods include chaotic image encryption [2] and Arnold transform [3]. However, it is not suitable to directly apply image retrieval technology in the plaintext domain for image retrieval on the encryption domain. Therefore, how to protect image information in the cloud computing while quickly retrieving the images that users need is an urgent problem that needs to be solved in the field of encrypted image retrieval.

In the field of image retrieval, most previous approaches exploit the frequency domain feature [4, 5], SIFT [6]. However, these approaches are based on hand-crafted features which cannot represent the image content comprehensively because of the low retrieval accuracy.

With the development of deep learning, the CNNs [7–11] have shown significant improvements in the performance on various tasks. However, the most CNNs usually have hundreds of layers, thus making networks more inefficient. Most state-of-the-art lightweight architectures, such as MobileNet [12] and ShuffleNet [13], become more efficient because of their network architectures. These networks can be carried out in a timely fashion on a computationally limited platform.

Even though the CNN-based representation is an appealing solution for image retrieval in the plaintext domain, it is inefficient to directly compute the similarity between two CNN features, such as 4096-dimensional vectors of the full connection layer in AlexNet. Recently, some approaches have been using deep architectures for hash learning for image retrieval [14, 15]. However, most of them are used for the plaintext domain, but lacks research on the encryption domain.

In order to address the above issues, this paper proposes an improved CNN-based hashing method for encrypted image retrieval (DLHEIR). In our method, we increase the size of the input image of the CNN to obtain better features and replace a part of the structure of the DenseNet network with inverted residual block to reduce the computational cost and parameters. The improved CNNs are used to generate hash codes for encrypted image retrieval.

Our main contributions are as follows:

(1) This paper proposes an improved CNN-based hashing method for encrypted image retrieval (DLHEIR). This network can learn image representations to generate the binary hash code for rapid image retrieval.

(2) We used images with larger sizes as input to the CNN to obtain better features. Moreover, the inverted residual block is introduced into our method, which can reduce the computational cost and parameters.

The organization of the remaining part is given as follows. Section 2 discusses the related works. Section 3 introduces the proposed method. Section 4 shows our experimental results, and we conclude this paper in Section 5.

2. Related Work

Content-based image retrieval (CBIR) refers to the retrieving of the needed information in large-scale multimedia data according to the content of the image. Recently, image retrieval has been applied in many fields, such as image search [16, 17] and image steganography [18]. However, it cannot be applied in cloud computing due to the privacy of images.

The searchable encryption (SE) method enables the users to store encrypted data in the cloud computing and supports data search in the encrypted domain. Xia et al. [19] proposed an encrypted image retrieval scheme (PSSE) in the cloud environment, which uses MPEG-7 visual descriptors as image features. The KNN is used to protect features, and the local sensitive hashing is used to improve retrieval efficiency. Qin et al. [20] proposed an encrypted image retrieval approach in the cloud computing environment, which employs the improved Harris algorithm and Local Sensitive Hash (LSH) to retrieve encrypted images. Shen et al. [21] proposed a secure content-based image retrieval method, which uses a secure multiparty computation technique to encrypt image features. Cheng et al. [4] proposed an encrypted JPEG image retrieval scheme based on the Markov process, which uses encryption to encrypt DCT coefficients to protect the confidentiality of the JPEG image content. Xia et al. [22] proposed an outsourcing CBIR scheme based on the BOEW model. Ferreira et al. [23] proposed a secure framework for outsourcing privacy-protected storage and retrieving in a large shared image repository. Lu et al. [24] proposed a privacy protection image retrieval method based on an encrypted image collection which uses a set of visual words to represent images, and the Jaccard distance is used to measure the similarity between images. Xia et al. [25] proposed a privacy-preserving image retrieval method based on Scale Invariant Feature Transform (SIFT) features and Earth Mover's Distance (EMD). Weng et al. [26] proposed a privacy preserving framework for an application called outsourcing media search. The framework relies on multimedia hashing and symmetric encryption to protect image information. However, these approaches are based on hand-crafted features, which do not consider the global information of the image, resulting in low accuracy for encrypted image retrieval.

CNNs have recently provided an attractive solution for many version tasks. The previous approaches are attributed to the ability of CNN to learn the rich image representations, which can be applied to the field of image retrieval [27, 28]. However, due to the high-computational cost of computing the similarity between two CNN features, some approaches use CNNs to automatically learn binary hashing codes [29–31]. However, these approaches are applicable only in the plaintext domain, and there are few approaches that focus on CNN-based encryption image retrieval.

In this paper, CNNs are applied to the field of encrypted image retrieval. With the powerful representation ability of CNNs' features, the accuracy of encrypted image retrieval is improved. At the same time, the retrieval efficiency is greatly improved by using the hash code.

3. Proposed Method

3.1. System Model. The system model is shown in Figure 1, and the system model mainly consists of three parts: data owner, cloud server, and query user.

Data owner has the image dataset $M = \{m_1, m_2, \ldots, m_n\}$. To preserve the image content, the dataset needs to be encrypted, generating the encrypted dataset $E = \{e_1, e_2, \ldots, e_n\}$. where n is the number of images in the dataset. To achieve rapid image retrieval, the data owner needs to generate the hash code corresponding to the image dataset. Both the encrypted image and hash code are outsourced to the cloud server. The data owner also needs to send the key to the query user when receiving the retrieval request.

Cloud server stores the encrypted dataset and hash code from the data owner. When receiving the retrieval request from the query user, the cloud server needs to calculate the similarity between the hash code from the data owner and the trapdoor of the query image and returns the top k retrieval results to the query user.

Query user generates the trapdoors for the query images and uploads it to the cloud server. We define the trapdoor as the hash code for query images, which utilize the same method as the data owner does. After receiving the resulting images, the query user sends a request to the data owner and obtains the key, and the user can decrypt the encrypted image with the key.

3.2. Overview of the Proposed Method. The proposed method mainly includes six functions, which are executed by the data owner, cloud server, and query user.

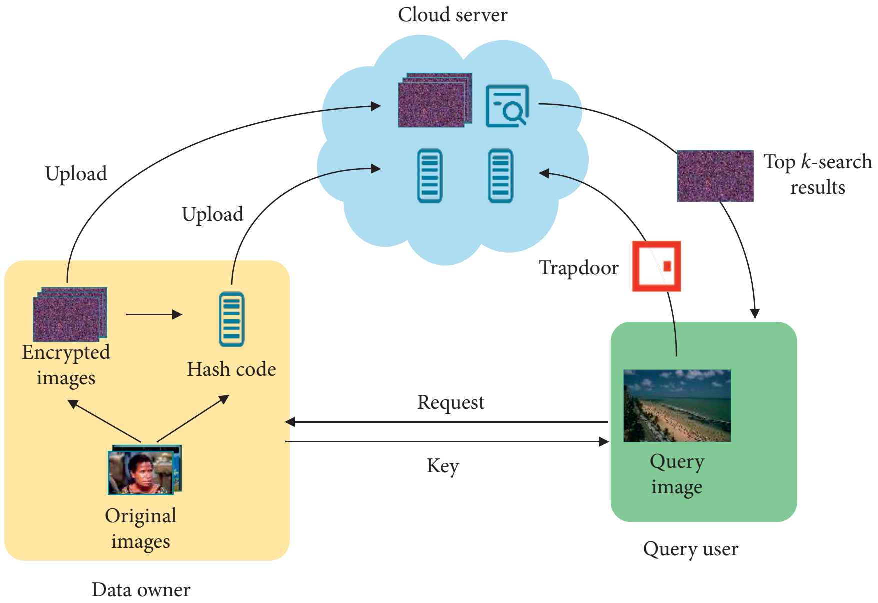


Figure 1: System model.

The following functions are executed in the data owner:

(1) *Key Generation*. KeyGen $(1^k) \longrightarrow K$. The input of the function is parameter k, and it returns the key K. After the user authorization, the data owner sends the key K to the user for decrypting the encrypted image.

(2) *Image Encryption*. EncImg $(K, M) \longrightarrow E$. The inputs of the function are the key K and the image dataset M, and it returns the encrypted image dataset E.

(3) *Hash Code Generation*. hashgen $(M, \text{bits}) \longrightarrow \text{HC}$. By adopting our method, the input of the function is the image dataset M, and this function returns the hash code HC.

The following functions are executed in the query user:

(1) *Trapdoor Generation*. TrapGen $(Q) \longrightarrow \text{HC}_q$. The input of this function is the query image Q. Construct trapdoor and generate hash code HC_q of query image.

(2) *Image Decryption*. Dec $(K, R) \longrightarrow \text{Img}_R$. The inputs of this function are the key K and the similar encrypted image R returned by the cloud server, and it decrypts the similar encrypted image to return a similar image Img_R.

The following function is executed in the cloud server:

(1) *Search*. Search $(\text{HC}, \text{HC}_q) \longrightarrow R$. The function calculates the similarity between HC_q corresponding to the query image and the *HC* corresponding to the encrypted image dataset and it returns similar encrypted image set R.

3.3. Improved Convolutional Neural Network Hashing. In this section, we will introduce our method, which consists of two main components, image preprocessing and network architecture.

3.3.1. Image Preprocessing. Before training or testing the network, the input images should be resized to the same size. For example, when training and testing DenseNet, all images should resize to 224×224 before feeding into the network.

The large image is resized to 224×224 or 299×299 by cropping or warping. The cropping may lose important information of the image, while the warping may change the aspect ratio of the image, and this will affect the features extracted by the CNN.

Consequently, in this paper, we increase the input image size of CNNs. Specifically, for the Corel10K dataset, we calculate the maximum image height and width, and then, the largest value height and width are taken as the image size. For example, for the Corel10K dataset, the maximum image height and width in the Corel10K dataset is 384 and 256, so the size of the input image is resized to 384×384.

3.3.2. Network Architecture. Inverted Residual Block. The network architecture of our method is shown in Figure 2. Specifically, the image is resized to 384×384 as the input of the DenseNet201. Then, the inverted residual block is introduced to replace a part of the architecture in the DenseNet, which can greatly reduce computational cost and parameters.

The inverted residual block consists of depthwise separable convolution. The computational cost c_d of depthwise separable convolution is shown in the following equation:

$$c_d = k \times k \times C_{\text{in}} \times W_{\text{in}} \times H_{\text{in}} + C_{\text{in}} \times W_{\text{out}} \times H_{\text{out}} \times C_{\text{out}}. \quad (1)$$

The parameter p_d of depthwise separable convolution is computed in the following equation:

$$p_d = k \times k \times C_{\text{in}} + 1 \times 1 \times C_{\text{in}} \times C_{\text{out}}. \quad (2)$$

For standard convolutions, the computational cost c_s and parameter p_s are computed by the following equation:

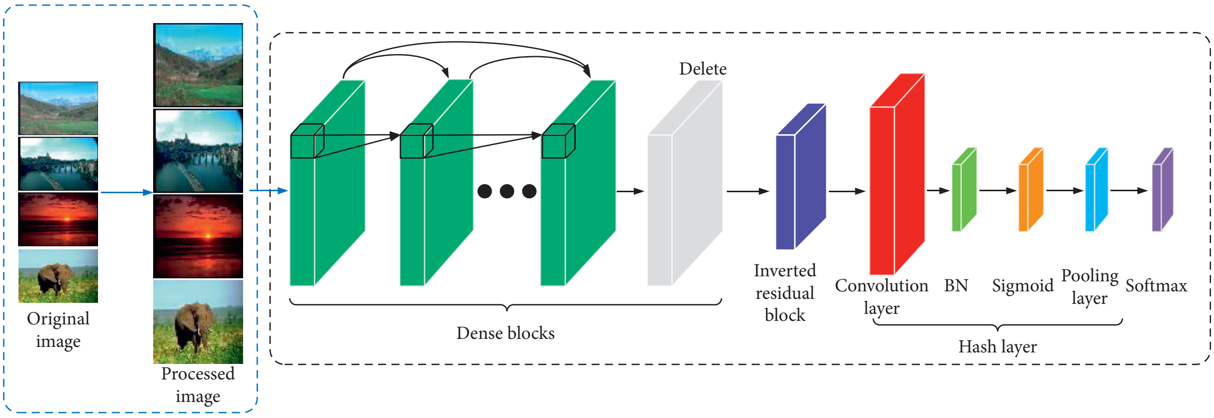


Figure 2: Network architecture.

$$c_s = k \times k \times C_{\text{in}} \times C_{\text{out}} \times H_{\text{out}} \times W_{\text{out}}, \quad (3)$$

$$p_s = k \times k \times C_{\text{in}} \times C_{\text{out}}. \quad (4)$$

Suppose the input feature map of depthwise separable convolution has the size $W_{in} \times H_{in} \times C_{in}$ and the output feature map has the size $W_{\text{out}} \times H_{\text{out}} \times C_{\text{out}}$, where C_{in} and C_{out} are the channel of the feature map, W_{in} and W_{out} are the width of the feature map, H_{in} and H_{out} are the height of the feature map, respectively, and $k \times k$ denotes the kernel size. The computational cost ratio r_c of the depthwise separable convolution to standard convolution is shown in the following equation:

$$r_c = \frac{C_{\text{dsc}}}{C_{\text{std}}} = \frac{1}{C_{\text{out}}} + \frac{1}{K^2}. \quad (5)$$

The parameters' cost ratio r_p of the depthwise separable convolution to standard convolution is shown in the following equation:

$$r_p = \frac{P_{\text{dsc}}}{P_{\text{std}}} = \frac{1}{C_{\text{out}}} + \frac{1}{K^2}. \quad (6)$$

Equations (4) and (5) show that the depth separable convolution uses less computational cost and parameters than standard convolution.

Densenet201 consists of four dense blocks, which consists of 6, 12, 48, and 32 BN-ReLU-Conv (1 × 1)-BN-ReLU-Conv (3 × 3) structures, respectively, where BN indicates batch normalization, ReLU indicates linear rectifier function, and Conv (1 × 1) indicates a Conv2D layer with N filters of kernel size 1-by-1. In order to reduce computational cost and parameters of the network, the last 14 BN-ReLU-Conv (1 × 1)-BN-ReLU-Conv (3 × 3) were replaced by an inverted residual block. Then, a hash layer is added, which consists of a convolution layer, batch normalization, sigmoid activation, and pooling layer. Finally, SoftMax is added to form our network.

Hash Layer. In this section, we will systematically describe the hash layer. It consists of three main layers, which are a convolutional layer, a batch normalization layers, activation function, and a global average pooling layer. The convolutional layer is a Conv2D layer with N filters of kernel size 1-by-1. For the activation function, we choose sigmoid so that parameters are approximated to (0, 1).

Suppose the input feature map has size of $l \times r \times q$, where l, r, and q are height, width, and channel of the feature map, respectively. The output of the feature map hash layer has size of $1 \times 1 \times N$, and the feature $AP = \{ap_1, ap_2, \ldots, ap_q\}$ is obtained.

In feature extraction, firstly, all images are resized to 384 × 384 before being fed into the network, and the feature AP of the global average pooling layer is extracted, and the binary codes are obtained by using the hash function to binarize AP by a threshold. The hash function is shown in the following equation:

$$\text{HC} = \begin{cases} 1, & \text{if } ap_i \geq \text{th}, \\ 0, & \text{if } ap_i < \text{th}, \end{cases} \quad i = 1, 2, \ldots, q, \quad (7)$$

where ap_i is the parameter in AP and th is the threshold of the hash function.

4. Experimental Results and Analysis

The experiments were performed on the Corel10K dataset [32]. Corel10K is a benchmark dataset for image retrieval. It includes 100 categories, and each category contains 100 similar images.

The experiment code was written in Python and Matlab R2016a under the Windows 10 system, using Intel(R) Core (TM) i7-9700KF CPU @ 3.60GHz, 16.00 GB RAM, and a Nvidia GeForce GTX 2080Ti GPU.

In the experiment, 80 images were randomly selected from each category of the Corel10K dataset as the training set, and the remaining images were used as the test set. DenseNet201 was selected as the backbone network. In the fine-tuning, we use the pretrain model which is trained on the ImageNet dataset. The stochastic gradient descent (SGD) is used as the network optimizer, the learning rate is set to 0.01, the momentum is set to 0.9, the batch size is set to 64, and epochs are set to 200.

4.1. Retrieval Precision

4.1. Retrieval Precision. In our experiments, "precision" was used as the evaluation metric, which is defined as $P_k = k'/k$, where k' is the number of real similar images in the k retrieved images. In the experiment, we use the test set as the query image and the training set as the query image collection to test the retrieval precision. We compare our method with the other methods [6, 17]. The experimental results are reported in Figure 3.

As shown in Figure 3, it is clear that our method outperforms conventional methods [6, 17]. This is because these methods all utilize the hand-craft feature, which limits their performance. In particular, the performance gap is not obvious as k increases, except $k = 100$. Also, note that our method with 48 bits has better performance than the model with others.

We also evaluate the role of image size for retrieval precision. The experimental results are shown in Table 1.

It is clear from Table 1 that the increase in the image scale consistently improves retrieval precision in different hash bits. This is because using larger images is beneficial for performance improvement. A scale larger than our method would instead increase the memory consumption of GPU and computational cost and parameters.

4.2. Comparison of Model Parameters and MFLOPs. In this section, we compare parameters and MFLOPS of our method with the original CNN combined with the hash layer. The experimental results are reported in Table 2.

Floating point operations per second (FLOPs) is a measure of computer performance, which is widely used to measure the computation cost in CNN models, such as ShuffleNet [13]. As can be seen from Table 2, we can find that our method has less parameters and MFLOPs.

4.3. Efficiency. The time consumptions of the retrieval, feature extraction, index construction, and trapdoor generation are compared in this section.

Time Consumption of Retrieval. In order to utilize the powerful computing power of the cloud server, the retrieval is applied in the cloud server, and the most similar k images are returned by calculating the Euclidean distance between two hash codes. Table 3 presents the time consumption of retrieval when retrieving images $k = 20$.

It can be seen from Table 3, the retrieval time consumption increases as the retrieval collection increases. It is clear that our method achieves better efficiency [6, 17]. This is because our method utilized the low-dimension binary hash code, which achieved efficiency in image retrieval.

Time Consumption of Feature Extraction. We also compared the time consumption of feature extraction with the CSD and SCD descriptors in the MPEG-7 feature extraction method of [17], and the time consumption of SIFT feature extraction in [6]. The experimental results are shown in Figure 4.

Figure 4 shows the feature extraction times for different numbers of image collection. Compared with [6, 17], the time consumption of our method is shorter on different numbers of image collections in most cases. This is because the time consumption of feature extraction in our method mainly consists of two parts: time consumption of the load model and hashing. Compared with complex conventional methods, our method is more efficient.

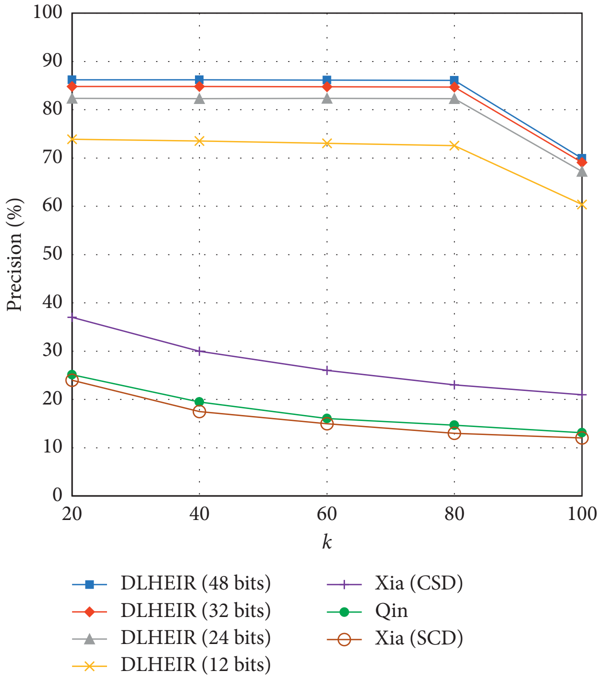


Figure 3: Top-k retrieval precision.

Table 1: Retrieval precision of different input image sizes (%).

Method		k				
Image size	Bits	20	40	60	80	100
384 × 384	12 bits	73.92	73.56	73.04	72.55	60.38
	24 bits	82.33	82.31	82.36	82.29	67.21
	32 bits	84.82	84.80	84.75	84.69	69.06
	48 bits	86.23	86.21	86.15	86.08	69.91
224 × 224	12 bits	72.55	72.29	72.13	71.85	59.63
	24 bits	80.10	80.02	80.01	79.90	65.23
	32 bits	81.50	81.42	81.38	81.34	66.37
	48 bits	82.26	82.24	82.19	82.11	66.01

Table 2: Compare the number of parameters and MFLOPs for different methods.

Method	Parameters	MFLOPs
DLHEIR (48 bits)	176.64	158.25
DLHEIR (32 bits)	176.54	158.25
DLHEIR (24 bits)	176.49	158.21
DLHEIR (12 bits)	176.41	158.13
Original CNN + hash layer (48 bits)	184.19	164.97
Original CNN + hash layer (32 bits)	183.87	164.68
Original CNN + hash layer (24 bits)	183.71	164.54
Original CNN + hash layer (12 bits)	183.46	164.32

Time Consumption of Index Construction. In our method, the similarity is directly computed by two hash codes without index construction, so there is no time consumption of index construction in our method. The time consumption of index construction comparison with Xia and Qin is shown in Figure 5.

Table 3: Time consumption of retrieval.

Methods	Number of images in data collection				
	2000	4000	6000	8000	10000
DLHEIR (48 bits)	0.58	0.742	0.908	1.073	1.249
DLHEIR (32 bits)	0.537	0.657	0.776	0.894	1.016
DLHEIR (24 bits)	0.505	0.608	0.637	0.799	0.908
DLHEIR (12 bits)	0.483	0.554	0.637	0.703	0.785
Xia (CSD)	2.51	4.60	6.26	8.34	10.49
Xia (SCD)	1.65	3.79	5.36	6.53	8.78
Qin	0.95	1.46	2.25	2.88	3.63

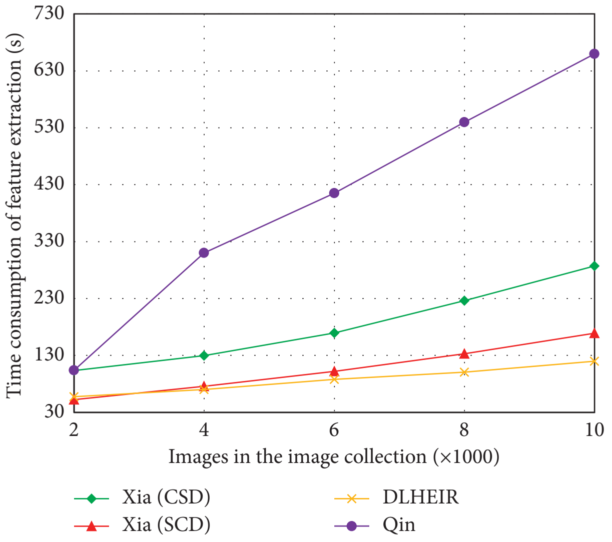


Figure 4: Time consumption of feature extraction (s).

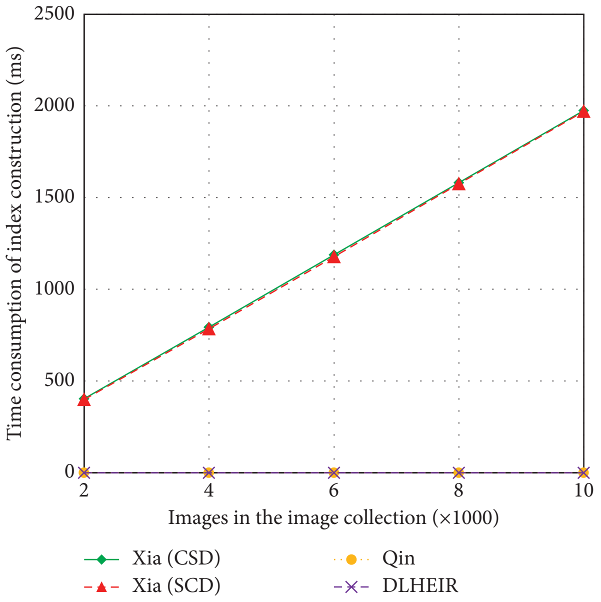


Figure 5: Time consumption of index construction (ms).

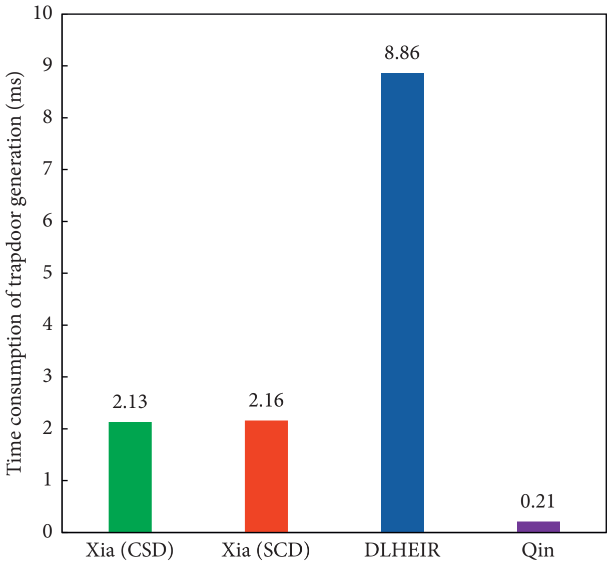


Figure 6: Time consumption of trapdoor generation (ms).

Time Consumption of Trapdoor Generation Time. Similar to the feature extraction, the trapdoor generation incurs the hash code generated by the data owner, so the time consumption of the trapdoor construction is the hash code generation time of the query image. The experimental results are shown in Figure 6**.**

We test the time consumption of trapdoor generation compared with the [6, 17] in Figure 6. Our method has more time consumption to these methods. This is because, in feature extraction, we need to extract features from the deep layer of DenseNet, so the time consumption is more than [6, 17].

4.4. Security Analysis

(i) *The Privacy of Image Content*. In our method, the images stored on the cloud server are encrypted with an encryption method. The key is generated by the data owner. Thus, the privacy of the image content in our scheme is well protected.

(ii) *The Privacy of Hash Code***.** The hash code may reveal the information about the image content. In our method, the hash code mapped from the feature vectors are protected by a one-way hash function. Thus, the hash code is well protected.

5. Conclusion

This paper proposes an improved CNN-based hashing method for encrypted image retrieval. In our method, we increase the size of the input image of the CNN to obtain better features and replace part of the structure of the DenseNet network with inverted residual block to reduce the computational cost and parameters, and a hash layer is added for hash code generation. These hash codes are used for encrypted image retrieval. The experimental results show that the method achieves better performance and greatly improves the retrieval efficiency. In the future, we plan to design more efficient methods to reduce the burden on users.

Data Availability

The Corel10K data used to support the findings of this study have been deposited in the "http://www-db.stanford.edu/~wangz/image.vary.jpg.tar."

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant nos. 61972205, U1836208, and U1836110, National Key R&D Program of China under Grant 2018YFB1003205, Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions Fund, Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET) Fund, China, and Ministry of Science and Technology (MOST), Taiwan, under Grant nos. 108-2221-E-259-009-MY2 and 109-2221-E-259-010.

References

[1] L. Wang, G. Von Laszewski, A. Younge et al., "Cloud computing: a perspective study," *New Generation Computing*, vol. 28, no. 2, pp. 137–146, 2010.

[2] X.-Y. Wang, L. Yang, R. Liu, and A. Kadir, "A chaotic image encryption algorithm based on perceptron model," *Nonlinear Dynamics*, vol. 62, no. 3, pp. 615–621, 2010.

[3] W. Chen, C. Quan, and C. J. Tay, "Optical color image encryption based on Arnold transform and interference method," *Optics Communications*, vol. 282, no. 18, pp. 3680–3685, 2009.

[4] H. Cheng, X. Zhang, J. Yu, and F. Li, "Markov process-based retrieval for encrypted JPEG images," *EURASIP Journal on Information Security*, vol. 2016, no. 1, 1 page, 2016.

[5] R. Bellafqira, G. Coatrieux, D. Bouslimi, G. Quellec, and M. Cozic, "Secured outsourced content based image retrieval based on encrypted signatures extracted from homomorphically encrypted images," arXiv preprint arXiv:1704.00457, 2017.

[6] J. Qin, Y. Cao, X. Xiang, Y. Tan, L. Xiang, and J. Zhang, "An encrypted image retrieval method based on SimHash in cloud computing," *Computers, Materials & Continua*, vol. 62, no. 3, pp. 389–399, 2020.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[9] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, June 2015.

[10] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. B. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[12] A. G. Howard, M. Zhu, B. Chen et al., "Mobilenets: efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

[13] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: an extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6848–6856, Salt Lake City, Utah, June 2018.

[14] K. Lin, H. F. Yang, J. H. Hsiao, and C.-S. Chen, "Deep learning of binary hash codes for fast image retrieval," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 27–35, Boston, MA, USA, June 2015.

[15] H. Liu, R. Wang, S. Shan, and X. Chen, "Deep supervised hashing for fast image retrieval," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2064–2072, Las Vegas, NV, USA, June 2016.

[16] Z. Zhou, Q. M. J. Wu, Y. Yang, and X. Sun, "Region-level visual consistency verification for large-scale partial-duplicate image search," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 2, pp. 1–25, 2020.

[17] A. Gordo, J. Almazán, J. Revaud, and D. Larlus, "Deep image retrieval: learning global representations for image search," in *Proceedings of the European Conference on Computer Vision*, pp. 241–257, Amsterdam, The Netherlands, October 2016.

[18] Z. Zhou, Y. Mu, and Q. M. J. Wu, "Coverless image steganography using partial-duplicate image retrieval," *Soft Computing*, vol. 23, no. 13, pp. 4927–4938, 2019.

[19] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2594–2608, 2016.

[20] J. Qin, H. Li, X. Xiang et al., "An encrypted image retrieval method based on Harris corner optimization and LSH in cloud computing," *IEEE Access*, vol. 7, pp. 24626–24633, 2019.

[21] M. Shen, G. Cheng, L. Zhu, X. Du, and J. Hu, "Content-based multi-source encrypted image retrieval in clouds with privacy preservation," *Future Generation Computer Systems*, vol. 109, pp. 621–632, 2020.

[22] Z. Xia, L. Jiang, D. Liu, L. Lu, and B. Jeon, "BOEW: a content-based image retrieval scheme using bag-of-encrypted-words in cloud computing," *IEEE Transactions on Services Computing*, no. 1, p. 1, 2019.

[23] B. Ferreira, J. Rodrigues, J. Leitao, and H. Domingos, "Practical privacy-preserving content-based retrieval in cloud image repositories," *IEEE Transactions on Cloud Computing*, vol. 7, no. 3, pp. 784–798, 2019.

[24] W. Lu, A. Swaminathan, A. L. Varna, and M. Wu, "Enabling search over encrypted multimedia databases. Media forensics and security," *International Society for Optics and Photonics*, vol. 7254, p. 725418, 2009.

[25] Z. Xia, Y. Zhu, X. Sun, Q. Zhan, and K. Ren, "Towards privacy-preserving content-based image retrieval in cloud computing," *IEEE Transactions on Cloud Computing*, vol. 6, no. 1, pp. 276–286, 2018.

[26] L. Weng, L. Amsaleg, and T. Furon, "Privacy-preserving outsourced media search," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 10, pp. 2738–2751, 2016.

[27] A. S. Razavian, J. Sullivan, S. Carlsson, and A. Maki, "[Paper] visual instance retrieval with deep convolutional networks," *ITE Transactions on Media Technology and Applications*, vol. 4, no. 3, pp. 251–258, 2016.

[28] M. Tzelepi and A. Tefas, "Relevance feedback in deep convolutional neural networks for content based image retrieval," in *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, pp. 1–7, Thessaloniki Greece, May 2016.

[29] V. A. Nguyen and M. N. Do, "Deep learning based supervised hashing for efficient image retrieval," in *Proccedings of the 2016 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, Seattle, WA, USA, July 2016.

[30] R. Zhang, L. Lin, R. Zhang, W. Zuo, and L. Zhang, "Bit-scalable deep hashing with regularized similarity learning for image retrieval and person re-identification," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4766–4779, 2015.

[31] X. Li, Q. Xue, and M. C. Chuah, "CASHEIRS: cloud assisted scalable hierarchical encrypted based image retrieval system," in *Proccedings of the IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9, Atlanta, GA, USA, May 2017.

[32] J. Z. Wang, *Semantics-sensitive Integrated Matching for Picture Libraries and Biomedical Image Databases*, Stanforduniversity, Stanford, CA, USA, 2000.

Hindawi
Security and Communication Networks
Volume 2021, Article ID 6635027, 8 pages
https://doi.org/10.1155/2021/6635027

Research Article

Recognition of Disease Genetic Information from Unstructured Text Data Based on BiLSTM-CRF for Molecular Mechanisms

Lejun Gong,[1,2] Xingxing Zhang,[1] Tianyin Chen,[1] and Li Zhang[3]

[1] *Jiangsu Key Lab of Big Data Security & Intelligent Processing, School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*
[2] *Zhejiang Engineering Research Center of Intelligent Medicine, Wenzhou 325035, China*
[3] *College of Computer Science and Technology, Nanjing Forestry University, Nanjing 210037, China*

Correspondence should be addressed to Lejun Gong; glj98226@163.com

Received 24 December 2020; Revised 20 January 2021; Accepted 5 February 2021; Published 19 February 2021

Academic Editor: Yuan Tian

Disease relevant entities are an important task in mining unstructured text data from the biomedical literature for achieving biomedical knowledge. Autism spectrum disorder (ASD) is a disease related to a neurological and developmental disorder characterized by deficits in communication and social interaction and by repetitive behaviour. However, this kind of disease remains unclear to date. In this study, it identifies entities associated with disease using the machine learning of a computational way from text data collection for molecular mechanisms related to ASD. Entities related to disease are extracted from the biomedical literature related to autism by using deep learning with bidirectional long short-term memory (BiLSTM) and conditional random field (CRF) model. Compared other previous works, the approach is promising for identifying entities related to disease. The proposed approach including five types of molecular entities is evaluated by GENIA corpus to obtain an F-score of 76.81%. The work has extracted 9146 proteins, 145 RNAs, 7680 DNAs, 1058 cell-types, and 981 cell-lines from the autism biomedical literature after removing repeated molecular entities. Finally, we perform GO and KEGG analyses of the test dataset. This study could serve as a reference for further studies on the etiology of disease on the basis of molecular mechanisms and provide a way to explore disease genetic information.

1. Introduction

With the rapid development of intelligent computing and machine learning technology, especially the development of deep learning technology [1], artificial intelligence technology has developed more widely involving algorithms and applications [2–6]. Moreover, it has been widely used in academia and industry such as communication security [7, 8] and opinion and text mining [9–11]. It is also popular in the biomedical field [12, 13]. Abundant experimental data in biomedical research are available [14]. A large number of terminological resources and knowledge bases can also be used in machine learning methods for biomedical text mining [15]. Hassanpour et al. [16] provided a semantic-based method for extracting concept definitions for scientific publications on autism phenotype. Thabtah et al. [17] proposed a new computational intelligence approach based on variable analysis to detect features for autism screening. Spencer et al. [18] found gene associations using frequent pattern mining specific to autism. Bush et al. [19] extracted ASD data from electronic health records for different workflows. Macedoni-Lukšič et al. [20] used ontology construction to identify the main concepts in autism by using the RaJoLink method based on Swanson's ABC model. In our previous work, we extracted candidate genes related to autism based on associated rules [21].

Autism is a neurodevelopmental disorder called autism spectrum disorder (ASD). ASD is a neurological and developmental disorder characterized by deficits in communication, social interaction, and repetitive behaviour. It is also syndrome about neurodevelopment with an as yet unknown unifying pathological or neurobiological etiology.

Zhang et al. [22] conducted genome-wide association study and integrated brain region-related enhancer-gene networks for ASD to explore the roles of chromosomal enhancer region in this disorder. Parr et al. [23] employed Bayesian frameworks to understand brain function formulate perception and action as inferential processes. Sato et al. [24] combined fuzzy spectral clustering and entropy analysis of functional MRI data to identify segregated regions in the functional brain connectome of individuals with autism. They also proved efficiency of this new tool to characterize neuropsychiatric disorders [25]. Rosenberg et al. [26] proposed that the alterations in nonlinear, canonical computations underlie the behavioural characteristics of individuals with autism. They believe that computational perspective on autism may aid in identifying physiological pathways to target in ASD treatment. The abovementioned computational approach can be employed to explore the etiology of autism without the need for expensive and time-consuming experimental validation. Although the etiology of ADS remains unclear, some studies have demonstrated that strong genetic components are involved in ASD development [27–29]. In the present study, we explored the molecular mechanisms related to ASD through computations to understand the etiology of this disorder.

To explore the underlying disease's mechanisms, we identified five disease entities related to autism based on deep learning using the hybrid model containing both bidirectional long short-term memory (BiLSTM) [30] and conditional random field (CRF) [31] model, and explored the molecular mechanism by analysing their relationships among molecular entities.

2. Materials and Methods

As a large unstructured data repository, the biomedical literature contains abundant biomedical information from which useful knowledge (specific and relevant interest points) can be obtained by subjecting unstructured text to natural language processing. In this study, molecular information related to autism was obtained from the biomedical literature. We first extracted molecular entities from experimental corpus by using a suitable computational model and then explored their relationships among molecular entities. Then, we divided these entities related to autism into confirmed and unknown samples. Finally, we explored known samples related to autism to understand the ethology of the disorder, which could offer a reference for understanding the unknown molecular mechanisms of the other samples related to autism.

Identifying molecular entities is a key factor in this study. Machine learning is the mainstream method. The task is considered a sequence tagging NLP problem. The output of taggers could be used for downstream input in sequence tagging. Some linear statistical models that have been applied in sequence tagging include the Hidden Markov model [32], maximum entropy Markov models [33], and CRF models. Recently, neural networks have been proposed to tackle the sequence tagging problem [34–36]. This study combined a hybrid network both BiLSTM and CRF to form a BiLSTM-CRF model for identifying molecular entities. The network could efficiently use past input features via a BiLSTM layer and sentence level tag information via a CRF layer. The following sections would describe the model of identifications.

2.1. LSTM Model. Long short-term memory (LSTM) [37] networks are similar to recurrent neural networks (RNNs). RNNs could not learn the relevant information of input data with sigma cells or tanh cells. The hidden layer updates are replaced by purpose-built memory cells in LSTM. Thus, LSTM is a special recurrent neural network model which could selectively store contextual information using a specially designed gate structure containing input gate, output gate, and forget gate. LSTM could handle the long-term dependencies well. The LSTM memory cell is illustrated in the works [30, 38]. By forgetting the information in the cell state and memorizing new information, this allows information that is useful for subsequent moments of computation to be transmitted, while useless information is discarded, and the hidden layer state is output at each time step. The values of the forgetting, memory and output are controlled by the state of the hidden layer at the last moment and the values of the memory gate, memory gate, and output gate calculated by the current input.

Generally, LSTM includes five computational processes: (1) calculating the forgetting gate and selecting the information to be forgotten; (2) calculating the memory gate and selecting the information to remember; (3) calculating the current cell state at the moment; (4) calculating the output gate and the state of the hidden layer at the current moment; (5) obtaining a hidden layer state sequence of the same length as the sentence. More details are described in [37]. The threshold mechanism of LSTM can effectively filter and memorize the information of the memory unit to solve the problem of RNN. However, the LSTM only captures the forward information from text. For the named entity identification tasks, the backward propagation information has also important reference values. Therefore, the hybrid network is applied in the work in the following section.

2.2. Hybrid Network. Hybrid network level contains the two parts: both the bidirectional LSTM network (BiLSTM) and CRF. The level of BiLSTM is utilized in the sequence tagging task to access both past and future input features. It mainly depends on forward and backward states resulting in two separate hidden states for capturing past and future information, respectively. In this study, the BiLSTM is used to obtain more contextual information. The input sequences $x = (x_1, x_2, \ldots, x_k)$ are put into the neural network. For each input sequence (x_i) in a sentence, it is converted into word embedding. These words in a given sentence are embedded into a BiLSTM network where the forward and backward representation of each word is computed. The symbol $\overrightarrow{h}_t$ is acted as the output of the forward LSTM at a t time, and the symbol $\overleftarrow{h}_t$ is referred as the output representation of the reverse LSTM at t time. The output representation of BiLSTM at t time is defined as $h_t = [\overrightarrow{h}_t, \overleftarrow{h}_t]$. Thus, this

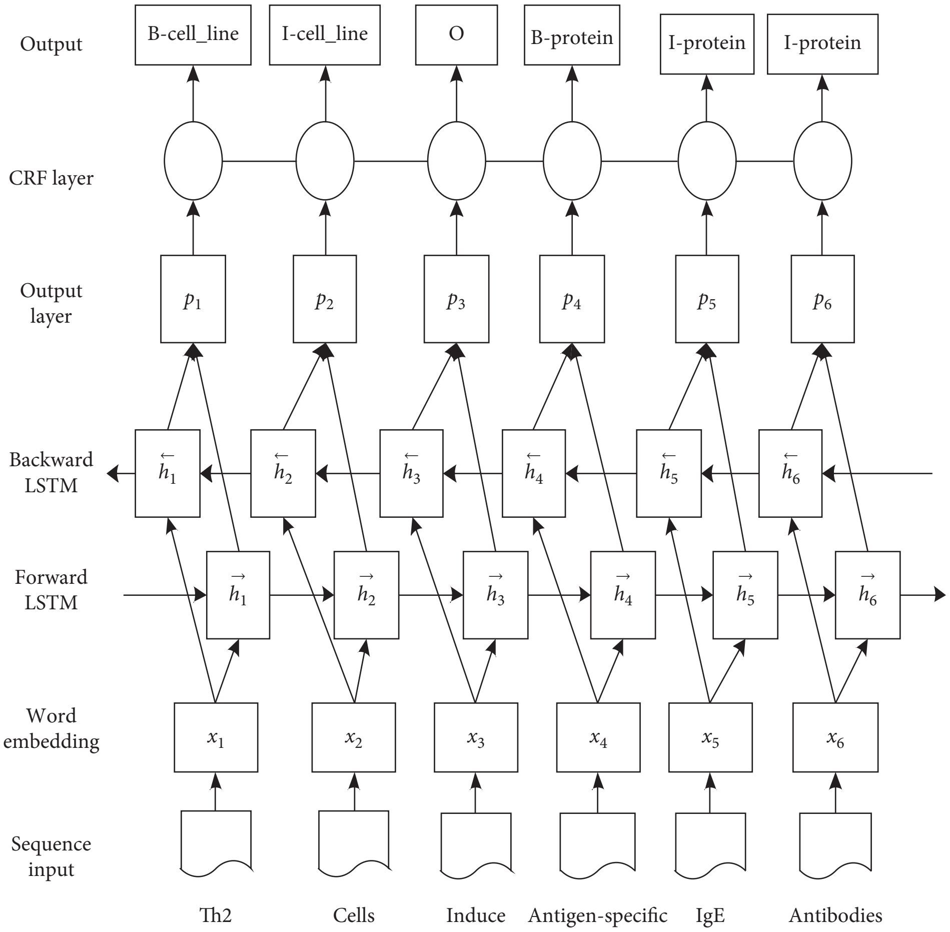


Figure 1: Pipeline of identified entities with the hybrid network.

output contains the more context information. It is used to labelling named entity in the text. The other network level is the condition random field (CRF) model, which focuses on sentence level instead of individual positions in sequence labelling tasks. It makes use of neighbour tag information for predicting the current tags. It is helpful that the correlation between labels in neighbourhoods and jointly decoding the best chain of labels for a given input sentence. Considering the relationship of adjacent labels, the linear CRFs can obtain a globally optimal labelling sequence which could maximize the relationship of adjacent tags. Moreover, it also optimizes the output tag sequences globally and demonstrates enhanced recognition performance for biologically named entities with larger lengths and modified vocabulary. The hybrid network integrated the two network's advantages for more identifying molecular entities.

2.3. Pipeline of Identified Entities. In this study, hybrid network contains both BiLSTM and CRFs. The output of the BiLSTM model is used as the input of the CRFs model to acquire the global optimal marker sequence. Word embedding which is a means of mapping a vocabulary to a real vector for capturing the distributed syntax and semantic information of the words launched by Google is used to switch words to vectors by word2vec. Aiming at the multiword entities, IOB tagging is used to detect entity boundary detection. The label "B" indicates the beginning of the boundary of the entity, the label "I" indicates the intermediate entity, and the label "O" indicates the nonbiological medical entity. Thus, the entity would be tagged as B-entity_category, I-entity_category, and O. For example, when the word is part of protein, it would be tagged as B-protein, I-protein, and O. The pipeline of identified instance "Th2 cells induce antigen-specific IgE antibodies" is shown in Figure 1.

3. Results and Discussion

This study used GENIA [39] corpus to evaluate which is annotated by professional researchers which is a semantically annotated dataset about the biomedical literature to validate the method of entity identification. It also provides the gold standard for the evaluation of text mining systems. GENIA corpus is extracted from the MEDLINE database with MEDLINE ID, title, and abstract encoded in an XML-database. Aiming at the abovementioned approach, we focused on five categories of entities, namely, DNA, protein, RNA, cell-type, and cell-line using three popular measurements which are used the works [40]. The experimental results are illustrated in Table 1. Our approach achieved an

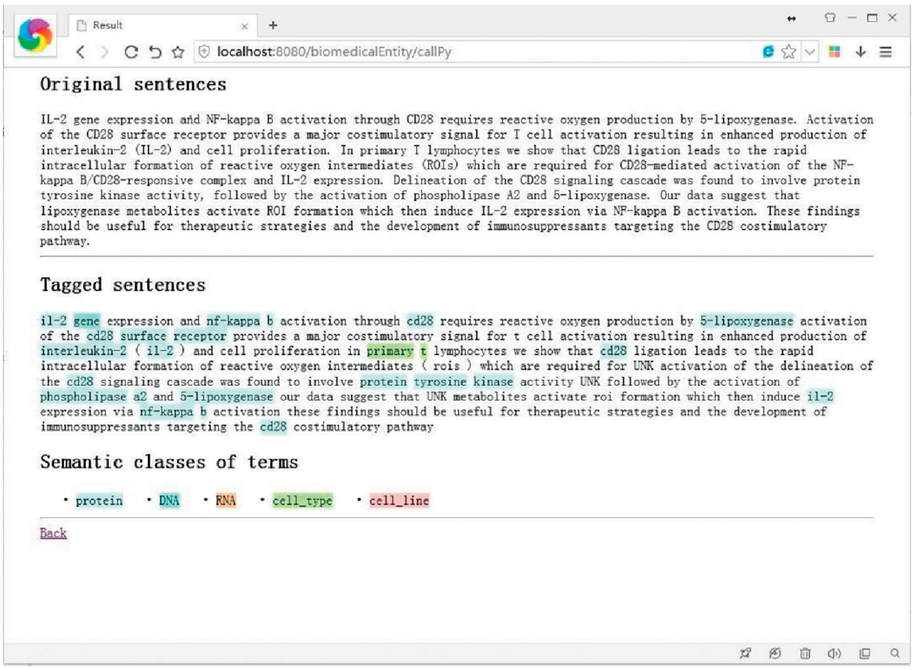
Table 1: Performance of identified molecular entities.

Molecular entity	P (%)	R (%)	F-score (%)
Protein	84.32	80.32	82.27
DNA	76.28	71.33	73.72
RNA	85.71	77.97	81.66
Cell-type	83.67	80.37	81.98
Cell-line	65.22	63.64	64.42
Overall	79.04	74.72	76.81

Compared other previous works, Table 2 illustrates the comparison between our approach and previous works.

Table 2: Comparisons between previous works and our approach.

Approach	P (%)	R (%)	F-score (%)
Zhou et al. [41]	75.99	69.42	72.55
Liao and Wu [42]	72.80	73.60	73.20
Tang et al. [43]	70.78	72.00	71.39
Yao et al. [44]	76.13	66.54	71.01
Li et al. [45]	74.77	70.85	72.76
Li and Guo [46]	**79.58**	69.86	74.40
Our approach	79.04	**74.72**	**76.81**



Original sentences

IL-2 gene expression and NF-kappa B activation through CD28 requires reactive oxygen production by 5-lipoxygenase. Activation of the CD28 surface receptor provides a major costimulatory signal for T cell activation resulting in enhanced production of interleukin-2 (IL-2) and cell proliferation. In primary T lymphocytes we show that CD28 ligation leads to the rapid intracellular formation of reactive oxygen intermediates (ROIs) which are required for CD28-mediated activation of the NF-kappa B/CD28-responsive complex and IL-2 expression. Delineation of the CD28 signaling cascade was found to involve protein tyrosine kinase activity, followed by the activation of phospholipase A2 and 5-lipoxygenase. Our data suggest that lipoxygenase metabolites activate ROI formation which then induce IL-2 expression via NF-kappa B activation. These findings should be useful for therapeutic strategies and the development of immunosuppressants targeting the CD28 costimulatory pathway.

Tagged sentences

il-2 gene expression and nf-kappa b activation through cd28 requires reactive oxygen production by 5-lipoxygenase activation of the cd28 surface receptor provides a major costimulatory signal for t cell activation resulting in enhanced production of interleukin-2 (il-2) and cell proliferation in primary t lymphocytes we show that cd28 ligation leads to the rapid intracellular formation of reactive oxygen intermediates (rois) which are required for UNK activation of the delineation of the cd28 signaling cascade was found to involve protein tyrosine kinase activity UNK followed by the activation of phospholipase a2 and 5-lipoxygenase our data suggest that UNK metabolites activate roi formation which then induce il-2 expression via nf-kappa b activation these findings should be useful for therapeutic strategies and the development of immunosuppressants targeting the cd28 costimulatory pathway

Semantic classes of terms

- protein
- DNA
- RNA
- cell_type
- cell_line

Back

Figure 2: Screen shot of identified molecular entities.

F-score of 76.81%. Table 2 illustrates the comparison between our approach and previous works and previously reported ones.

Zhou et al. [41] identified entities with 72.55% F-score. Liao and Wu [42] used artificial features to construct a skip-chain CRF model that considers long-distance dependencies with an F-score of 73.20% in GENIA corpus. Nevertheless, this paper proposes the BiLSTM-CRF model, which does not use any artificial features but obtains better results in GENIA corpus than the model used by Liao and Wu [42]. Yao et al. [44] used a multilayer neural network learning feature representation and achieved an F-score of 71.01%. Li and Guo [46] constructed a BiLSTM model with word and character vectors and obtained an F-score of 74.40%. Our proposed method obtained an F-score of 76.81%, indicating that our approach is better than those in previous works [42–46]. Thus, it is promising for extracting molecular entities from the biomedical literature.

In this study, we also used the key word "autism" to search the NCBI database, including 29767 literature studies until August 12, 2018. The approach have extracted 9146 proteins, 145 RNAs, 7680 DNAs, 1058 cell-types, and 981 cell-lines after removing repeated molecular entities. In these extracted molecular entities, the MECP2 gene appears

Table 3: The same 70 genes compared to the ripe genes in the work [11].

The same 70 genes compared to the ripe genes in the work [11]						
OT	ERK	RORA	FOXP1	TCF4	CDH13	VEGF
TRPV1	NLGN4	HMGB1	NRG3	UPS	HNF1B	ST8SIA
PAFAH1B1	TNF	FGF22	HDAC4	TLR3	NTK2	CDH8
SCN3A	DIA1	L1CAM	CRK	NOS1	VP	AGC1
CACNA1A	SHOX	ATP8A1	MVP	NR4AL	WNT1	FMR2
SOX5	CRBN	SUSD4	DAT1	MAPT	MTNRLA	ATRNL1
LRRTM3	DLG4	PCDH15	MKL2	RPP25	OGG1	CTCF
SLOS	GLUT1	KIF1A	GRIA1	ID3	BDMR	INS
TSGA14	CRHR1	CD28	GAS	TSC	BF	GATM
MDR1	SOX9	GAP43	ARA	PLA2	FOSB	WMS

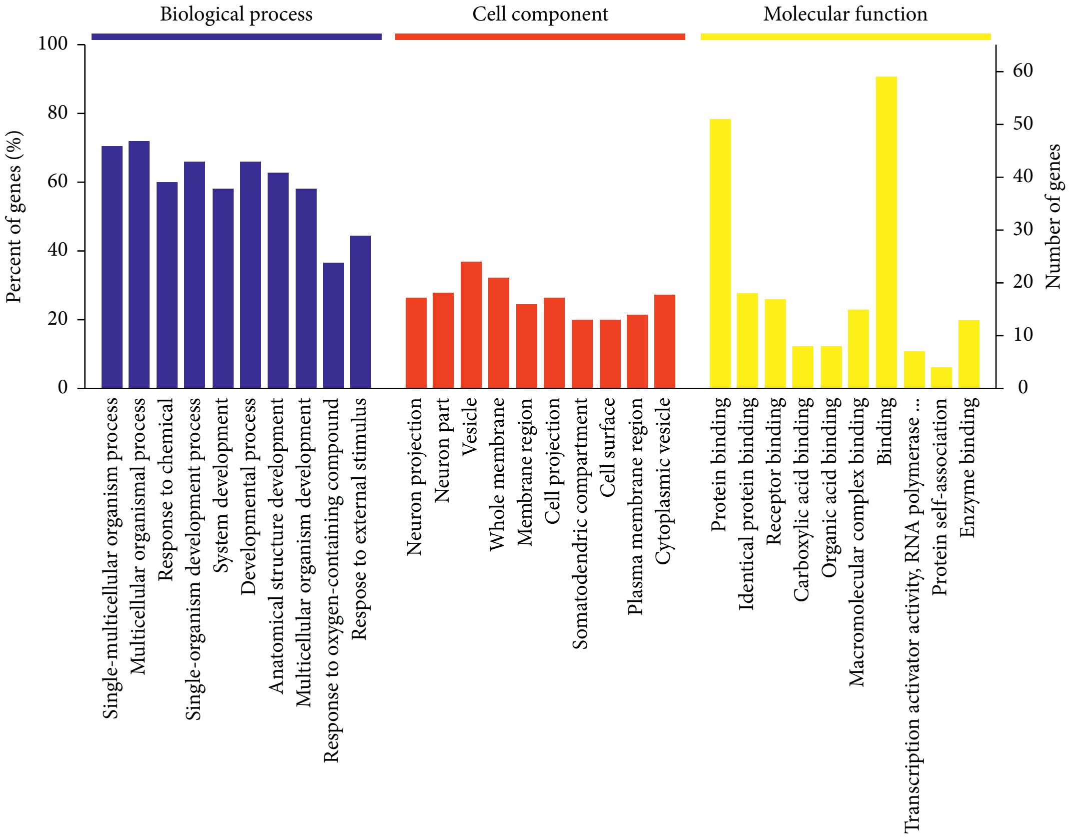


Figure 3: GO analysis related to 70 genes.

most frequently, followed by gene the OXYTOCIN gene in the experimental dataset. The two genes are confirmed as autism susceptibility genes. We used Python to extract molecular entities related to autism and developed an identification system. The screen shot is shown in Figure 2.

Compared to the ripe genes in the work [11], there are the same 70 genes in the extracted entities. They are shown in Table 3. GO and KEGG analyses of the 70 genes are shown in Figures 3 and 4, respectively.

GO analysis showed that about 70% of the genes participate in developmental process and nearly 50% of the genes participate in response to external stimulus as shown in Figure 3. Nearly 30 genes are located in neuron projection and partly in cell component. Finally, about 90% of the genes show binding and protein binding molecular functions.

KEGG analysis indicated that some of these genes are associated with long-term depression, glutamatergic synapse, dopaminergic synapse, and circadian entrainment in the nervous system as shown in Figure 4. About 9% of the genes participate in the MAPK signaling pathway. Both GO and KEGG analyses of the known genes related to autism provide a reference for understanding the molecular

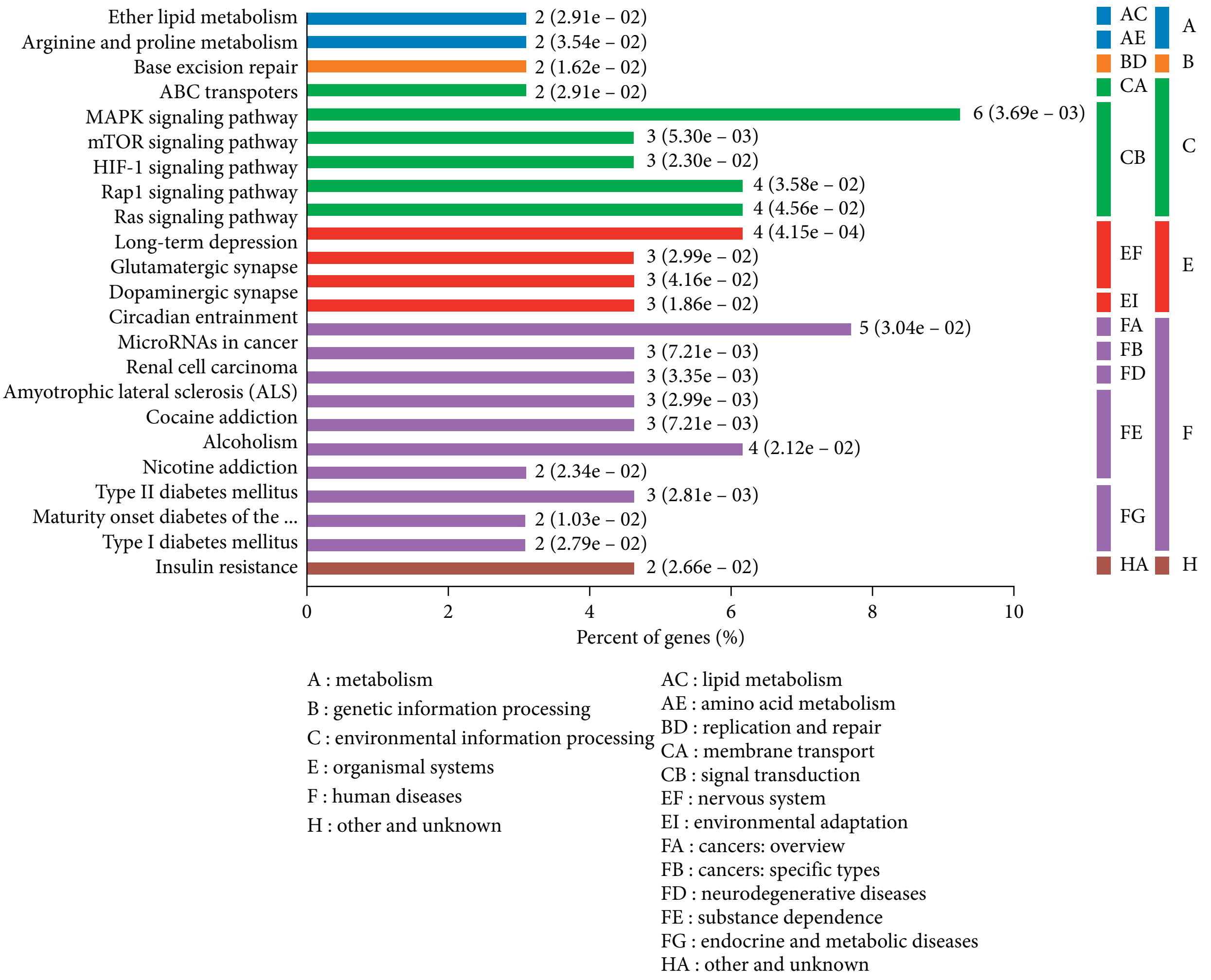


Figure 4: KEGG analysis related to 70 genes.

mechanism of the unknown samples, which could find new genes related to autism.

4. Conclusions

Entities related to disease were identified using the BiLSTM-CRF model, and the approach was evaluated with an F-score of 76.81%. To the best of our knowledge, the provided approach is state-of-art compared the previous works. Based on the approach, we also develop an identified system. Meanwhile, this study also analyses the extracted genes by GO and KEGG analyses. The proposed approach will be applied to explore other molecular mechanisms related to other neurological-diseases, such as Parkinson. This study can serve as a reference for understanding disease etiology, which is promising for identifying disease entities.

Data Availability

The experiment dataset related to the autism biomedical literature was extracted from the PubMed database with E-utilities (http://eutils.ncbi.nlm.nih.gov/corehtml/query/static/eutils_help.html) by using the key word "autism." The biomedical corpus plays an important role in biomedical text mining for achieving the biomedical knowledge domain. It promoted the blossom of text mining technology based on machine learning. GENIA corpus provides a reference material using natural language processing techniques for biomedical text mining. It is a semantically annotated dataset that provides evaluation criteria for text mining approaches. It is also annotated by authoritative domain experts for biological terms encoded in an XML-based markup scheme. This study applied GENIA corpus to build a method about the identification of molecular entities.

Conflicts of Interest

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

Acknowledgments

This research was supported by the National Natural Science Foundation of China (grant nos. 61502243 and 61802193), Natural Science Foundation of Jiangsu Province (BK20170934), Zhejiang Engineering Research Center of

Intelligent Medicine under 2016E10011, China Postdoctoral Science Foundation (2018M632349), NUPTSF (NY217136), and Natural Science Foundation of the Higher Education Institutions of Jiangsu Province in China (16KJD520003).

References

[1] R. Huan, T. Ma, J. Cao, Y. Tian, and A.-D. Abdullah, "Deep rolling: a novel emotion prediction model for a multi-participant communication context," *Information Sciences*, vol. 488, pp. 158–180, 2019.

[2] L. Fu, Z. Li, Q. Ye et al., "Learning robust discriminant subspace based on joint L2,p- and l2,s-norm distance metrics," *IEEE Transactions on Neural Networks and Learning Systems*, 2020, Early Access.

[3] L. Fu, D. Zhang, and Q. Ye, "Recurrent thrifty attention network for remote sensing scene recognition," *IEEE Transactions on Geoscience and Remote Sensing*, 2020, Early Access.

[4] Q. Ye, Z. Li, L. Fu, Z. Zhang, W. Yang, and G. Yang, "Nonpeaked discriminant analysis for data representation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3818–3832, 2019.

[5] Q. Ye, J. Yang, F. Liu, C. Zhao, N. Ye, and T. Yin, "L1-Norm distance linear discriminant analysis based on an effective iterative algorithm," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 114–129, 2018.

[6] Q. Ye, H. Zhao, Z. Li et al., "L1-Norm distance minimization-based fast robust twin support vector k -plane clustering," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4494–4503, 2018.

[7] B. Al-Otibi, N. Al-Nabhan, and Y. Tian, "Privacy-preserving vehicular rogue node detection scheme for fog computing," *Sensors*, vol. 19, no. 4, p. 965, 2019.

[8] Y. Tian, M. M. Kaleemullah, M. A. Rodhaan et al., "A privacy preserving location service for cloud-of-things system," *Journal of Parallel and Distributed Computing*, vol. 123, p. 215, 2019.

[9] Z. Pan, C.-N. Yang, S. Sheng Victor, N. Xiong, and W. Meng, "Machine learning for wireless multimedia data security," *Security and Communication Networks*, vol. 2019, Article ID 7682306, 2019.

[10] T. Ma, R. Huan, Y. Hao, J. Cao, Y. Tian, and Al-R. Mznah, "A novel sentiment polarity detection framework for Chinese," *IEEE Transactions on Affective Computing*, 2019.

[11] L. Gong, R. Yang, and X. Sun, "Prioritization of disease susceptibility genes using LSM/SVD," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3410–3417, Article ID 000327554000020, 2013.

[12] L. Gong, Y. Yan, J. Xie, H. Liu, and X. Sun, "Prediction of autism susceptibility genes based on association rules," *Journal of Neuroscience Research*, vol. 90, no. 6, pp. 1119–1125, Article ID 000302536300002, 2012.

[13] L. Gong, X. Sun, D. Jiang, and S. Gong, "AutMiner: a system for extracting ASD-related genes using text mining," *Journal of Biological Systems*, vol. 19, no. 1, pp. 113–125, Article ID 000288809600007, 2011.

[14] W. W. M. Fleuren and W. Alkema, "Application of text mining in the biomedical domain," *Methods*, vol. 74, pp. 97–106, 2015.

[15] A. Jimeno Yepes and R. Berlanga, "Knowledge based word-concept model estimation and refinement for biomedical text mining," *Journal of Biomedical Informatics*, vol. 53, pp. 300–307, 2015.

[16] S. Hassanpour, M. J. O'Connor, and A. K. Das, "A semantic-based method for extracting concept definitions from scientific publications: evaluation in the autism phenotype domain," *Journal of Biomedical Semantics*, vol. 4, no. 1, p. 14, 2013.

[17] F. Thabtah, F. Kamalov, and K. Rajab, "A new computational intelligence approach to detect autistic features for autism screening," *International Journal of Medical Informatics*, vol. 117, pp. 112–124, 2018.

[18] M. Spencer, N. Takahashi, S. Chakraborty, J. Miles, and C.-R. Shyu, "Heritable genotype contrast mining reveals novel gene associations specific to autism subgroups," *Journal of Biomedical Informatics*, vol. 77, pp. 50–61, 2018.

[19] R. A. Bush, C. D. Connelly, A. Pérez, H. Barlow, and G. J. Chiang, "Extracting autism spectrum disorder data from the electronic health record," *Applied Clinical Informatics*, vol. 8, no. 3, pp. 731–741, 2017.

[20] M. Macedoni-Lukšič, I. Petrič, B. Cestnik, and T. Urbančič, "Developing a deeper understanding of autism: connecting knowledge through literature mining," *Autism Research and Treatment*, vol. 2011, Article ID 307152, 10 pages, 2011.

[21] L. Gong, Y. Yan, J. Xie, H. Liu, and X. Sun, "Prediction of autism susceptibility genes based on association rules," *Journal of Neuroscience Research*, vol. 90, no. 6, pp. 1119–1125, 2012.

[22] L. Zhang, L. Liu, Y. Wen et al., "Genome-wide association study and identification of chromosomal enhancer maps in multiple brain regions related to autism spectrum disorder," *Autism Research*, vol. 12, no. 1, p. 26, 2018.

[23] T. Parr, G. Rees, and K. J. Friston, "Computational neuropsychology and bayesian inference," *Frontiers in Human Neuroscience*, vol. 12, p. 61, 2018.

[24] J. R. Sato, J. Balardin, M. C. Vidal, and A. Fujita, "Identification of segregated regions in the functional brain connectome of autistic patients by a combination of fuzzy spectral clustering and entropy analysis," *Journal of Psychiatry & Neuroscience*, vol. 41, no. 2, pp. 124–132, 2016.

[25] J. R. Sato, M. Calebe Vidal, S. de Siqueira Santos, K. Brauer Massirer, and A. Fujita, "Complex network measures in autism spectrum disorders," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 2, pp. 581–587, 2018.

[26] A. Rosenberg, J. S. Patterson, and D. E. Angelaki, "A computational perspective on autism," *Proceedings of the National Academy of Sciences*, vol. 112, no. 30, pp. 9158–9165, 2015.

[27] S. Jamain, H. Quach, H. Quach et al., "Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism," *Nature Genetics*, vol. 34, no. 1, pp. 27–29, 2003.

[28] A. M. Persico and T. Bourgeron, "Searching for ways out of the autism maze: genetic, epigenetic and environmental clues," *Trends in Neurosciences*, vol. 29, no. 7, pp. 349–358, 2006.

[29] J. F. Abelson, K. Y. Kwan, B. J. O'Roak et al., "Sequence variants in SLITRK1 are associated with Tourette's syndrome," *Science*, vol. 310, no. 5746, pp. 317–320, 2005.

[30] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," 2015, http://arxiv.org/abs/1508.01991.

[31] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: probabilistic models for segmenting and labeling sequence data," *Proceedings of ICML*, vol. 28, 2001.

[32] L. Patel, N. Gustafsson, Y. Lin, R. Ober, R. Henriques, and E. Cohen, "A hidden markov model approach to

characterizing the photo-switching behavior OF fluorophores," *The Annals of Applied Statistics*, vol. 13, no. 3, pp. 1397–1429, 2019.

[33] R. Cofré, C. Maldonado, and F. Rosas, "Large deviations properties of Maximum entropy Markov chains from spike trains," *Entropy*, vol. 20, no. 8, p. 573, 2018.

[34] M. Yin, C. Mou, K. Xiong, and J. Ren, "Chinese clinical named entity recognition with radical-level feature and self-attention mechanism," *Journal of Biomedical Informatics*, vol. 98, Article ID 103289, 2019.

[35] M. Basaldella, L. Furrer, C. Tasso, and F. Rinaldi, "Entity recognition in the biomedical domain using a hybrid approach," *Journal of Biomedical Semantics*, vol. 8, no. 1, p. 51, 2017.

[36] X. Wang, Y. Zhang, X. Ren et al., "Cross-type biomedical named entity recognition with deep multi-task learning," *Bioinformatics*, vol. 35, no. 10, pp. 1745–1752, 2019.

[37] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[38] X. Yang, Y. Li, L. Gong et al., "Bidirectional LSTM-CRF for biomedical named entity recognition," in *Proceedings of the 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, Huangshan, China, July 2018.

[39] J.-D. Kim, T. Ohta, Y. Tateisi, and J. Tsujii, "GENIA corpus--a semantically annotated corpus for bio-textmining," *Bioinformatics*, vol. 19, no. 1, pp. i180–i182, 2003.

[40] L. Gong, R. Yang, Q. Liu, Z. Dong, H. Chen, and G. Yang, "A dictionary-based approach for identifying biomedical concepts," in *Proceedings of the 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery*, Changsha, China, August 2005.

[41] G. Zhou, J. Zhang, J. Su, D. Shen, and C. Tan, "Recognizing names in biomedical texts: a machine learning approach," *Bioinformatics*, vol. 20, no. 7, pp. 1178–1190, 2004.

[42] Z. Liao and H. Wu, "Biomedical named entity recognition based on skip-chain CRFS," in *Proceedings of the 2012 International Conference on Industrial Control and Electronics Engineering*, pp. 1495–1498, Xi'an, China, August 2012.

[43] B. Tang, H. Cao, X. Wang, Q. Chen, and H. Xu, "Evaluating word representation features in biomedical named entity recognition tasks," *BioMed Research International*, vol. 2014, Article ID 240403, 2 pages, 2014.

[44] L. Yao, H. Liu, Y. Liu, X. Li, and M. W. Anwar, "Biomedical named entity recognition based on deep neutral network," *International Journal of Hybrid Information Technology*, vol. 8, no. 8, pp. 279–288, 2015.

[45] L. Li, L. Jin, Y. Jiang et al., "Recognizing biomedical named entities based on the sentence vector/twin word embeddings conditioned bidirectional LSTM," in *Proceedings of the China National Conference on Chinese Computational Linguistics*, pp. 165–176, Kunming, China, October 2019.

[46] L. Li and Y. Guo, "Biomedical named entity recognition based on CNN-BLSTM-CRF model," *Chinese Journal of Information*, vol. 1, pp. 116–122, 2018.

Hindawi
Security and Communication Networks
Volume 2020, Article ID 6679022, 12 pages
https://doi.org/10.1155/2020/6679022

Research Article

Research on Multidomain Authentication of IoT Based on Cross-Chain Technology

Dawei Li,[1,2] Jia Yu,[3] Xue Gao,[3] and Najla Al-Nabhan[4]

[1]*School of Computing Engineering, Nanjing Institute of Technology, Nanjing 211167, China*
[2]*Energy Research Institute, Nanjing Institute of Technology, Nanjing 211167, China*
[3]*State Grid Electric Power Research Institute, Nanjing, China*
[4]*Department of Computer Science, King Saud University, Riyadh, Saudi Arabia*

Correspondence should be addressed to Dawei Li; lidw@njit.edu.cn

Received 22 October 2020; Revised 17 November 2020; Accepted 29 November 2020; Published 10 December 2020

Academic Editor: Zhaoqing Pan

Blockchain is an innovated and revolutionized technology, which has attracted wide attention from academia and industry. At present, blockchain has been widely used in certificate management and credential delivery in network access authentication. In a large-scale multidomain Internet of Things (IoT) environment, one of the important issues is cross-domain key sharing and secure data exchange between different IoT. In this paper, aiming at the multidomain authentication requirements of the IoT, this paper introduces the blockchain cross-chain technology into the cross-domain authentication process of the IoT and proposes an effective cross-domain authentication scheme of the IoT based on the improved PBFT algorithm. First, an architecture of blockchain-based cross-domain authentication is proposed. Then, the block data structure is designed in order to enhance the function of access authentication. Third, the authentication process is realized by intelligent contract. The authentication information is encrypted and distributed by a key sharing method to ensure the security of authentication data. Simulation results show that the proposed scheme has significant advantages in security and availability.

1. Introduction

With the rapid development of 5G and other information and communication technologies (ICTs), the intelligence level as well as deployment scale of Internet of Things (IoT) is increased accordingly. In IoT application scenario, a large number of intelligent terminals work together to collect and process data. The wide area interconnection of IoT brings convenience to users; at the same time, it also brings security risks such as wide attack area, fuzzy security boundary, and poor node controllability [1–4].

The access control of terminals is a key aspect in the security of IoT [5]. With the expansion of the scale of the IoT, the secure access of IoT nodes is not limited to the small-scale trusted authentication of a single security domain, but multi security domain interactive authentication scenarios with business association are becoming more and more common [6–9].

In the traditional IoT authentication method based on centralization, there is an authentication center as an authoritative node for key escrow and certificate management. However, in the multidomain authentication scenario of the IoT, it is difficult to find a trusted authentication authority. A secure mechanism is needed to share the credentials in each security domain for cross domains [10–12].

The most common way of cross-domain authentication is to realize distributed public key authentication through digital certificate and PKI technology. But this method often involves complex certificate management process and has a large cost of computing and storage, which is not suitable for the deployment of low-power IoT systems [13].

In recent years, blockchain technology has been widely used in all walks of life and has produced huge economic and social benefits [14]. Especially in the field of IOT authentication [15], there have been many implementation schemes based on blockchain [16–18]. However, in the current application scenario, each security domain often deploys blockchain system separately, and the blockchain architecture, data structure, and authentication certificate are different. In the large-scale IOT system with multidomain interconnection, there are challenges of authentication data communication and value transfer between public chain, private chain, and alliance chain with different architectures [19–22].

Cross-chain technology is the supporting technology of data asset interconnection and interworking in different blockchain systems. Through the establishment of cross-chain protocol between chains, the cross-domain transmission of data assets or value can be fully trusted. Cross-chain technology provides a feasible solution for the transfer of cross-domain authentication credentials in the IoT.

In order to meet the needs of large-scale multidomain authentication of the IoT, this paper applies the cross-chain technology to the cross-domain authentication certificate transfer, opens up the chain data channel in the authentication system of the IoT, and proposes an effective authentication scheme.

The motivation of the paper is that, on the one hand, the decentralized features of cross chain can decrease the overload of the CA and reduce the problem of single failure; on the other hand, the block chain can transfer cross-domain authentication credentials with the associated blocks in a credible and tamper proof way.

The contributions of the paper are as follows:

(1) We analyze the security requirements of the IoT and propose a blockchain-based cross-domain authentication architecture.

(2) We introduce the cross-chain technology into the multidomain authentication process of the IoT and realize the effective cross-domain transmission and use of the authentication certificate.

(3) We design a block data structure in order to enhance the function of access authentication.

(4) Based on cross-chain technology and distributed consensus mechanism, we realize the authentication process by intelligent contract.

2. Related Work

Compared with the traditional database-based data management, blockchain is a relatively closed system. At present, most of the security application scenarios based on blockchain use independent blockchain systems, which are isolated from each other and difficult to achieve interconnection and horizontal expansion, which hinders the effective transfer and circulation of digital assets between systems. With the popularity of blockchain applications and the complexity of its functions, more and more cross-chain requirements are proposed [23]. In the field of IoT access authentication, when the authentication requirements span multiple security domains of heterogeneous blockchain systems, the interoperability of cross-chain authentication data and remote authentication is particularly important [24, 25]. However, there are few multidomain authentication solutions based on cross-chain technology.

Cross-chain technology can be divided into three mechanisms: notary schemes, side chains/relays, and hash locking. Different cross-chain methods are suitable for different application scenarios. The recommended cross-chain technology in multidomain authentication of the IoT is based on the side chain and relay chain [26]. This technology supports lightweight client-side verification. Through smart contract, it verifies the validity of cryptographic hash tree in the cross-chain system to determine the validity of a specific authentication event and state.

In terms of the implementation mechanism of cross-chain technology, Blockstream put forward the concept of pegged sidechain and studied the transfer mechanism between different blockchain assets in 2014 [26]. Jae and Ethan [27] proposed cosmos, an interoperability architecture between blockchains, which can access different blockchains through inter-blockchain communication protocol. In 2018, Joseph and Vitalik [28] proposed the blockchain expansion design mode for the 2-layer expansion of blockchain and designed the plasma cross-chain system framework with the main chain as the tree root and the slave chain as the branch, which has become the research foundation of many cross-chain technologies. Eykholt et al. [14] proposed an enterprise-level parallel cross-chain platform with high scalability. The platform runs smart contracts concurrently through RhoVM virtual machine and name space to realize multichain interoperability.

Cross-chain technology can realize data interoperability and interoperability between different blockchain systems. It is of great practical significance to apply it to the multidomain authentication field of the IoT. However, this research is still in its infancy, and no mature program has yet emerged.

IoT is an open system with distributed deployment, and its access security is particularly important; traditional access authentication is centralized scheme-based PKI [29–31]. At present, most of the common cross-domain authentication protocols of the IoT are based on distributed public key system, which uses digital certificate for identity authentication [32]. For example, literature [33] established the trust link based on the third-party trust CA to realize the cross-domain authentication of PKI. Literature [34] proposed the cross-domain trust model of PKI based on P2P grid network. Literature [13] proposed a public key infrastructure based on blockchain distributed ledger for the first time. On this basis, the follow-up researchers put forward various improvement schemes, such as PB-PKI [29].

The existing multidomain authentication of the IoT is a large-scale deployment of the same type of system, and the cross-domain authentication of heterogeneous systems is rarely involved. Especially for the IOT deployed with different blockchain platforms, the existing cross-domain authentication schemes are difficult to achieve satisfactory authentication effect.

So, aiming at the security issues in cross-domain authentication of IoT, based on cross-chain technology, we improved the PBFT mechanism using secret sharing protocol and addressed the practical multidomain authentication scheme.

3. Algorithm Description

The basic idea of the proposed algorithm is to improve the traditional PBFT consensus mechanism through identity-based secret sharing algorithm to achieve group authentication for access requests. IBE algorithm is a common public key encryption algorithm in the field of Internet of Things. It can achieve high security strength with short key. As an encryption algorithm, public key and private key appear in pairs, which can only be used in point-to-point encryption and authentication scenarios. When IBE algorithm is applied to the IoT cross-domain distributed authentication scenario, it needs to combine a distributed key management scheme to fragment the single key. In this scheme, a secret sharing algorithm based on Lagrange interpolation is adopted, which encapsulates the key information as the authentication certificate to form a subkey, and each node takes the subkey as the voting basis in the PBFT consensus algorithm. If the node votes in favor, it submits the correct subkey. The number of votes that meet the threshold number indicates that the group authentication has passed.

4. Preparatory Knowledge

IBE public key encryption system takes the character string representing identity as the encryption public key [34, 35]. The algorithm can be implemented by elliptic curve and has the semantic security of adaptive selection ciphertext attack (IND-ID-CCA). The algorithm consists of four algorithms.

(1) Setup: with the security parameter k, generate the system parameter params and master key. The system parameters determine the plaintext space M and ciphertext space C. The system parameters are published through public channels, while the system master key is only secretly stored by the key generation center (PKG).

(2) Extract: with params and master key and identity $\mathrm{ID} \in \{0,1\}^*$ as input, the corresponding private key d is returned.

(3) Encrypt: input params, ID, and plaintext $m \in M$ and output ciphertext $c \in C$.

(4) Decrypt: input params, ciphertext $c \in C$, and private key d and output plaintext $m \in M$.

5. Authentication Mechanism

5.1. Authentication Scenario. In the scenario, terminals of the IoT are divided into several domains, and each domain has a local blockchain which contains the local authentication information of this domain. There is an alliance blockchain which stores metadata of local authentication data in each domain. If cross-domain authentication is required, the authenticator can read the metadata of the authenticated terminal from the federation chain to confirm its access rights.

The common scenario of IoT is shown in Figure 1. The left side is two authentication domains, each maintaining a local authentication blockchain, and each domain has three Internet of Things terminals. On the right is the public authentication blockchain, whose form of existence is alliance chain.

When cross-domain interaction is needed, the local blockchain first verifies the identity of the requester. After the verification, the local authentication information is exchanged to the alliance chain through cross-chain technology. According to the authentication strategy and distributed authentication algorithm based PBFT, the public authentication blockchain completes the authentication of the requesting node. After the authentication is passed, the authentication information is recorded, and the authentication information is exchanged to the local chain of the other domain through the cross-chain technology so as to realize the transfer of the authentication certificate. The specific process of certification is described in detail below.

5.2. Cross-Chain Data Exchange. The data transfer between local chain and alliance chain is realized by side chain of cross-chain technology. Side chain is a technology that allows token to exchange assets safely between different blockchains. The side chain is connected with the main chain through a two-way pegging mechanism. After the connection, the assets on the main chain can be operated to a certain extent through the two-way peg technology.

Through side chain technology, digital assets can be transferred from the first blockchain to the second blockchain and can be safely returned from the second blockchain to the first blockchain at a later time point. The first block chain is usually called the main chain, and the second block chain is called the side chain. By connecting different blockchains together, side chain technology extends the technology of single blockchain, realizes the interoperability between accounts, and ensures the controllable sharing of information in the local domain. The advantage of side chain architecture is that the code and data are independent, do not increase the burden of the main chain, and avoid excessive data expansion. It is a natural fragmentation mechanism.

The core of side chain technology is to realize the cooperation and data interaction between the main chain and the slave chain, which is called “two-way peg.” Two-way peg realizes the flow of the same data assets on the main chain and side chain. When the assets on the main chain are locked, the equivalent side chain assets can be released on the side chain. When the assets on the side chain are locked, the equivalent assets on the main chain are released.

There are several ways to achieve two-way peg.

5.2.1. Symmetrical Mechanism. The main chain and side chain have equal data exchange mode. The two directions carry out equivalent simplified payment verification (SPV) to ensure the authenticity of data in a chain. In data

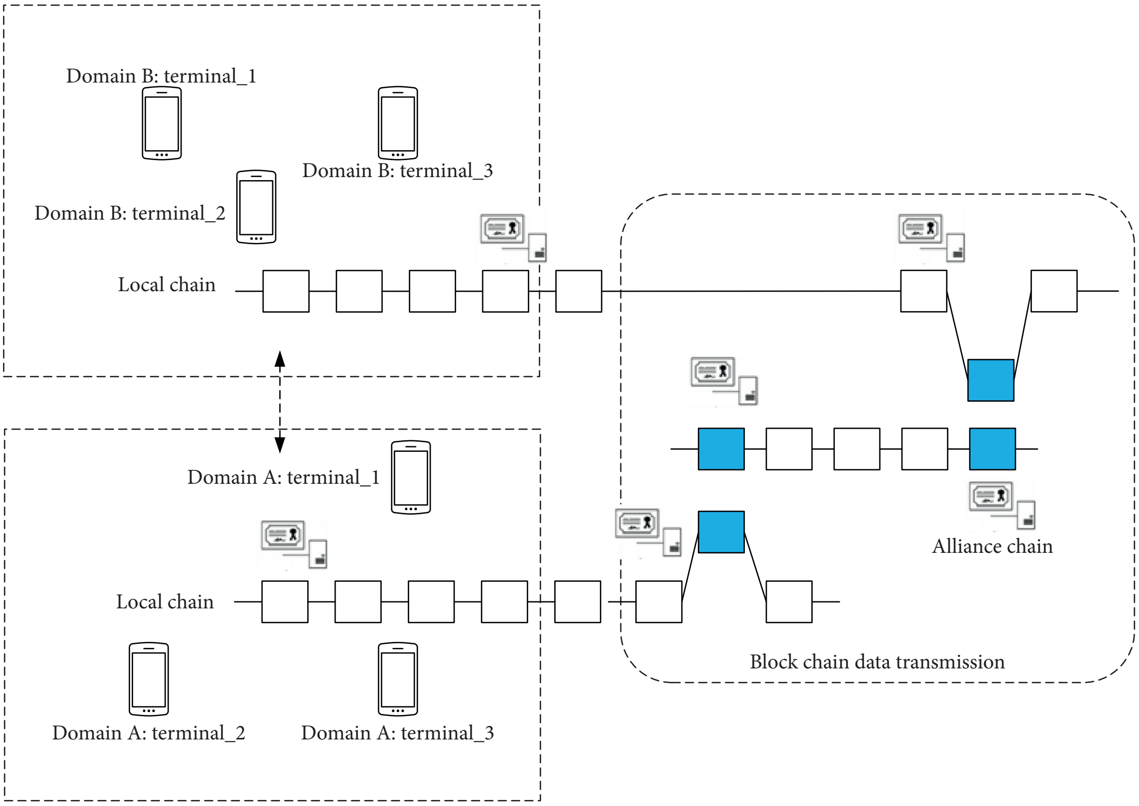


Figure 1: The proposed authentication scenario.

exchange, the status of the main chain and side chain is equal, so it is called symmetric two-way peg.

5.2.2. Asymmetric Mechanism. The information between the side chain and the main chain is asymmetric. The users on the side chain can fully verify the main chain, while the data on the main chain need to be verified by SPV when the data on the main chain are transferred to the side chain. In this mode, the verifier of the side chain needs to synchronize with the main chain.

5.2.3. Single Hosting Mode. A trustee is designated on the main chain to realize the information locking, asset synchronization, and unlocking functions when the side chain of the main chain is synchronized.

5.2.4. Joint Hosting Mode. In this mode, there are multiple hosting centers, and cross-chain data exchange is confirmed in a joint way. In order to achieve security, multisignature mechanism is often used.

5.2.5. SPV Mode. The user sends the data to the main chain. After the confirmation of six blocks on the main chain, the information in the ledger is stored as main chain block. The main chain starts the side chain data update by creating SPV verification.

5.2.6. Driving Chain Mode. Users drive the data interaction between the chains, monitor the status of the side chain, and ensure the data consistency through consensus algorithm.

In the cross-domain authentication of power IoT, each authentication domain has its own authentication policy and certificate, which is stored in the distributed ledger of local chain, namely, side chain. The power IoT system composed of multiple authentication domains maintains an authentication chain as the main chain of the system.

5.3. Data Structure. Block is a data structure for storing ledger. The cross-domain authentication credentials recorded in the block have publicly verifiable and unforgeable attributes. In the cross-chain authentication information exchange, the block data structure defines the description specification, security policy, and security level of authentication certificate.

Cross-domain authentication block is composed of header and data part. Header contains several fields, which are (1) the data used to connect the previous block and index the hash value from the parent block; (2) the timestamp to determine the session aging; (3) the random number used for authentication algorithm; and (4) Merkle tree root data that can summarize and verify all transaction data in the block. As a data carrier, block body stores authentication information through Merkle tree. The data structure is illustrated in Figure 2.

Smart contract of local chain calculates hash value of every node's certificate and forms Merkle tree in blocks.

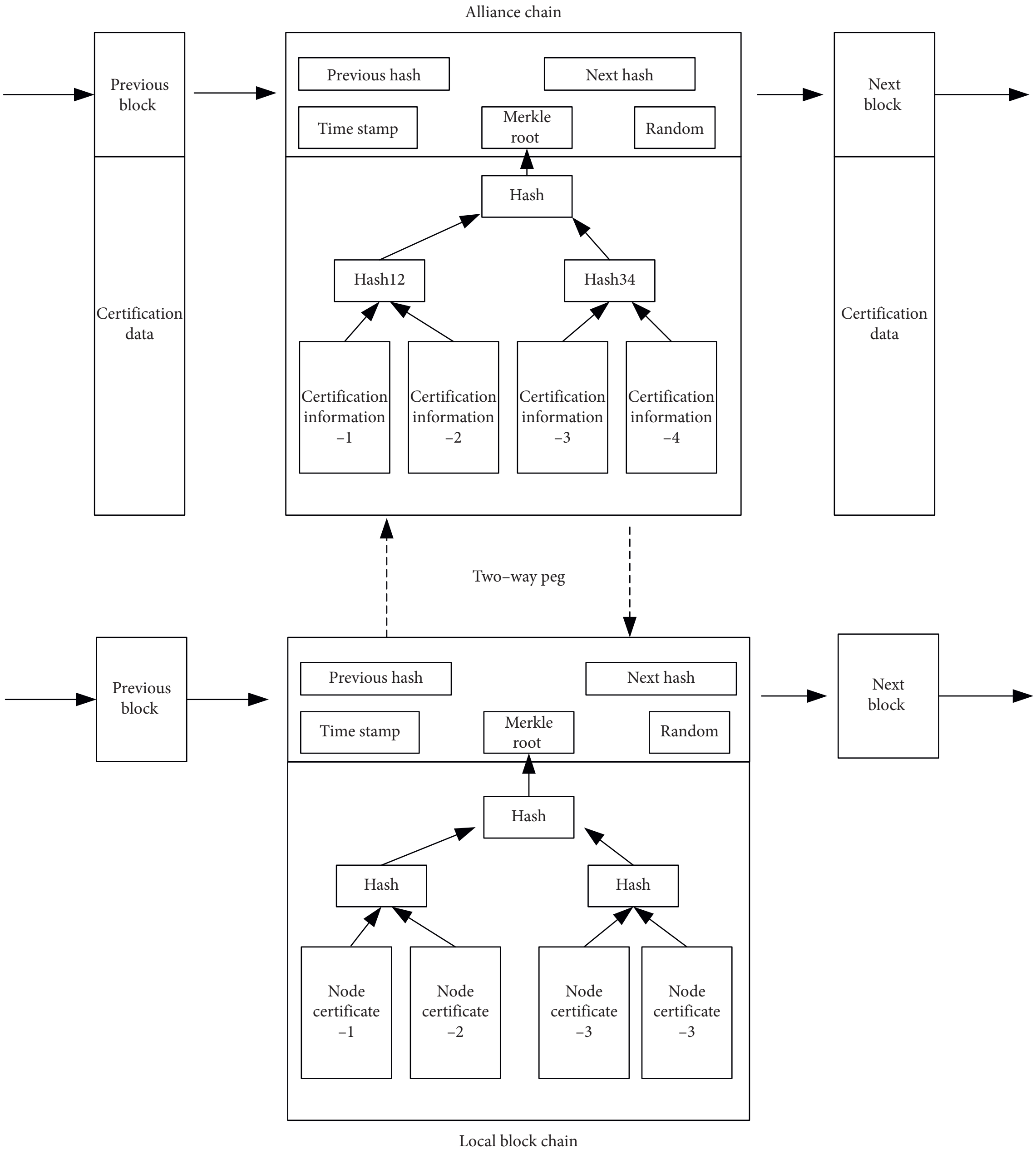


Figure 2: The structure of the certified blockchain.

The certificate includes certificate serial number, public key information, issuer, validity period, signature information, domain ID, etc. Nodes in alliance chain store certification information of cross-domain authentication, and all the information is saved in Merkle tree of the block chain.

5.4. Certification Process. The cross-domain authentication process includes registration, authentication request, credential transfer, distributed authentication, and authentication passing, as shown in Figure 3.

5.4.1. Register. In the registration phase, terminal A in local domain A initiates registration according to the unique ID and triggers the smart contract on the local chain.

The system calls the smart contract and returns the registration information encrypted with A's public key, i.e., Res = $\mathrm{EN}_{\mathrm{Pub}_A}(\mathrm{UUID}||\mathrm{timestamp}||n_1||\mathrm{cert}_A)$, where n_1 is a random number selected by the system.

Terminal A returns the digital signature, $\mathrm{Sig}(\mathrm{UUID}||\mathrm{timestamp}||n_1)$, and completes the three handshake registration interaction processes, and the smart contract on the local chain A is activated and executes registration processing.

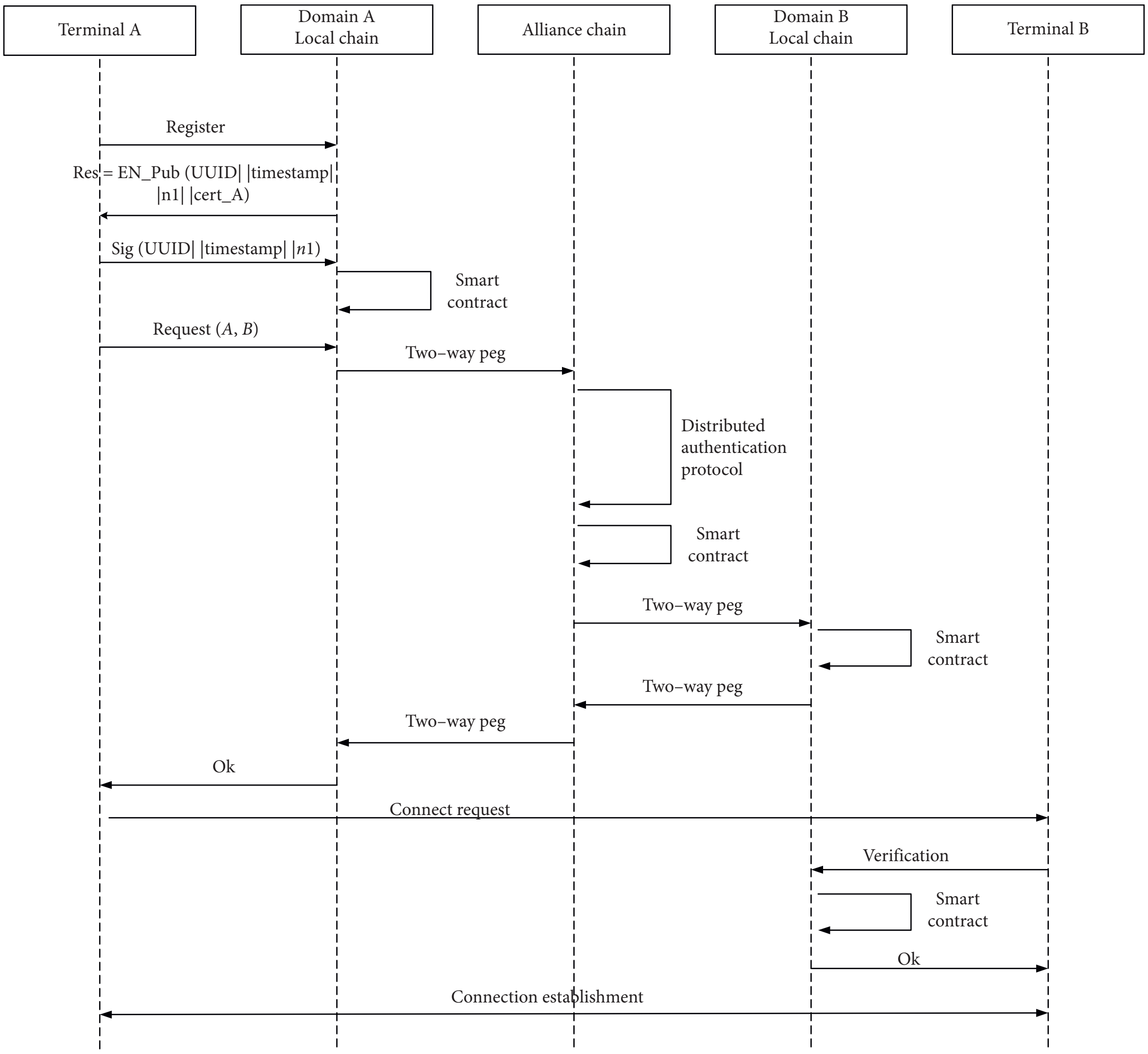


Figure 3: Access authentication process.

5.4.2. Authentication. When terminal A needs to access resources in remote domain B, terminal A initiates authentication request $\text{Request}(A, B) = \text{EN}_{\text{Pri}_A}(\text{ID}_A||\text{Pub}_A||\text{time\_stamp}||\text{businuss\_code})$, which includes local domain ID and remote domain ID, as well as business type.

After the local chain confirms, it initiates the two-way peg with the authentication alliance chain and transmits the authentication request information as well as certificate information of A to the alliance chain synchronously. Thus, the authentication certificate of terminal A and the cross-domain authentication request of A are in the alliance chain.

5.4.3. Verification. The smart contract on the alliance chain performs distributed authentication for the nodes applying for access. Firstly, the nodes meeting the threshold number are selected from the alliance chain to form the authentication group. Secondly, the optimized PBFT consensus algorithm is used for distributed authentication. Finally, the authentication results are stored in the chain. The optimized PBFT consensus algorithm will be described in detail in the next section.

Trigger the smart contract and store the authentication results on the distributed ledger of the alliance chain. Then, the authentication credential information is transferred to the blockchain B in remote domain B by two-way peg. Therefore, the nodes in remote domain B also have the ability to authenticate A.

5.4.4. Connection Establishment. After node A in area A initiates a remote access request to terminal B in remote domain B, terminal B firstly verifies the request and calls smart contract of blockchain B to authenticate A's access request. Because there are communication permissions between A and B in blockchain B, based on the distributed consensus of the blockchain and the unforgeable security attribute, terminal B is very easy to confirm the identity and authority of terminal A, which can be established after authentication secure communication.

5.5. Optimized PBFT Algorithm. Practical Byzantine fault tolerance (PBFT) is a common consensus algorithm in many blockchain application scenarios. It solves the Byzantine error problem in a limited number of nodes by election, and the algorithm performance can be applied to mainstream IoT scenarios.

The distributed cross-domain authentication process of the IoT based on PBFT algorithm is divided into four steps: request, prepreparation, preparation, and submission, as shown in Figure 4.

In the initialization phase, the key management system generates identity-based encryption keys for each IOT terminal, as follows.

Given a security parameter κ, select the large prime number $p(\kappa bit)$, find a hyper singular elliptic curve $E/GF(p)$ that satisfies the CDH security assumption, and generate the order $q (q > 2^{\kappa})$ subgroup $(G, +)$ and its generator P; bilinear mapping $\hat{e}$: $G \times G \longmapsto GF(P^2)^*$.

Select one-way hash functions H_1, H_2, H_3:

$$H_1\colon \{0,1\}^* \longmapsto G^*, H_2, H_3\colon \mathrm{GF}\left(P^2\right)^* \longmapsto \{0,1\}^l. \tag{1}$$

Select the master key $s \in Z_q^*$, calculate the system public key $P_{\text{pub}} = sP$, and return the system parameters: para = $(G, q, P, \hat{e}, H_1, H_2, H_3, P_{\text{pub}})$.

When the IoT terminal with ID as identification in the local domain makes an authentication request to the remote domain, it firstly generates request message by $X = E_{pk}(\text{ID})$, where pk is the public key of ID and $E(\cdot)$ represents the encryption function. The authentication request is transmitted to the master node of the remote domain through the cross-chain technology. The master node runs the smart contract to verify the authenticity of the authentication data transferred across the domain.

After accepting the authentication request, the primary node first finds the legitimate nodes in the security domain to form the authentication group $G = \{P_1, P_2, \ldots, P_t\}$. The master node packages the authentication request data and publishes the subauthentication message to the members of the authentication group, and the system enters the prepreparation stage.

The (t, n) secret sharing mechanism is used to generate the subkey. Assuming that the number of members of the authentication group G is n and the authentication threshold is t, if and only if not less than t nodes submit confirmation, the authentication is deemed to have passed.

The subkey generation process is as follows.

Let IBE ciphertext be $C = <U, V> = \ <rP, R \oplus H_2 (g^r)>$, where $\oplus$ is XOR operation. Choose $t-1$ elements randomly, i.e., $a_1, \ldots, a_{t-1} \in Z_p^*$, and let Lagrange interpolation polynomial be $f(x) = sH_1(\text{ID}) +\ a_1 x + a_2 x^2 + \ldots + a_{t-1} x^{t-1}$; for each voting node, calculate $S_i = f(x_i), 1 \leq i \leq n$ and send S_i and $y_i = \hat{e}(\mathrm{S}_i, P)$ as subkeys.

Calculate and verify key $U_j = \hat{e}(a_j, P_{\text{pub}})$, where $1 \leq j \leq t-1$ and $U_0 = \hat{e}(sH_1(\text{ID}), P_{\text{pub}})$.

In the preparation stage, all nodes in the authentication group conduct P2P broadcast and exchange their own subkeys with each other. The rule is if the cross-domain authentication request of the IoT terminal is agreed and the request information is verified, the subkey held by itself will be disclosed. All participating nodes collect the shared subkeys in the network. When a node in the authentication group collects more than the threshold number of subkeys, the authentication key can be recovered by secret sharing algorithm. At this time, the status is set to the submitted state.

Suppose the subkeys are $(x_1, y_1), \ldots, (x_t, y_t)$, and let the authorization subset of t members be Φ. Subkey receiving node can be calculated as

$$R' = \prod_{j \in \Phi} \hat{e}(S_i, U)^{C_{0,j}^{\Phi}}, \tag{2}$$

where $C_{0,j}^{\Phi}$ is the Lagrange coefficient, defined as: $C_{x,j}^{\Phi} = \prod_{l \in \Phi, l \neq j} ((x-l)/(j-l)) \in Z_q$. Set $\Phi \subset \{1, 2, \ldots, n\}$, $|\Phi| \geq t$.

According to IBE encryption algorithm, $R = V \oplus H_2(R')$.

6. Security Analysis

6.1. Correctness Analysis. If there is authorization subset in access structure Φ, satisfying $|\Phi| \geq t$, then the peer node decrypter can decrypt the ciphertext C to get R according to the shadow secret provided by the member in Φ.

The participant who needs to decrypt R sends the decryption request to the member of Φ and gets the verified shadow secret $\{S_i | i \in \Phi\}$ after the authentication.

Execute bilinear operation of U and S_i; according to Lagrange interpolation theorem, we have

$$R' = \prod_{j \in \Phi} \hat{e}\left(S_j, U\right)^{C_{0,j}^{\Phi}} = \hat{e}\left(\sum_{j \in \Phi} C_{0,j}^{\Phi} S_j, U\right) = \hat{e}\left(D_{\text{ID}}, U\right). \tag{3}$$

Thus, $R = V \oplus H_2(R')$.

6.2. Security Analysis. The security analysis focuses on several attack types which are common in IoT systems. For example, internal and external data source attack, anti-counterfeiting attacks, mutual authentication, middleman attack, Sybil attack, generation attack, single point failure, and so on.

6.2.1. Anti-Internal and Anti-External Data Source Attacks. Through the double-layer structure of local chain and alliance chain, data t_{pub} in the domain are stored in the local chain, and only metadata m_{pub} of local block are stored in the alliance chain, which can be controlled and retrieved through the smart contract using hash function $m_{\text{pub}} = H(t_{\text{pub}})$. The feature of this structure is that the searcher can query and parse the authentication information through metadata specification and get the results that can be publicly verified but cannot get the detailed data, thus protecting the sensitive information in the domain. In addition, the alliance chain uses hash function and other cryptographies to ensure data security and to prevent tampering by illegal users.

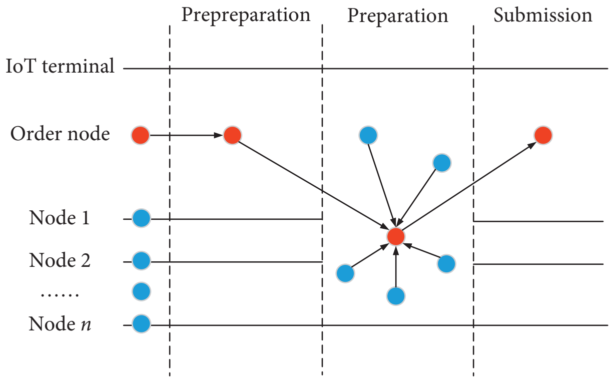


Figure 4: The core stages of consensus algorithm.

6.2.2. Anticounterfeiting Attack. The proposed scheme-based IBE threshold secret sharing encryption system can obtain high security with short key length. The digital certificates are encrypted and stored in distributed ledgers; because of the characteristics of blockchain, it is easy to verify the integrity of certificate. When the authentication group votes for distributed access authentication, the cross-domain authentication key can be calculated only when the adversary obtained more than the threshold number of subkeys, and the attack of malicious nodes can be effectively prevented as long as the threshold t is controlled within a reasonable range.

6.2.3. Antireplay Attack. Replay attack is one of the common attacks in IOT access authentication. By intercepting and resending the information, the adversary can cheat the system. There are three forms of replay attack: one is direct replay, that is, replay to the original verification end; the second is reverse replay, which replays the message originally sent to the receiver to the sender; the third is the third-party replay, which replays messages to other verifiers in the domain. In the scheme, there are timestamps and serial numbers as the basis of message freshness in different stages, such as cross-domain request, intradomain agent encapsulation, cross-domain authentication, etc. If the system finds that there are random numbers used before in the message, it can identify replay attacks easily.

6.2.4. Anti-Sybil Attack. In Sybil attacks, attackers rely on a single node with multiple identities and control most nodes of the system to gain the advantage of voting, which is a common attack in cooperative IOT scenarios. In the proposed scheme, the blockchain is a distributed database that only writes and does not delete. Through redundant data of multiple nodes, network security and nontamperability can be achieved. Multiple identity information of attacking nodes can be easily found by consensus algorithm.

In this scheme, the original PBFT consensus algorithm is improved so that the weight of verifier's voting corresponds to its historical trust value. When the threshold is set to be greater than 2/3 of the number of nodes, it can effectively resist witch attacks. In addition, the verification message U_j of voting broadcast in the proposed scheme can also effectively prevent witch attack.

Table 1: Experimental system configuration.

Item	Configuration parameter
Blockchain platform	Hyper ledger Fabric
Docker version	18.06
Kubernetes	1.9
Operation system	CentOS 7
Authentication station	Intel i7 CPU 16 GB RAM
IoT terminal (type 1)	Raspberry pie 3b+ 1 GB LPDDR2

7. Simulation Analysis

The proposed algorithm is based on alliance blockchain, which requires all the nodes and users to be authenticated and authorized. For example, there are ECert (Enrollment Cert), TCert (Transaction Cert), and TLSCert (Transport Layer Security Cert) integrated by CA of Membership component of Hyperledger Fabric. The ECert certificate is used for identity authentication, which can confirm the identity of nodes and users when logging in the system. TCert certificate is used for signature and verification of transactions. Each transaction contains the signature and transaction certificate of the sender. To ensure that the third party cannot trace the specific sender from the transaction certificate, different TCert certificates can be used for each transaction. TLSCert certificate is used for SSL/TLS communication between system components.

In the simulation environment, multiple x86 servers are used to simulate multiple blockchain nodes in the security domains. Each server is deployed with Hyperledger alliance chain system instance, and the data exchange between multiple instances is realized through cross chain. Each server is interconnected with its IoT terminals. The specific configuration parameters of the system are shown in Table 1.

The simulation environment is the application scenario of the IoT for video capture and monitoring, and the terminal is the camera. The remote camera must pass the cross-domain authentication before sharing data. The authentication scheme is the multidomain authentication scheme based on cross chain proposed above. The simulation topology is shown in Figure 5. Among them, Figure 5(a) shows the network connection mode. The four terminals belong to four authentication domains, respectively, and are connected through switches. The authentication application server is set on the uplink network node and managed and configured by the configuration terminal. Figure 5(b) shows the physical device diagram, including the IoT video terminal, authentication node, and authentication server.

Delay is an important indicator of the efficiency of IoT terminal multidomain access authentication, which directly affects the performance of the upper business system.

We use the scheme which is proposed by Chen et al. in literature [36] as a comparative scheme. Chen's scheme combines the key sharing and distribution protocol in secure multiparty computing with Hyperledger platform and addresses a trusted access authentication scheme for power IoT

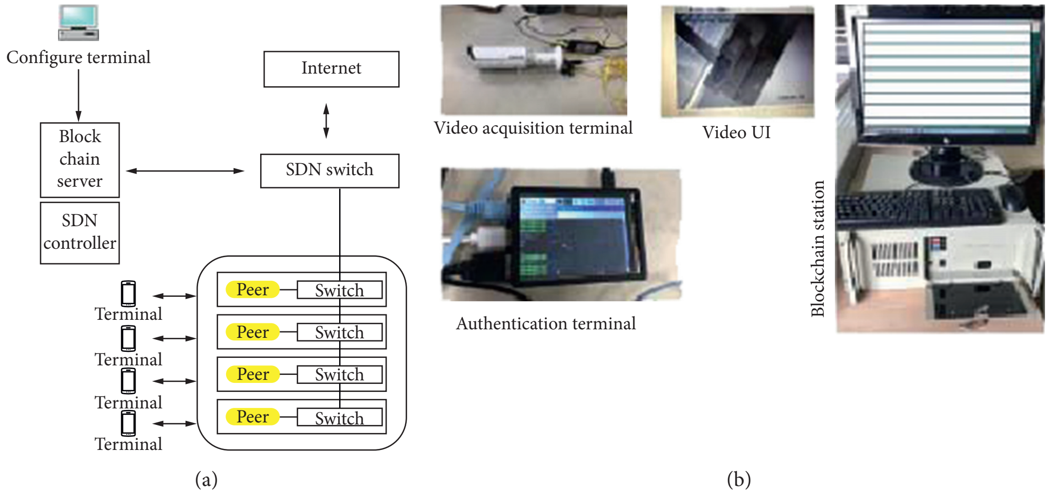


Figure 5: Logical architecture and physical composition of the simulation IoT. (a) Logical architecture. (b) Physical composition.

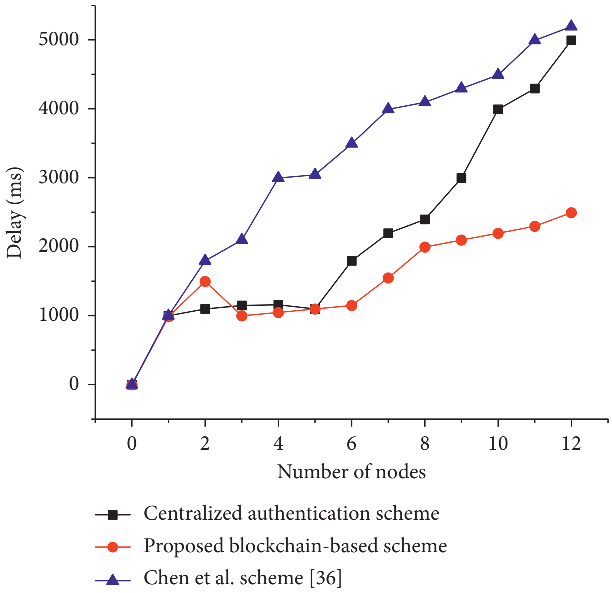


Figure 6: Certification time curve.

terminals. We simulate the system of literature [36] and the proposed cross-chain system in the same environment. The simulation results are shown in Figure 6. From the simulation data, we can see that with the increase of concurrent access requests, the total delay of authentication presents an increasing trend. When the number of nodes increases from 0 to 12, the total delay increases from 2000 ms to 4000 ms. From the horizontal comparison of the three schemes, when the number of concurrent requests exceeds 3, the proposed scheme begins to have performance advantages, that is, in each concurrency level, the proposed scheme is better than the comparison scheme.

In terms of performance, we tested the CPU load in the experimental environment, and the test results are shown in Figure 7. When the number of concurrent nodes is from 0 to 15, the CPU load in the scheme proposed in this paper and in reference [36] increases to about 80%. The difference is that the growth rate of the proposed scheme is slower than that of the comparison scheme. For example, under the condition of less than 5 concurrent nodes, the CPU load of the proposed scheme is less than 5%, but the load of the comparison scheme has increased to 60%. Obviously, the proposed scheme has advantages in the occupation of system resources.

Figure 8 shows the relationship between authentication threshold size and authentication delay. Generally, the larger the threshold value is, the more the legitimate nodes are required for authorization and credit endorsement, and the system security will be improved accordingly. However, the increase of information interaction between nodes in the system with large threshold will directly lead to the increase of authentication delay. In this experiment, the proposed scheme and the contrast scheme have the same change trend. About 10% of the proposed scheme is slightly better than the contrast scheme.

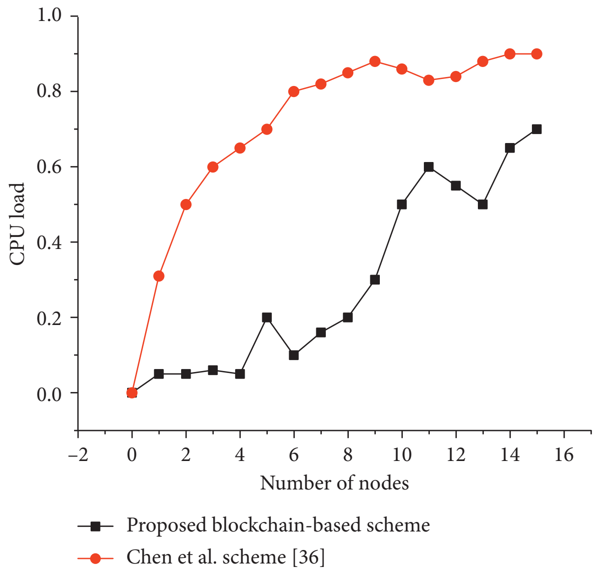


Figure 7: CPU load curve.

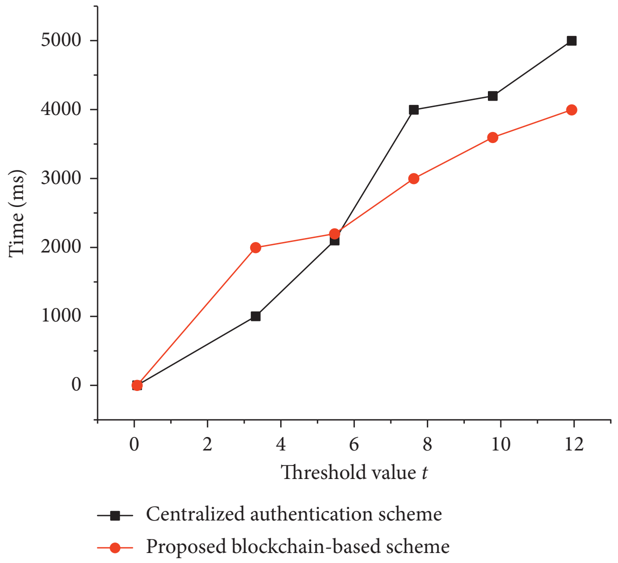


Figure 8: The impact of certification group size.

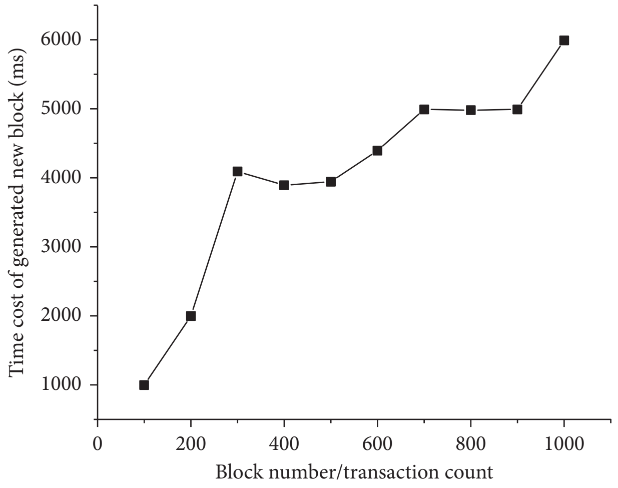


Figure 9: Block generation curve.

An important evaluation index in blockchain system is block speed, which is often associated with TPS. However, from the perspective of blockchain system performance, Figure 9 shows the time of block output under different blockchain heights. It can be seen that with the increase of the height of the blockchain, the retrieval and processing efficiency of the data on the chain decreases, leading to an upward trend in the time delay. Due to the improvement and optimization of the PBFT consensus mechanism, the reduction of the block speed is acceptable.

8. Conclusions

In this paper, a cross-domain authentication method and a model for distributed shared authentication factors are constructed by using the double blockchain structure. The scheme stores authentication data in untouchable blockchains and shares them through public alliance chains. It has high security and good system stability. It can be directly deployed in existing systems and is compatible with local systems. On the basis of ensuring security, it realizes the interoperability of cross-domain terminals.

Further research is to improve the consensus mechanism, improve the authentication efficiency, and adapt to 5G and other new IoT application scenarios. In addition, for the authentication mechanism, distributed identity (DID) and zero trust principle will be added to realize a more flexible authentication mechanism.

Data Availability

The processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Talent Start-Up Project of Nanjing Institute of Technology (YKJ201721) and Natural Science Foundation of Colleges in Jiangsu Province (19KJB520036). The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group no. RG-1441-331.

References

[1] B. Al-Otaibi, N. Al-Nabhan, and Y. Tian, "Privacy-preserving vehicular rogue node detection scheme for fog computing," *Sensors*, vol. 19, no. 4, 2019.

[2] R. Aljably, Y. Tian, and M. Al-Rodhaan, "Preserving privacy in multimedia social networks using machine learning anomaly detection," *Security and Communication Networks*, vol. 2020, Article ID 5874935, 14 pages, 2020.

[3] Y. Tian, M. M. Kaleemullah, M. A. Rodhaan, B. Song, A. Al-Dhelaan, and T. Ma, "A privacy preserving location service for cloud-of-things system," *Journal of Parallel and Distributed Computing*, vol. 123, pp. 215–222, 2019.

[4] Y. Tian, B. Song, M. A. Rodhaan et al., "A stochasticlocation privacy protection scheme for edge computing," *Mathematical Biosciences and Engineering*, vol. 17, no. 3, pp. 2636–2649, 2020.

[5] L. Zhang, H. Li, L. Sun, Z. Shi, and Y. He, "Poster: towards fully distributed user authentication with blockchain," in *Proceedings of the 2017 IEEE Symposium on Privacy-Aware Computing (PAC)*, Washington, DC, USA, August 2017.

[6] O. Abdulkader, A. M. Bamhdi, V. Thayananthan, F. Elbouraey, and B. Al-Ghamdi, "A lightweight blockchain based cybersecurity for IoT environments," in *Proceedings of the The 6th IEEE International Conference on Cyber Security and Cloud Computing (IEEE CSCloud 2019)*, Paris, France, June 2019.

[7] W. N. Qian, Q. F. Shao, Y. C. Zhu, C. Q. Jin, and A. Y. Zhou, "Research problems and methods in blockchain and trusted data managemen," *Journal of Software*, vol. 29, no. 1, pp. 150–159, 2018.

[8] T. Ma, H. Rong, Y. Hao, J. Cao, Y. Tian, and M. A. Al-Rodhaan, "A novel sentiment polarity detection framework for Chinese," *IEEE Transactions on Affective Computing*, vol. 99, p. 1, 2019.

[9] B. Song, et al., A two-stage approach for task and resource management in multimedia cloud environment,"*Computing*, vol. 98, no. 1-. 119–145, 2016.

[10] D. Li, W. Peng, W. Deng, and F. Gai, "A blockchain-based authentication and security mechanism for IoT," in *Proceedings of the 2018 27th International Conference on Computer Communication and Networks (ICCCN)*, Hangzhou, China, July 2018.

[11] Z. Liehuang, G. Feng, and S. Meng, "Survey on privacy preserving techniques for blockchain technology," *Journal of Computer Research an Delelopment*, vol. 54, no. 10, pp. 2170–2186, 2017.

[12] G. L. Millan, M. G. Perez, G. M. Perez, and A. F. G. Skarmeta, "PKI-based trust management in inter-domain scenarios," *Computers & Security*, vol. 29, no. 2, pp. 278–290, 2010.

[13] C. Fromknecht, D. Velicanu, and S. Yakoubov, "CertCoin: a NameCoin based decentralized authentication system," 2014, http://courses.csail.mit.edu/6.857/2014/files/19-fromknecht-velicann-yakoubov-certcoin.pdf.

[14] E. Eykholt, L. Meredith, and J. Denman, "RChain architecture documentation," 2017, https://media.readthedocs.org/pdf/rchain-architecture/stable/rchain-architecture.pdf.

[15] M. Samaniego, U. Jamsrandorj, and R. Deters, "Blockchain as a service for IoT," in *Proceedings of the 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Chengdu, China, December 2016.

[16] A. Ouaddah, A. Abou Elkalam, and A. Ait Ouahman, "FairAccess: a new Blockchain-based access control framework for the internet of things," *Security and Communication Networks*, vol. 9, no. 18, pp. 5943–5964, 2016.

[17] B. Lee and J.-H. Lee, "Blockchain-based secure firmware update for embedded devices in an internet of things environment," *The Journal of Supercomputing*, vol. 73, no. 3, pp. 1152–1167, 2017.

[18] X. Chen, X. Hu, Y. Li, X. Gao, and D. Li, "A blockchain based access authentication scheme of energy internet," in *Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing, China, October 2018.

[19] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of blockchains in the internet of

things: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1676–1717, 2019.

[20] A. Mohammed, V. Potdar, and L. Yang, "Key factors affecting blockchain adoption in organizations," in *Big Data and Security. ICBDS 2019. Communications in Computer and Information Science*Springer, Singapore, 2020.

[21] C. Y. Guan Zhenyu, D. Li, W. Liu, and D. Yu, "A cross-domain authentication scheme for Internet of vehicles based on blockchain," *Cyberspace Security*, vol. 11, no. 9, p. 8, 2020.

[22] A. Moinet, B. Darties, and J. L. Baril, "Blockchain based trust & authentication for decentralized sensor networks," 2017, https://arxiv.org/pdf/1706.01730.pdf.

[23] S. Guo, X. Hu, S. Guo, X. Qiu, and F. Qi, "Blockchain meets edge computing: a distributed and trusted authentication system," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1972–1983, 2020.

[24] L. Kan, Y. Wei, A. Hafiz Muhammad, W. Siyuan, G. Linchao, and H. Kai, "A multiple blockchains architecture on inter-blockchain communication," in *Proceedings of the 2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, Lisbon, Portugal, July 2018.

[25] D. Li and X. Gao, "A blockchain based terminal security of IoT," *ICBDS 2019, Communications in Computer and Information Science*, Springer, Singapore, pp. 445–454, 2020.

[26] B. A. Sidechains, et al., Enabling blockchain innovations with pegged sidechains," 2014, https://blockstream.com/sidechains.pdf.

[27] K. Jae and B. Ethan, "Cosmos: a network of distributed ledgers," 2020, https://github.com/cosmos/cosmos/blob/master/WHITEPAPER.md.

[28] P. Joseph and B. Vitalik, "Plasma: scalable autonomous smart contracts," 2018, https://plasma.io/plasma.pdf.

[29] L. Axon, "Privacy-awareness in blockchain-based PKI," 2015, https://ora.ox.ac.uk/objects/uuid:f8377b69-599b-4cae-8df0-f0cded53e63b/download_file?file_format=pdf&safe_filename=21-15.pdf&type_of_work=Working+paper.

[30] S. Matsumoto and R. M. Reischuk, "IKP: turning a PKI around with decentralized automated incentives," in *Proceedings of the 2017 IEEE Symposium on Security and Privacy (SP)*, San Jose, California, May 2017.

[31] H. Orman, "Blockchain: the emperors new PKI?" *IEEE Internet Computing*, vol. 22, no. 2, pp. 23–28, 2018.

[32] M. Wu, K. Wang, X. Cai, S. Guo, M. Guo, and C. Rong, "A comprehensive survey of blockchain: from theory to IoT applications and beyond," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8114–8154, 2019.

[33] Z. Wen-Fang, W. Xiao-Min, G. Wei, and H. E. Da-Ke, "An efficient inter-enterprise authentication scheme for VE based on the elliptic curve cryptosystem," *Tien Tzu Hsueh Pao/Acta Electronica Sinica*, vol. 42, no. 6, pp. 1095–1102, 2014.

[34] X. M. Lu and D. G. Feng, "An identity-based authentication model for multi-domain grids," *Tien Tzu Hsueh Pao/Acta Electronica Sinica*, vol. 34, pp. 577–582, 2006.

[35] B. Yang, G. Q. Chen, and Y. H. Sun, "Research of a new identity-based authentication model for multi-domain," *Computer Security*, vol. 1, no. 8, pp. 15–18, 2010.

[36] X. Chen, X. Xiaohai, and G. Feng, "Research on distributed authentication of power IoT based on hyperledger blockchain," *Application of Electronic Technique*, vol. 45, no. 5, pp. 57–60, 2019.