IJNS
International Journal of Network Security

ISSN 1816-353X (Print)
ISSN 1816-3548 (Online)

Vol. 25, No. 6 (November 2023)

INTERNATIONAL JOURNAL OF NETWORK SECURITY

PUBLISHING OFFICE

Min-Shiang Hwang
Department of Computer Science & Information Engineering, Asia University, Taichung 41354, Taiwan, R.O.C.
Email: mshwang@asia.edu.tw

International Journal of Network Security is published both in traditional paper form (ISSN 1816-353X) and in Internet (ISSN 1816-3548) at http://ijns.jalaxy.com.tw

PUBLISHER: Candy C. H. Lin

Volume: 25, No: 6 (November 1, 2023)

International Journal of Network Security

Generating Signatures of Unknown Attacks via Flow(s)-Level Anomaly Detection Technology

Xibin Sun[1], Jiwei Qiu[2], Zhenyang Yu[1], and Jianhua Xie[1]
(Corresponding author: Jiwei Qiu)
Guangdong Polytechnic of Science and Technology, Zhuhai, China[1]
Beijing Institute of Technology, Zhuhai 519090, China[2]
Email: jacky5555@qq.com
(Received Jan. 26, 2023; Revised and Accepted Aug. 5, 2023; First Online Sept. 3, 2023)

Abstract

Signature-based intrusion detection usually has a high detection rate for known attacks. However, they cannot detect unknown attacks whose patterns aren't present in their predefined pattern memory. Therefore, extending the capability of signature-based intrusion detection to detect unknown network attacks is crucial. This paper proposes a signature-generation approach that mainly uses an anomaly-based method to find unknown attacks and then extract their patterns to generate signatures in the signature-based intrusion detection system. Unlike other packet-level anomaly detectors, we design a flow(s)-level anomaly detector to find unknown attacks in our signature-generation approach. Experiments on the 1999 Defense Advanced Research Projects Agency (DARPA) dataset show that our anomaly detector performs better in detecting Dos and Probe attacks than packet-level anomaly detectors. Moreover, we also generate new Snort rules for Dos and Probe attacks using a signature-generation approach. The experimental results show that the newly generated rules enhance Snort's ability to detect network attacks and reduce the false positive rate.

Keywords: Flow(S)-Level; Intrusion Detection System; Signature-Based; Unknown Attacks

1 Introduction

As we know, two main approaches, the signature-based method and the anomaly-based method, are used to build a network intrusion detection system (IDS). In the anomaly-based method IDS, the normal profiles are created in advance from the attack-free dataset. If the current activity deviates from the normal profiles, it is considered as abnormal activity. Therefore, the anomaly-based IDS can find unknown network attacks. The signature-based method IDS is usually very powerful in detecting known attacks and does not find unknown attacks whose patterns are not included in the pattern repository.

As described in [20], signature-based IDS usually has the absolute advantages over anomaly-based IDS when deployed in the actual network environment. For example, Snort [18] is a lightweight signature-based IDS that uses predefined rules to detect attacks. However, given the growing number of novel attacks or the increasing diversity of network attacks, signature-based IDS is not suitable for detecting network attacks in a complex network environment. Therefore, we develop a the signature-generation approach to extract patterns of unknown network attacks and then generate signatures in signature-based IDS. The contribution of this paper is given below:

1) We design the flow(s)-level anomaly detector (FLAD) to find novel network attacks. Compared to other anomaly detectors, such as the packet-level head anomaly detector (PHAD), the application-layer anomaly detectors (ALAD) and exponential smoothing Kullback-Leibler distance detector (ES-KLD), our FLAD can achieve better performance on novel network attacks, especially those involving multiple flows.

2) A signature-generation approach is introduced to extract the patterns of novel or unknown attacks and then convert them into signatures that can be used by signature-based IDS to detect subsequent similar network attacks.

3) We apply our FLAD to find Dos and Probe attacks in the 1999 DARPA dataset [21] and use the proposed signature-generation approach to generate rules in Snort according to the extracted patterns. Experimental results show that the newly generated rules not only expend Snort's ability to detect network attacks, but also reduce the false positive rate compared to existing Snort rules.

The remainder of this paper is organised as follows. Section II presents related work on anomaly detection meth-

ods for finding novel attacks. Section III presents our flow(s)-level anomaly detector in detail. Section IV describes our proposed signature-generation approach in detail. Section V mainly analyses the experimental results. Finally, Section VI summarises the work in this paper and points out future work.

2 Related Work

Anomaly detection technologies are effective means of detecting unusual or unexpected activities. Therefore, many researchers have adopted them in the field of network security to detect unknown attacks. In [19], the taxonomy of anomaly-based IDS was presented, which mainly can be divided into three categories: Classification-based methods in [2,4,5], Clustering-based methods in [7,16,17] and Statistical-based methods in [6,13,14]. Classification-based methods are semi-supervised algorithms, which establish normal profiles from the attack-free training dataset. If the current behavior deviation from the normal profiles can be classified as abnormal. For example [5] introduced the enhanced J48 algorithm for anomaly detection, and experiments on NSL-KDD intrusion dataset showed that the proposed approach got 99.88% accuracy rate (AR) with 10-fold cross validation policy. [4] proposed a deep learning approach, Long-Short-Term Memory (LSTM) for anomaly detection and eventually achieved 85% accuracy on the coburg intrusion detection data sets (CIDDS).

Clustering-based anomaly detectors are mainly related to unsupervised learning algorithms that don't require a pre-labelled dataset. Clustering for anomaly detection is usually based on the following strong assumptions: normal instances have an overwhelmingly large fraction as attack instances and there is a clear boundary between normal and attack instances. For example, in [7], density-based spatial clustering of applications with noise (DBSCAN) was proposed for anomaly detection. Compared with the traditional k-means algorithm, DBSCAN algorithm-based IDS has higher accuracy. [16] integrated Sub-Space with One Class Support Vector Machine (OCSVM) to detect unknown attacks, and experiments with the NSL-KDD intrusion dataset showed that the proposed method has better performance compared to existing approaches.

Statistical-based methods mainly use statistical properties to structure the normal profile from the attack-free network dataset and apply statistical inference to determine whether the new instance deviates from the normal profile. the statistical-based approaches presented in [6, 13, 14] monitor protocol packet fields for anomaly detection. For example, packet header anomaly detector (PHAD) [13] monitors 33 fields of the entire transfer control protocol (TCP)/internet protocol (IP) stack packet header to establish normal profiles, and [14] examines the text-based fields of application layer protocols (e.g. Hypertext Transfer Protocol, File Transfer Protocol, and Simple Mail Transfer Protocol) to build models of normal behavior for anomaly detection. In [6], an anomaly detection mechanism using the exponentially smoothing kullback-leibler distance (ES-KLD) is proposed to detect Dos and DDos attacks. Experiments on three datasets showed that the ES-KLD approach could achieve the best detection rate and the lowest false positive rate compared to other anomaly detection solutions.

The hybrid IDS presented in [3, 11, 15] combine the anomaly-based detector simultaneously with the signature-based detector. [3] proposed an improved hybrid IDS that combined the results of the signature-based detector and the anomaly-based detector (decision tree based on the C4.5 algorithm) to obtain the final results. The experimental results showed that the proposed hybrid IDS achieved an accuracy rate of 99.8% and a false positive rate of 0.1%. [15] presented a hybrid IDS for detecting known and unknown attacks in the Internet of Things (IOT) and the proposed hybrid IDS achieved a detection rate of 96.9%. The above hybrid IDS only combine the results of the signature-based detector and the anomaly detector to improve the detection accuracy, but don't generate signatures of unknown attacks in Signature-based IDS.

3 Proposed Anomaly Detector

Statistical-based anomaly detectors often have better semantic analysis of the detected results than other anomaly detectors, such as classification-based and clustering-based anomaly detectors. For example, statistical-based anomaly detectors can tell which features contribute most to the currently detected attack, and network security administrators can use these features to create rules in the signature-based IDS to detect subsequent similar attacks. Therefore, we developed a statistical-based anomaly detector, flow(s)-level Anomaly Detector (FLAD), to detect new or unknown attacks.

3.1 Packet-Level Anomaly Detector

The packet-level anomaly detector mainly detects anomalies in the packet fields, such as packet header fields and packet payload fields. in [13], a packet header anomaly detector (PHAD) was proposed to detect anomalies, which mainly learns the normal profile by monitoring 33 packet header fields, such as source IP address, destination IP address, source port destination port, protocol, and so on. In PHAD, the anomalous score of each packet header field is formally defined by

$$Score_{field} = \frac{tn}{r} \tag{1}$$

Where r is the distinct anomaly values during the training period, n is the total observed times in the training phase, and t is the time since the last anomaly in the current field. In PHAD, the anomalous score of the whole packet

is formally defined by

$$\text{Score}_{\text{packet}} = \sum_{i \in \text{anomalous fields}} \frac{t_i n_i}{r_i} \quad (2)$$

3.2 Proposed Flow(s)-Level Anomaly Detector

The network flow is defined in [12] and consists of a series of IP packets that share the following characteristics: Source IP address, Destination IP address, Source port, Destination port, and Protocol identifier. Packet-level features contain the information about the packet header or payload of the packet. Flow-level features usually express the aggregated information of related packets in a flow, and flow(s)-level features contain the aggregated information of multiple flows. The packet-level detector can find attacks whose patterns are contained in the packet header or payload. However, it often suffers from encrypted packet payload. Moreover, it cannot find attacks whose patterns occur in multiple flows, such as denial-of-service(Dos) attacks, worm attacks, flooding attacks, etc. Therefore, the flow(s)-level detector has more advantages in identifying communication patterns between hosts than the packet-level detector.

We extracted 28 different statistical flow(s)-level features from the KDDCup'99 dataset [9]. Table 1 shows the details of the extracted flow(s)-level features. We develop the flow(s)-level Anomaly Detector (FLAD) to find anomalies affecting single or multiple traffic flows. Inspired by PHAD, our FLAD creates a non-stationary normal profile based on the 28 extracted flow(s)-level features for anomaly detection. We can use [1] tools to generate the flow(s)-level datasets from the offline file (pcap). The flow(s)-level datasets provide information about the semantic information of the network communication at a higher level than the packet-level traffic. Therefore, FLAD has more advantages in finding anomalies whose patterns span multiple flows compared to PHAD.

In addition, we found that the PHAD approach ignores the high-frequency characteristics of nonanomalous fields. For example, sending an email once in the mail bomb attacks can be considered normal behavior. However, high-frequency email behavior within a short period of time may belong to aggressive behavior. When normal email traffic occurs in the attack-free training dataset, the PHAD method considers the packet-level fields associated with the email traffic as normal. Therefore, PHAD doesn't compute anomalous values in the detection phase, resulting in a high false negative rate for mail bomb attack detection. To address the shortcomings of PHAD, we choose flow(s)-level features that can express the frequency characteristics of communication between hosts. Regardless of whether the current flow(s)-level feature belongs to the anomalous fields, our FLAD counts its anomalous score in the detection phase. Therefore, for each flow(s)-level feature, FLAD calculates the anomalous score as follows:

$$\text{Score}_{\text{feature}} = \begin{cases} \frac{\text{tn}}{\text{r}}, & \text{feature} \in \text{anomalous features} \\ \frac{\text{f}}{\text{t}'}, & \text{feature} \notin \text{anomalous features} \end{cases}$$

Where r is the distinct anomaly values during the training period, n is the total observed times in the training phase, and t is the time since the last anomaly in the current field. t' is the time since the last equal value in the current field, and f is the observed times of the current value of the feature during the detection phase. In FLAD, the anomalous score of the flow is formally defined by

$$\text{Score}_{\text{flow}} = \sum_{\text{feature} \in \text{flow}} \text{Score}_{\text{feature}} \quad (3)$$

3.3 Proposed Signature-Generation Approach

In this work, we developed a signature generation approach that mainly uses FLAD to first find unknown attacks and then extract their patterns that can be used by signature-based IDS to detect similar attacks. Our FLAD belongs to the statistics-based anomaly detection methods, so it provides a better explanation for the predicted results. For example, FLAD can tell us which top N features contribute to the current detected anomalies. Therefore, we use FLAD to find the unknown attacks and then extract their patterns from the flow(s)-level features that contribute most to their detection. For example, we use FLAD to find mail bomb attacks from the 1999 DARPA intrusion detection offline evaluation dataset. The experimental results show that the flow(s)-level feature (srv_count=8) contributes the most to mail bomb attack detection. Therefore, the statistical values (srv_count=8, destination_ip=SMTP_server, destination_port=25, protocol_type=tcp) of FLAD can be considered as the final patterns of mail bomb attacks. According to the generated patterns, the signature-based IDS can use the following rule (srv_count=8, destination_ip=SMTP Server, destination_port=25) to detect mail bomb attacks in the future. The flowchart of the proposed signature-generation approach is shown in Figure 1.

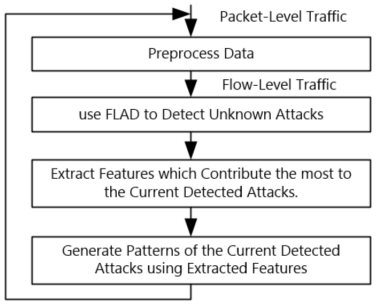


Figure 1: The flowchart of signature-generation approach

Table 1: The selected flow(s)-level features from KDDCup'99 dataset

No.	Feature Name	Granularity	Type
1	duration	flow-level	continuous
2	protocol_type	flow-level	symbolic
3	service	flow-level	symbolic
4	flag	flow-level	symbolic
5	src_bytes	flow-level	continuous
6	dst_bytes	flow-level	continuous
7	land	flow-level	discrete
8	wrong_fragment	flow-level	continuous
9	urgent	flow-level	continuous
10	count	flows-level	continuous
11	srv_count	flows-level	continuous
12	serror_rate	flows-level	continuous
13	srv_serror_rate	flows-level	continuous
14	rerror_rate	flows-level	continuous
15	srv_rerror_rate	flows-level	continuous
16	same_srv_rate	flows-level	continuous
17	diff_srv_rate	flows-level	continuous
18	srv_diff_host_rate	flows-level	continuous
19	dst_host_count	flows-level	continuous
20	dst_host_srv_count	flows-level	continuous
21	dst_host_same_srv_rate	flows-level	continuous
22	dst_host_diff_srv_rate	flows-level	continuous
23	dst_host_same_src_port_rate	flows-level	continuous
24	dst_host_srv_diff_host_rate	flows-level	continuous
25	dst_host_serror_rate	flows-level	continuous
26	dst_host_srv_serror_rate	flows-level	continuous
27	dst_host_rerror_rate	flows-level	continuous
28	dst_host_srv_rerror_rate	flows-level	continuous

4 Experiments and Results

4.1 Datasets

We evaluated our FLAD approach using the 1999 DARPA offline evaluation dataset [21]. The 1999 DARPA dataset contains three weeks of offline traffic as a training dataset and two weeks as a test dataset. In the training dataset, the first and third week data do not contain attacks, so they are normally used to train the normal profiles in the anomaly detection methods. The test dataset consists of the fourth and fifth week data and contains 201 instances of 58 attack types. We use the test data to evaluate the performance of the anomaly detection methods.

4.2 Experiments and Results

The C++ language is used to implement our proposed FLAD approach. In FLAD, the selected flow(s)-level features from the KDDCup'99 dataset can be divided into three categories: discrete, symbolic, and continuous. For the discrete or symbolic features, such as land, urgent, flag, and service, we directly store all the values that appear in the training dataset when we construct the normal profile model. For the continuous features, such as src_bytes, dst_bytes, and count, which are inspired by PHAD, we use the cluster-32 approach to store all values.

We select 7 days of the attack-free dataset (week 3) and 9 days of the dataset (weeks 4 and 5, except for day 2 of week 4, which is missing) from the 1999 DARPA dataset as training and test datasets and use them to evaluate the performance of three statistical-based anomaly detection approaches: FLAD, packet header anomaly detector (PHAD) [13] and application layer anomaly detector (ALAD) [14] and exponential smoothing Kullback-Leibler distance detector (ES-KLD) [6].

Table 2 illustrates the final normal model created using our FLAD with cluster-32 method. Figure 2 and Figure 3 show the detection rate (DR) of four statistics-based anomaly detection approaches for detecting Dos attacks and Probe attacks under different thresholds (the number of false positives). From Figure 2 and Figure 3, it can be seen that our FLAD approach has a higher AR than PHAD, ALAD and ES-KLD methods. As shown in Figure 4 and Figure 5, our FLAD method can find more types of Dos and Probe attacks compared to the other

three approaches. This is mainly because FLAD uses flow(s)-level features to find network anomalies. Compared to packet-level features, flow(s)-level features contain the aggregated information of multiple network traffic flows and provide deeper insight into the patterns of Dos and Probe attacks. Moreover, PHAD and ALAD ignore the frequency characteristics of non-anomalous fields and don't count their anomalous values. FLAD counts anomalous values of non-anomalous fields in the detection phase, which reduces the false negative rate in detecting Dos and Probe attacks.

4.3 Extract Patterns of Unknown Attacks

For FLAD, which belongs to statistical-based anomaly detectors, FLAD can show better analysis for the detected network attacks than other anomaly methods. Table 3 and Table 4 show the detected Dos and Probe attacks using our FLAD cluster-32. The column "FA" describes the number of false alarms before the current attack is detected, and the "How Detected" column shows which features contribute the most to the current detected attack. Therefore, we extract the values in the "How Detected" column as patterns or signatures of the currently detected attack. These extracted patterns can be used by the signature-based IDS to detect other similar network attacks.

Moreover, compared to PHAD and ALAD, we found that the features used by FLAD better describe the patterns of Dos and Probe attacks. For example, PHAD detects the mail bomb attacks based on the anomalous value of 253 in the TTL field, and ALAD finds the mail bomb attacks based on the application field in which the SMTP command "main" is lowercase. However, the two fields cannot reflect the high frequency characteristics of mail bomb attacks well. Therefore, they have poor generalization performance in detecting the subsequent variants of mail bomb attacks. As shown in Table 3, the feature srv_count at the flow(s)-level contributes the most to mail bomb detection in our FLAD. As we know, the feature srv_count expresses the number of connections to the same service as the current connection in the last two seconds, so its values can provide insight into the high-frequency characteristics of mail service flooding in a given time interval.

4.4 Performance Analysis of Proposed Signature-generation Approach

In proposed signature-generation approach, we mainly use FLAD to detect unknown attack and then extract the features contained in the "How Detected" column as the final patterns of attacks. To prove the effectiveness of proposed signature-generation approach, we convert the extracted patterns into Snort's rules, and [20] presented the details of generating Snort's rules using the extracted flow(s)-level features. Due to limited space, Table 5 shows the partially regenerated Snort's rules for detecting Dos attacks using the extracted patterns.

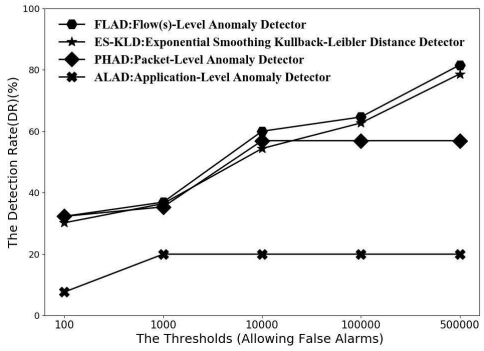


Figure 2: The detection rate (DR) of DoS attacks on different thresholds (Allowing False Alarms)

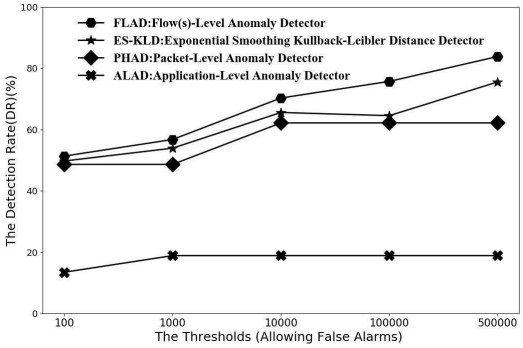


Figure 3: The detection rate (DR) of probe attacks on different thresholds (Allowing False Alarms)

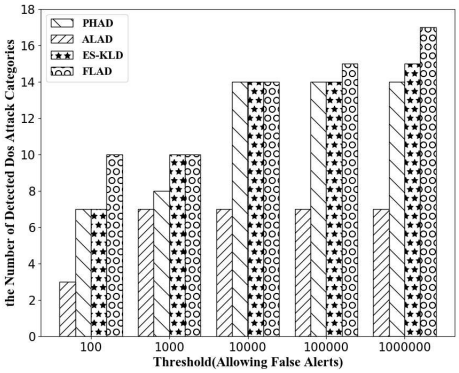


Figure 4: The number of detected DoS attack categories on different thresholds (Allowing False Alarms)

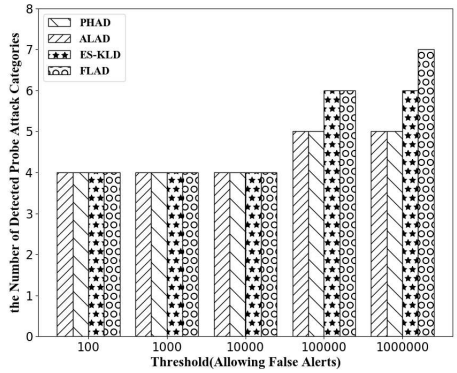


Figure 5: The number of detected probe attack categories on different thresholds (Allowing False Alarms)

To verify the effectiveness of proposed signature-

Table 2: The built normal profile using FLAD method

Feature Name	r/n	Values
duration	1740/354788	0-47726 47929-50894 ...
protocol_type	3/354788	udp, tcp, icmp
service	20/354788	pop_3,X11,netbios_ssn,smtp,telnet,ftp,ftp_data, http,ecr_i,finger,auth,time,ssh,printer,oth_i, eco_i,urh_i,other, ntp_u,domain_u
src_bytes	446/354788	60-390277 407503-649293 ...
dst_bytes	609/354788	0-816043 845350-898194 ..
land	1/354788	0
wrong_fragment	1/354788	0
urgent	7/354788	0 1 3 .....
count	1285/354788	0-114 116-237 239 ...
srv_count	1286/354788	0-109 111-219 221 ...
serror_rate	23/354788	0 0.03 0.04 ...
srv_serror_rate	25/354788	0 0.02 0.03 ...
rerror_rate	21/354788	0 0.04 0.05 ...
srv_rerror_rate	38/354788	0-0.11 0.12-0.24 ...
same_srv_rate	38/354788	0-0.2 0.22-0.48 ...
diff_srv_rate	38/354788	0-0.06 0.07-0.12 ...
srv_diff_host_rate	86/354788	0-0.02 0.03-0.42 ...
dst_host_count	101/354788	0 -31 33-70 72 ...
dst_host_srv_count	101/354788	0-28 30-68 70 ...
dst_host_same_srv_rate	51/354788	0-0.39 0.47-0.68 0.71 ...
dst_host_diff_srv_rate	79/354788	0-0.26 0.37-0.69 0.73 ...
dst_host_same_src_port_rate	45/354788	0-0.14 0.17-0.28 ...
dst_host_srv_diff_host_rate	87/354788	0-0.12 0.15-0.24 ...
dst_host_serror_rate	60/354788	0-0.3 0.31-0.52 ...
dst_host_srv_serror_rate	59/354788	0-0.27 0.28-0.49 ...
dst_host_rerror_rate	63/354788	0-0.41 0.45-0.68 0.73 ...
dst_host_srv_rerror_rate	78/354788	0-0.38 0.43-0.5 0.57 ...

Table 3: The detected DoS attacks using FLAD

Attack Name	FA	How Detected
Teardrop	1	wrong_fragment=1
Land	1	land=1
Neptune	11	dst_host_same_srv_rate=0.99
Apache2	20	serror_rate=0.02
Tcpreset	26	rerror_rate=0.18
Back	28	srv_rerror_rate=0.99
Smurf	47	srv_count=1286
Pod	91	wrong_fragment=1, src_bytes=64000
Processtable	142926	count=3
Dosnuke	276769	srv_count=5
Mailbomb	304885	srv_count=8
Sshprocesstable	308540	srv_count=9

generation approach, we evaluate the performance of our newly generated rules, existing Snort rules [8], and third-party rules [10] separately. As shown in Table 6, the newly generated rules receive a lower FPR and a higher AR than the existing Snort rules and the third-party rules. This is mainly because the flow(s)-level keywords used in our newly generated rules provide better insight into the semantics of Dos and Probe attacks than the packet-level features used in the existing Snort rules and third-party rules.

Table 4: The detected probe attacks using FLAD

Attack Name	FA	How Detected
Insidesniffer	17	urgent=27, serror_rate=0.01
Satan	19	srv_rerror_rate=0.92
Queso	26	state=RSTOS0
Mscan	35	dst_host_same_src_port_rate=0.98
Resetscan	35644	count=2
Ntinfoscan	57489	count=5
Ipsweep	359335	srv_count=80

Table 5: The generated snort rules for detecting DoS attacks using the extracted flow(s)-level features

Attack Name	Generated Snort Rule
Teardrop	alert ip $EXTERNAL_NET any ->$HOME_NET any (msg:"teardrop dos attack";overlapfragment; classtype:attempted-dos; sid:80000; rev:1)
Smurf	alert icmp $EXTERNAL_NET any ->$HOME_NET any (msg:"smurf dos attack";twoSecondsSameDstServiceSessions: $\geqslant$ 1286;itype:0; classtype:attempted-dos; sid:80001; rev:1)
Land	alert tcp $EXTERNAL_NET any ->$HOME_NET [79,23] (msg:"land dos attack";sameip;classtype:attempted-dos; sid:80002; rev:1)
Neptune	alert tcp $EXTERNAL_NET any <>$HOME_NET any (msg:"neptune dos attack";twoSecondsSameDstHostServiceRate: $\geqslant$ 0.99; sessionFlags:S0,REJ,RST0;classtype:attempted-dos; sid:80003; rev:1)
Apache2	alert tcp $EXTERNAL_NET any ->$HOME_NET any (msg:"apache2 dos attack";twoSecondsSameDstHostSynErrorRate: $\geqslant$ 0.2;classtype:attempted-dos; sid:80004; rev:1)
Mailbomb	alert tcp $EXTERNAL_NET any ->$SMTP_SERVER 25 (msg:"mailbomb dos attack";twoSecondsSameDstServiceSessions: $\geqslant$ 8;classtype:attempted-dos; sid:80005; rev:1)
Processtable	alert tcp $EXTERNAL_NET any ->$HOME_NET 23 (msg:"processtable dos attack";twoSecondsSameDstHostSessions: $\geqslant$ 3;classtype:attempted-dos; sid:80006; rev:1)

Table 6: The performance evaluation of snort rules

Source of Rules	Accuracy Rate	False Positive Rate
New Generated Rules	84.16%	16.48%
Existing Snort Rules	72.52%	21.37%
Third-party Rules	74.56%	17.51%

5 Conclusions and Future Work

In this paper, we propose a signature generation approach that mainly uses FLAD to find novel network attacks and generate their signatures in the signature-based IDS using the features that contribute most to their detection. Finally, we take the 1999 DARPA dataset and Snort as experimental subjects to validate our approach. The experimental results show that the proposed FLAD has better performance in detecting Dos and Probe attacks than the packet-level anomaly detector. Moreover, the newly generated rules with the proposed signature generation approach not only expand Snort's detection capabilities, but also improve Snort's performance (e.g., accuracy rate and false positive rate).

At present, our approach mainly focus on the protocol based unknown attacks, and the business-based unknown attacks aren't involved. How to extract the abnormal patterns from the system business data (e.g., log file) is also a problem to be solved in future work.

Acknowledgments

This study is mainly supported by the Guangzhou Science and Technology Program Project "Research on Feature

Engineering Technology and Machine Learning algorithm Applied in the Intrusion Detection System" (Project No.: 202102080586). Furthermore, it is also supported in part by the Special Scientific Research Projects in Key Areas of Education Department of Guangdong Province, China (Project No.: 2021ZDZX1127).

References

[1] Abrahamfi, "The tool for extracting the flow(s)-level features from traffic sniffied with tcpdump," *https://github.com/abrahamfikire/tcpdump2kddcup99*, 2021.

[2] M. Ahmed, M. Mahfouz, and A. Abuhussein, "Network Intrusion Detection Model Using One-Class Support Vector Machine, Advances in Machine Learning and Computational Intelligence," *Algorithms for Intelligent Systems*, vol. 13, no. 1, pp. 79–86, 2020.

[3] M.R.A-G Ahmed and F.M.A Ali, "Enhancing Hybrid Intrusion Detection and Prevention System for Flooding Attacks Using Decision Tree," in *2019 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, pp. 1–4, Khartoum, Sudan, Sep 2019.

[4] I. AI-Turaiki and N. Altwaijry, "A Convolutional Neural Network for Improved Anomaly-Based Network Intrusion Detection," vol. 9, pp. 233–254, 2021.

[5] S. Aljawarneh, M. B. Yassein, and M. Aljundi, "An enhanced J48 classification algorithm for the anomaly intrusion detection systems," *Cluster Computing*, vol. 22, no. 5, pp. 49–65, 2019.

[6] B. Bouyeddou, F. Harrou, B. Kadri, and Y. Sun, "Detecting network cyber-attacks using an integrated statistical approach," *Cluster Computing*, vol. 24, no. 1, pp. 1435–1453, 2020.

[7] D.S Deng, "Research on Anomaly Detection Method Based on DBSCAN Clustering Algorithm," in *5th International Conference on Information Science Computer Technology and Transportation*, pp. 439–442, Shenyang, China, Nov 2020.

[8] Eldon, "The original Snort's rules for detecting Dos attacks," *https://github.com/eldondev/Snort*, 2020.

[9] A. Gharib, I. Sharafaldiny, and A. Habibi, "An Evaluation Framework for Intrusion Detection Dataset," *in Proc. Int. Conf. Info. Sci. Secur. (ICISS)*, vol. 36, no. 2, pp. 1–6, 2016.

[10] A. Gupta and L.S. Sharma, "Mitigation of Dos and Port Scan Attacks Using Snort," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 4, pp. 248–258, 2019.

[11] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "A Novel Ensemble of Hybrid Intrusion Detection System for Detecting Internet of Things Attacks," *electronics*, vol. 8, no. 11, pp. 1–18, 2019.

[12] C. Liu, L. He, G. Xiong, Z. Cao, and Z. Li, "FS-Net: A Flow Sequence Network For Encrypted Traffic Classifiction," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pp. 1171–1179, Paris, France, Apr 2019.

[13] M.V. Mahoney and P.K. Chan. "Phad: Packet header anomaly detection for identifying hostile network traffic,". Tech. Rep. Florida Tech. technical report CS-2001-04, Nov. 2001.

[14] M.V. Mahoney and P.K. Chan. "Learning nonstationary models of normal network traffic for detecting novel attacks,". Tech. Rep. Florida Tech. technical report CS-2001-06, Jul. 2002.

[15] Y. Otoum and A. Nayak, "AS-IDS: Anomaly and Signature Based IDS for the Internet of Things," *Journal of Network and Systems Management*, vol. 29, no. 23, pp. 1–26, 2021.

[16] G. Pu, L. J. Wang, J. Shen, and F. Dong, "A Hybrid Unsupervised Clustering-Based Anomaly Detection Method," *Tsinghua Science and Technology*, vol. 26, no. 2, pp. 79–86, 2021.

[17] B. Ray, S. Ghosh, S. Ahmed, and R. Sarkar, "Outlier detection using an ensemble of clustering algorithms," *Multimedia Tools and Applications*, vol. 81, no. 3, pp. 2681–2709, 2021.

[18] M. Roesch, "Snort-Light Weight Intrusion Detection for Networks," in *Proceedings of LISA'99:13th Systems Administration Conference*, pp. 229–238, Seattle, Washington, USA, Nov 1999.

[19] J. Sen, *Computer and Network Security*. London: IntechOpen, 2020.

[20] X. B. Sun, D. Zhang, H. O. Qin, and J. H.Tang, "Bridging the Last-Mile Gap in Network Security via Generating Intrusion-Specific Detection Patterns through Machine Learning," vol. 2020, pp. 1–20, 2020.

[21] L. Wei, "Cybersecurity Data Science: Concepts, Algorithms, and Applications," in *4th International Conference on Wireless, Intelligent and Distributed Environment for Communication*, pp. 21–30, Cham, Feb 2022.

Biography

XiBin Sun received his Ph.D of Computer Technology and Application in 2022 from Faculty of Information Technology, Macau University of Science and Technology, Macau, China. He is a lecturer in the computer science department of Guangdong Polytechnic of Science and Technology. His current research involves the machine learning and network intrusion detection.

JiWei Qiu received her master's degree. in Philosophy of Science and Technology in 2009 from the Nanchang University, China. She is now a lecturer in the Beijing Institute of Technology, Zhuhai, China. Her current research involves network security technologies.

ZhenYang Yu received his master's degree in Software Engineering in 2019 from the South China University of

Technology, China. He is a lecturer in the computer science department of Guangdong Polytechnic of Science and Technology. His current research involves information hiding and network security.

JianHua Xie received his master's degree in Computer Technology in 2005 from the Chongqing University, China. He is an associate professor in the computer science department of Guangdong Polytechnic of Science and Technology. His current research involves data mining and network security.

Mobile RFID Authentication Protocol Based on Parity Check Patching Operation

Zhen-Hui Li

(Corresponding author: Zhen-Hui Li)

School of Data Science and Computer Science, Guangdong Peizheng College, Guangzhou 510830, China

Email: 2602441@peizheng.edu.cn

(Received Dec. 12, 2022; Revised and Accepted Aug. 5, 2023; First Online Oct. 6, 2023)

Abstract

The traditional RFID authentication protocol can't meet the security requirements of mobile RFID systems. To address this problem, this paper proposes an ultra-lightweight authentication protocol for mobile RFID systems. The proposed protocol encrypts and transmits important information by parity check patching operation. The parity checks patching operation is implemented by a bitwise operation that cleverly combines the hamming-weight parameter of the encrypted information itself. The operation can reduce the introduction of parameters and storage space and increase the difficulty of protocol cracking. Moreover, the GNY formal analysis proves the proposed protocol's security. The performance comparison results show the proposed protocol has advantages in time overhead. The security comparison results show that the proposed protocol can resist common attacks and performs better.

Keywords: Hamming-weight; Internet of Things; Parity Check Patching Operation; Radio Frequency Identification (RFID) System; Ultra-lightweight

1 Introduction

Radio frequency identification (RFID) is a mainstream automatic identification technology [14]. It transmits data through radio signals, so as to realize non-contact identification, real-time positioning and intelligent management of objects [4, 7]. Radio frequency identification system typically includes RFID tags, RFID readers, servers [13]. According to the different application occasions, the system appropriately adds some physical equipment to meet the corresponding needs. RFID technology has been widely used in the Internet of Things due to the advantages of low cost [18], small size [8], long service life and easy deployment [3, 9].

RFID systems can generally be divided into two categories: fixed RFID systems and mobile RFID systems [6]. In the former, a wired channel is used to exchange data between the servers and the readers, it is safe and reliable, and the server is generally fixed. In the latter, a wireless channel is used to exchange data between the servers and the readers. Due to the inherent openness of the wireless channel itself, the channel is unsafe and unreliable, and the servers are generally in flux according to the requirements [2, 17]. With the development of society and the progress of science and technology. The fixed RFID system gradually can't meet the complex and changeable needs of human beings, while the emergence of mobile RFID system can perfectly solve the complex and changeable needs of human beings. In order to guarantee the data security of the system in the wireless channel, various protocols are designed.

Based on the differences between the above two systems, many protocols are only applicable for fixed RFID systems, and can't be used in mobile RFID systems. In order to solve this problem, this paper proposes an ultra-lightweight authentication protocol for mobile RFID system. The protocol uses an ultra-lightweight encryption algorithm to encrypt the transmitted private data, that is, parity check patching operation. This operation is an independent research and design of ultra-lightweight encryption operation based on bitwise operation. Combined with the hamming-weight parameters of the information to be encrypted, the odd parity checks patching operation or even parity checks patching operation in different ways according to the size of the hamming-weight can ensure the security, which reduces the introduction of parameters and the data storage.

2 Related Research Works

In order to solve the problems existing in the application process of RFID system, many experts and scholars designed different types of protocols to ensure the privacy and security of information.

In literature [5], an authentication protocol is designed, which can be used in mobile RFID systems. However, the analysis shows that this protocol has a large amount of computation and cannot be used in tags with low-cost and limited computing power.

Literature [12] propose a mobile RFID authentication protocol based on pseudorandom numbers is provided. However, since the shared key used in the authentication process is not stored in the back-end database, the attacker can cause some privacy data between the label and the back-end database to lose consistency in some ways, that is, the protocol can't resist asynchronous attacks.

In literature [19], an authentication protocol is proposed, which can be used in the mobile RFID system of smart campus. After analyzing the protocol, it is found that part of the private data is transmitted in plaintext mode during the protocol information transmission, and the attacker can obtain the private data by eavesdropping. An attacker, combined with other relevant information, can launch an impersonation attack.

In literature [15], a mobile RFID system authentication protocol applicable to supply chain management is provided, which cleverly uses the physical unclonable function to encrypt the private data to be sent. Due to the advantage that the physical unclonable function can't be cloned, the protocol can resist many attack modes such as impersonation attack and replay attack. However, one end of the database lacks the shared key used for authentication. Therefore, an attacker can launch an asynchronous attack, forcing the privacy data between the database and the label to lose consistency.

In literature [16], an authentication protocol is designed to resist the attacker initiated by the attacker by reducing the communication times. The protocol has certain advantages in terms of computation and communication time. However, because some important data of the protocol is sent in plaintext mode, the attacker can obtain it by eavesdropping, which brings certain security risks to the protocol. For example, the protocol can't prevent impersonation attacks.

In literature [1], a protocol for the mobile RFID system is proposed, But the design of the protocol has defects, which makes the attacker replay some information for many times, resulting in the loss of information consistency between the mobile reader and the server, that is, unable to resist replay attacks and asynchronous attacks.

In literature [11], an ultra-lightweight authentication protocol is proposed, which has certain advantages in terms of computation amount, but it still has some shortcomings. Because the calculation of some messages can be derived from the XOR operation of other messages, the attacker can derive other important messages on the premise of obtaining some messages, and then launch impersonation attacks to obtain more private data.

Aim at shortcomings of the existing protocols, such as the large amount of computation or the inability to resist certain attacks, this paper designs an ultra-lightweight authentication protocol that can be applied to mobile RFID systems. In order to reduce the computation amount, an original encryption algorithm, the parity check patching operation, is presented. This operation is implemented by bitwise operation, so it can greatly reduce the overall calculation amount. The protocol skillfully combines the Hamming weight information carried by the encryption parameter itself without adding new parameters, and carries out different combination and patching operation according to the different hamming-weight sizes, which increases the difficulty of the attacker to crack.

3 The Party Check Patching Operation Definition

In this paper, the symbol $Pcpo(x, y)$ is used to represent Parity check patching operation (Pcpo). The operation is implemented as follows:

1) x, y are binary strings of length L bits; $h(x), h(y)$ are hamming-weights of x, y respectively.

2) When $h(x) \geq h(y)$
Take $h(x)-h(y)$ bits of x from left to right and place it in the corresponding position of z, then take $h(x)-h(y)$ bits after a backward number of $h(x)-h(y)$ bits and place it in the corresponding position of z, and repeat this operation until the value of x is taken. Take $h(x)-h(y)$ bits of y from left to right and place it in the corresponding position of z, then take $h(x)-h(y)$ bits after a backward number of $h(x)-h(y)$ bits and place it in the corresponding position of z, and repeat this operation until the value of y is taken.

   If $h(x)-h(y)$ is odd, the even parity check patching operation is carried out. That is, if the total number of 1s in z is odd, a 1 should be added to the last digit of z to obtain the final encryption operation result. If $h(x)-h(y)$ is even, is only need to add a 0 to the last bit of z to get the final encryption result.

3) When $h(x) \leq h(y)$
Take $h(x)-h(y)$ bits of y from left to right and place it in the corresponding position of z, then take $h(x)-h(y)$ bits after a backward number of $h(x)-h(y)$ bits and place it in the corresponding position of z, and repeat this operation until the value of y is taken. Take $h(x)-h(y)$ bits of x from left to right and place it in the corresponding position of z, then take $h(x)-h(y)$ bits after a backward number of $h(x)-h(y)$ bits and place it in the corresponding position of z, and repeat this operation until the value of x is taken.

   If $h(x)-h(y)$ is odd, the even parity check patching operation is carried out. That is, if the total number of 1s in z is odd, a 1 should be added to the last digit of z to obtain the final encryption operation result. If $h(x)-h(y)$ is even, is only need to add a 0 to the last bit of z to get the final encryption result.

This can be explained by the following example. For example: $x = 010111010110, y = 001010101001$, then $h(x) = 7, h(y) = 5$, based on the above can be obtained $h(x) \geq h(y)$ and $h(x) - h(y) = 2$, according to the above definition can be obtained $z = 011011100101$. Since there

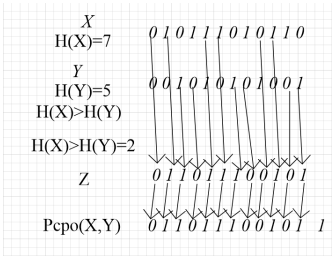


Figure 1: $Pcpo(x, y)(h(x) \geqslant h(y))$

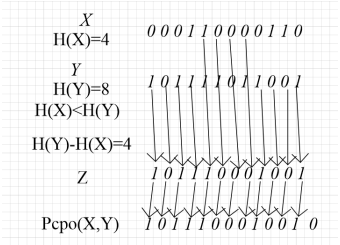


Figure 2: $Pcpo(x, y)(h(x) < h(y))$

are seven ones in z, the number of ones is odd, so the final result is $Pcpo(x, y) = 0110111001011$. As shown in Figure 1.

Again: $x = 000110000110, y = 101111011001$, then $h(x) = 4, h(y) = 8$, based on the above can be obtained $h(x) < h(y)$ and $h(x) - h(y) = 4$, according to the above definition can be obtained $z = 101110001001$. Since there are six ones in z, the number of ones is even, so the final result is $Pcpo(x, y) = 1011100010010$. As shown in Figure 2.

4 RFID Mobile Authentication Protocol

This section describes the meanings of protocol symbols and then describes the protocol steps in detail.

1) Convention of Signs

R represents mobile reader;

T represents label;

DB represents mobile server;

K represents the shared key among R, T and DB;

K_{new} represents the current shared key among R, T and DB;

K_{old} represents the upper round shared key among R, T and DB;

K_{R_DB} represents the shared key between R and DB;

K_{T_DB} represents the shared key between T and DB;

IDR represents the identifier of R;

IDT represents the identifier of T;

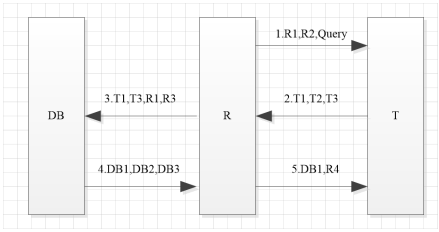


Figure 3: Schematic Diagram of RFID Mobile Authentication Protocol

x represents the random number generated by T;

y represents the random number generated by R;

z represents the random number generated by DB;

$\oplus$ represents the nonequivalence operation;

& represents the and operation;

Ti represents the message calculated by T;

Ri represents the message calculated by R;

DBi represents the message calculated by DB;

$Pcpo(x, y)$ represents the Party check patching operation.

2) Protocol Process

In mobile RFID system, data exchange between mobile reader and label, mobile reader and mobile server is based on wireless channel. The openness of wireless channel itself makes the data exchanged have security risks and insecurity.

Before the protocol starts, each communication entity will store certain data information, such as label stores K, K_{T_DB}, IDT; Mobile reader stores K, K_{R_DB} and IDR; Mobile server stores $K, K_{R\_DB}, K_{T\_DB}, IDR$, and IDT.

The authentication protocol applicable to mobile RFID system in this paper is shown in Figure 3.

In combination with Figure 3, the protocol steps can be described as follows:

Step one. R generates a random number y, calculates the message $R1, R2$, and sends $R1, R2, Query$ to T.

$$R1 = y \oplus K, R2 = Pcpo(y, K).$$

Step two. T received a message, $R1$ deformation processing to obtain $y1$, combined with $y1$, K using the same algorithm to calculate $r2$, compare whether $r2$ and $R2$ are equal.

If $r2$ and $R2$ are equal, $y1 = y$ can be explained. T perform the following operations. T generates a random number x, gets message $T1, T2, T3$ through calculation in turn, and sends $T1, T2, T3$ to R as response information.

If $r2$ and $R2$ are not equal, the protocol stops.

$$\begin{aligned} y1 &= R1 \oplus K \\ r2 &= Pcpo(y1, K) \\ &= Pcpo(R1 \oplus K, K) \\ T1 &= x \oplus K \\ T2 &= Pcpo(K, x) \\ T3 &= Pcpo(x \oplus K_{T\_DB}, IDT). \end{aligned}$$

Step three. R received a message, $T1$ deformation processing to obtain $x1$, combined with $x1$, K using the same algorithm to calculate $t2$, compare whether $t2$ and $T2$ are equal.

If $t2$ and $T2$ are equal, $x1 = x$ can be explained. R perform the following operations. Get message $R3$ through calculate R and send $T1, T3, R1, R3$ to DB.

If $t2$ and $T2$ are not equal, the protocol stops.

$$\begin{aligned} x1 &= T1 \oplus K \\ t2 &= Pcpo(K, x1) \\ &= Pcpo(K, T1 \oplus K) \\ R3 &= Pcpo(y \oplus IDR, K_{R\_DB}). \end{aligned}$$

Step four. After receiving the message, DB verifies R. After the authentication is passed, T is verified. DB performs subsequent operations only when both R and T pass the authentication.

DB verifies R in the following ways:

DB to $R1$ deformation processing to obtain $y2$, combined with $y2, IDR, K_{R\_DB}, K$, using the same algorithm to calculate, get $r3$. Compare whether $r3$ and $R3$ are equal.

If $r3$ and $R3$ are equal,$y2 = y$ can be explained. DB continues the following steps, that is, starts to verify T.

If $r3$ and $R3$ are not equal, the protocol stops.

DB verifies T in the following ways:

DB to $T1$ deformation processing to obtain $x2$, combined with $x2, IDT, K_{T\_DB}, K$, using the same algorithm to calculate, get $t3$. Compare whether $t3$ and $T3$ are equal.

If $t3$ and $T3$ are equal, $x2 = x$ can be explained. DB continues the following steps.

If $t3$ and $T3$ are not equal, the protocol stops.

If and only if both R and T pass the authentication, DB will perform the following operations:

DB generates a random number z, gets message $DB1, DB2, DB3$ through calculation, then updates the shared key K, and finally sends $DB1, DB2, DB3$ to R.

When $K = K_{new}$, and both R and T pass the authentication, the shared key K is updated as follows:

$$\begin{aligned} K_{old} &= K_{new} \\ K_{new} &= Pcpo(x\&K_{new}, y\&z). \end{aligned}$$

When $K = K_{old}$, and both R and T pass the authentication, the shared key K is updated as follows:

$$\begin{aligned} K_{old} &= K_{old} \\ K_{new} &= Pcpo(x\&K_{old}, y\&z) \\ y2 &= R1 \oplus K \\ r3 &= Pcpo(y2 \oplus IDR, K_{R\_DB}) \\ &= Pcpo((R1 \oplus K) \oplus IDR, K_{R\_DB}) \\ x2 &= T1 \oplus K \\ t3 &= Pcpo(x2 \oplus K_{T\_DB}, IDT) \\ &= Pcpo((T1 \oplus K) \oplus K_{T\_DB}, IDT) \\ DB1 &= z \oplus K \\ DB2 &= Pcpo(z \oplus K_{R\_DB}, IDR\&y) \\ DB3 &= Pcpo(z \oplus IDT, K_{T\_DB}\&x). \end{aligned}$$

Step five. After receives the message, R to $DB1$ deformation processing to obtain $z1$, combined with $z1, IDR, K_{R\_DB}, K$, using the same algorithm to calculate, get $db2$. Compare whether $db2$ and $DB2$ are equal.

If $db2$ and $DB2$ are equal, $z1 = z$ can be explained. R continues the following steps. Get message $R4$ through calculate R, and then the shared key K is updated. Finally, $DB1, R4$ are sent to T.

If $db2$ and $DB2$ are not equal, the protocol stops.

$$\begin{aligned} z1 &= DB1 \oplus K \\ db2 &= Pcpo(z1 \oplus K_{R\_DB}, IDR\&y) \\ &= Pcpo((DB1 \oplus K) \oplus K_{R\_DB}) \\ R4 &= Pcpo(y\&DB3, x\&DB3) \\ K &= Pcpo(x\&K, y\&z). \end{aligned}$$

Step six. After receives the message, T to $DB1$ deformation processing to obtain $z2$, combined with $z2, IDT, K_{T\_DB}, K$, using the same algorithm to calculate, get $db3$. Then combined with $db3, x, y$, using the same algorithm to calculate, get $r4$. Compare whether $r4$ and $R4$ are equal.

If $r4$ and $R4$ are equal, $z2 = z$ can be explained. T continues the following steps. T updates the shared key K. After the shared key is updated, the entire authentication protocol is complete.

If $r4$ and $R4$ are not equal, the protocol stops.

$$\begin{aligned} z2 &= DB1 \oplus K \\ db3 &= Pcpo(z2 \oplus IDT, K_{T\_DB}\&x) \\ &= Pcpo((DB1 \oplus K) \oplus IDT, K_{T\_DB}\&x) \\ r4 &= Pcpo(y\&db3, x\&db3) \\ K &= Pcpo(x\&K, y\&z). \end{aligned}$$

5 Formal Analysis Based on GNY Logic

This section combines GNY [10] logical formalization to analyze the proposed protocol, which is done as follows

1) Protocol Formal Description

$MSG1 : R \rightarrow T : R1, R2, Query$

$MSG2 : T \rightarrow R : T1, T2, T3$

$MSG3 : R \rightarrow DB : T1, T3, R1, R3$

$MSG4 : DB \rightarrow R : DB1, DB2, DB3$

$MSG5 : R \rightarrow T : DB1, R4$

The above protocol is specified in GNY logical language as follows:

$MSG1 : T < *\{R1, R2, Query\}$

$MSG2 : R < *\{T1, T2, T3\}$

$MSG3 : DB < *\{T1, T3, R1, R3\}$

$MSG4 : R < *\{DB1, DB2, DB3\}$

$MSG5 : T < *\{DB1, R4\}$

2) Protocol Initialization Assumption

$SUP1 : (K, IDT, K_{T\_DB}) \in T$

$SUP2 : (K, IDR, K_{R\_DB}) \in R$

$SUP3 : (K, IDR, K_{R\_DB}, IDT, K_{T\_DB}) \in DB$

$SUP4 : R| \equiv \#(x, y, z)$

$SUP5 : T| \equiv \#(x, y, z)$

$SUP6 : DB| \equiv \#(x, y, z)$

$SUP7 : T| \equiv DB \overset{K}{\longleftrightarrow} T$

$SUP8 : DB| \equiv T \overset{K}{\longleftrightarrow} DB$

$SUP9 : R| \equiv DB \overset{K}{\longleftrightarrow} R$

$SUP10 : DB| \equiv R \overset{K}{\longleftrightarrow} DB$

$SUP11 : T| \equiv R \overset{K}{\longleftrightarrow} T$

$SUP12 : R| \equiv T \overset{K}{\longleftrightarrow} R$

$SUP13 : T| \equiv DB \overset{K_{T\_DB}}{\longleftrightarrow} T$

$SUP14 : DB| \equiv T \overset{K_{T\_DB}}{\longleftrightarrow} DB$

$SUP15 : R| \equiv DB \overset{K_{R\_DB}}{\longleftrightarrow} R$

$SUP16 : DB| \equiv R \overset{K_{R\_DB}}{\longleftrightarrow} DB$

$SUP17 : T| \equiv DB \overset{IDT}{\longleftrightarrow} T$

$SUP18 : DB| \equiv T \overset{IDT}{\longleftrightarrow} DB$

$SUP19 : R| \equiv DB \overset{IDR}{\longleftrightarrow} R$

$SUP20 : DB| \equiv R \overset{IDR}{\longleftrightarrow} DB$

3) Agreement Proof Objectives

$GOAL1 : T| \equiv R| \sim \#\{R1, R2\}$

$GOAL2 : R| \equiv T| \sim \#\{T1, T2, T3\}$

$GOAL3 : DB| \equiv R| \sim \#\{T1, T3, R1, R3\}$

$GOAL4 : R| \equiv DB| \sim \#\{DB1, DB2, DB3\}$

$GOAL5 : T| \equiv R| \sim \#\{DB1, R4\}$

4) Agreement Certification Process

Due to the limited length of the text, only the first target of proof $GOAL1 : T| \equiv R| \sim \#\{R1, R2\}$ is chosen as an example to prove.

$\because MSG1 : R \rightarrow T : R1, R2, Query$ and rules $P1 : \frac{P<X}{X \in P}$

$\therefore \{R1, R2\} \in T$

$\because SUP4 : R| \equiv \#(x, y, z)$ and rules $F1 : \frac{P|\equiv(X)}{P|\equiv(x,y), P|\equiv\#F(X)}$

$\therefore T = \#\{R1, R2\}$

$\because$ Rules $P2 : \frac{X \in P, Y \in P}{(X,Y) \in P, F(X,Y) \in P}$, $SUP1 : (K, IDT, K_{T\_DB}) \in T, SUP2 : (K, IDR, K_{R\_DB}) \in R$

$\therefore \{R1, R2\}\# \in T$

$\because$ Rules $F10 : \frac{P|\equiv(X), X \in P}{P|\equiv\#(H(X))}$ and derived $T = \#(R1, R2), \{R1, R2\}\# \in T$

$\therefore T| = \#\{R1, R2\}$

$\because$ Rules $I3 : \frac{P<H(X,<S>)>, (X,S) \in P, P|\equiv P \leftrightarrow Q, P|\equiv\#(X,S)}{P|\equiv Q|\sim(X,S), P|\equiv Q \sim H(X,<S>)}$

$\because SUP11 : T| \equiv R \overset{K}{\longleftrightarrow} T, SUP12 : R| \equiv T \overset{K}{\longleftrightarrow} R$ and $MSG1 : R \rightarrow T : R1, R2, Query$

$\therefore T| = \{R1, R2\}$

$\because$ Freshness definition and derived $T| = \{R1, R2\}, T| = R \sim \{R1, R2\}$

$\therefore GOAL1 : T| \equiv R| \sim \#\{R1, R2\}$ freshness definition and derived

6 Performance Analysis

In this section, the amount of calculation and storage capacity of one label end are analyzed. Meanwhile, the performance related aspects of the protocol are analyzed in combination with the total amount of data exchanged during a complete session. The comprehensive comparative analysis results are shown in Table 1.

Explain some symbols that appear in Table 1: L represents the length of data, $Pcpo()$ represents the amount of calculation for parity check patching operation, *xor* represents the amount of computation for bit-by-bit XOR operation, *and* represents the amount of computation for bit-by-bit and operation, *or* represents the amount of computation for bit-by-bit or operation, *bit* represents bits, $PDO()$ represents the amount of computation for tag identity encryption operation, mod represents the amount of computation for modular operation, $PUF()$

Table 1: Performance Comparison between Different Protocols

Reference	Calculations at One End of Tag	Tag End Storage	Total Data for a Full Session
Ref [15]	$6PUF() + 5xor + 3or$	$3L$	$12L + 2bit$
Ref [16]	$5hash() + 5xor + 4and$	$4L$	$16L + 1bit$
Ref [1]	$7PRNG() + 4xor$	$3L$	$15L + 3bit$
Ref [11]	$1PDO() + 4xor + 2mod + 3and$	$2L$	$13L + 1bit$
This Protocol	$6Pcpo() + 3xor$	$3L$	$14L + 1bit$

represents the amount of computation for physical unclonable function, and $hash()$ is the computation of the hash function, $PRNG()$ is the computation of the pseudo-random function.

In this paper, the calculation amount and the storage capacity of the tag end, and the origin of the total data amount in a complete session are analyzed.

Amount of calculation at one label end: When calculating $y1$, the xor operation is used for the first time. When calculating $r2$, the $Pcpo()$ operation is used for the first time. When calculating $T1$, the xor operation a second time. When calculating $T2$, the $Pcpo()$ operation is used for the second time. When calculating $T3$, the $Pcpo()$ operation is used for the third time. When calculating $z2$, the xor operation is used for the third time. When calculating $db3$, the $Pcpo()$ operation is used for the fourth time. When calculating $r4$, the $Pcpo()$ operation is used for the fifth time. When calculating K, the $Pcpo()$ operation is used for the sixth time. Based on the above, the amount of calculation at one end of the label is $6Pcpo() + 3xor$.

Storage capacity of one end of the label: In this paper, there are three parameters K, K_{T_DB} and IDT stored at one end of the protocol label. The length of each parameter is l, and the storage capacity of the label end is $3L$.

Total data for a complete session: In Step 1, R sends $R1, R2, Query$ messages to T. In Step 2, T sends $T1, T2, T3$ messages to R. In Step 3, R sends $T1, T3, R1, R3$ messages to DB. In Step 4, DB sends $DB1, DB2, DB3$ messages to R. In Step 5, R sends $DB1, R4$ messages to T. Based on the above, the total data of sessions for a complete session is $14L + 1bit$, where $Query$ only needs one bit.

Compared with other classical protocols, this protocol is similar to other protocols in terms of the total data of a complete session and the storage capacity of label end, but it has some advantages in terms of the calculation amount of label end . The encryption algorithms used by other protocols are all lightweight operations, such as hash functions and physical unclonable functions. According to the principle of the encryption algorithm $PDO()$ used by the protocol in literature [11], this algorithm belongs to the ultra-lightweight encryption algorithm, but the protocol also uses modular operation for encryption, resulting in a sharp increase in the overall calculation amount at label end. However, the protocol in this paper only adopts ultra-lightweight $Pcpo()$ operation and bitwise XOR operation, which can make the overall calculation amount of the label end reach the ultra-lightweight level, far less than that of other protocols.

7 Security Analysis

This section will analyze the resistance of the protocol to different attack types from different perspectives.

Mutual Authentication

R verifies T as follows: R verifies T with message $T1, T2$.

R verifies DB as follows: R verifies DB with message $DB1, DB2$.

T verifies R twice, the first time through message $R1, R2$ verifies R, the second time through message $R4$ verifies R.

T verifies DB as follows: T verifies DB with message $DB1, R4$.

DB verifies R as follows: DB verifies R with message $R1, R3$.

DB verifies T as follows: DB verifies T with message $T1, T3$.

Impersonation Attack: An attacker can pose as any entity in the system to launch an impersonation attack. Here the attacker is selected to impersonate as a label to attack.

After receiving the message $R1, R2$, the attacker attempts to crack the message $R1, R2$, but the cracking fails. Therefore, the attacker can only randomly select a parameter as the random number generated by R to participate in the calculation. The attacker obtains message $T1, T2, T3$ through calculation and sends the message to R. The attacker thinks that he can pass the authentication of R, but in fact, the attacker can't succeed at all.

After receiving the message of $T1, T2, T3$, R first deforms $T1$ to obtain the random number generated by the attacker, and then calculates a $t2$ by combining with other parameters. Because the shared key used by the attacker must be different from the shared key

used by R, the $t2$ calculated by R must be different from the $T2$ sent by the attacker. Therefore, the attacker fails to impersonate.

Replay Attack: Due to the open nature of the wireless channel, anyone can listen and get all the messages of a complete session. When the next round of communication begins, the attacker can replay the intercepted message in an attempt to analyze useful privacy information through an entity's authentication.

In the proposed protocol in this paper, all transmitted data is encrypted first and then transmitted. Random numbers are mixed in the encryption process to ensure the freshness of the message, so that the message value of the last round is not the same as that of the current round. When an attacker replays a message, an entity only needs to perform a simple computational authentication to identify the authenticity of the source. Attacker replay attack exposed.

Asynchronous Attack: In order to prevent the inconsistency of shared key values among labels, mobile readers and mobile servers, the protocol saves the shared key used in previous communication on the mobile server. When the mobile server fails to authenticate labels and mobile readers with the current shared key, the mobile server selects the shared key used in the previous session to authenticate them again. In this way, the information consistency between the three can be restored, making the attacker fail to perform asynchronous attacks.

Exhaustive Attack: Message $T1, T2$ is used as an example to perform exhaustive attack analysis.

The attacker listens to get message $T1, T2$, first processes message $T1$ to get $x1 = T1 \oplus K$, and then puts $x1 = T1 \oplus K$ into $T2$ to get $t2 = Pcpo(K, x1) = Pcpo(K, T1 \oplus K)$. The attacker thinks that there is only one unknown parameter K in $t2$ and that the correct value of K can be obtained by means of exhaustion, but the attacker can't succeed at all, for the following reasons:

First, the attacker doesn't know the hamming-weight of K, that is, the value of $H(K)$. Secondly, the attacker doesn't know the hamming-weight of $T1 \oplus K$, that is, the value of $H(T1 \oplus K)$. Third, without knowing $H(K)$ and $H(T1 \oplus K)$, the attacker doesn't know which way the parity make up operation is carried out to get the final result.

Based on the above analysis, the attacker can't obtain the value of the shared key by exhaustive method at all, and the exhaustive attack fails.

Tracking Attack: As long as the attacker has the patience to continuously listen to the message sent by the label, the message can locate the actual location of the label.

In order to avoid this situation, each information is encrypted before the label is sent, so that the data captured by the attacker is ciphertext and the location of the label can't be directly cracked. In addition, during the encryption process of the protocol information, random numbers are introduced, which are generated randomly by the random number generator and can't be predicted, so that the messages values of each round are different, giving the attacker a feeling that the label is changing at any time, which can ensure the privacy of the label position.

Backward-secure: In order to prevent an attacker from backtracking the privacy information used before from the current intercepted message, it is necessary to ensure that there is no correlation between the two rounds of messages. The proposed protocol solves this problem by adding random numbers. With the addition of random numbers, the values of the messages before and after are different and can't be predicted. In this way, the attacker can't reverse deduce the previous private data information from the current message.

Forward-secure: To make it impossible for an attacker to predict future session communication message values, it is necessary to ensure that the attacker can't know the values of some parameters used in the next round of message encryption. The proposed protocol solves this problem by randomly generating random numbers through a random number generator, which makes it impossible for the attacker to predict or know in advance the random numbers used in future communication, thus ensuring that the attacker can't predict the future message value.

The proposed protocol can be compared with other classical protocols in terms of security, the comparison results are shown in Table 2.

8 Conclusion

This paper proposed an ultra-lightweight authentication protocol for mobile RFID system. We designed an ultra-lightweight encryption algorithm, which is uses the parity check patching operation. This operation cleverly uses the hamming-weight parameters of the encrypted information, and performs odd party check or even party check in different ways according to the different hamming-weights of the two encryption parameters. This protocol can reduce the storage requirements and increase the difficulty of cracking without adding parameters . The proposed protocol is analyzed with GNY logic formal analysis. Finally, the protocol is analyzed from several specific attack types, and the results show that the protocol has high security; The analysis of tag computation shows that the protocol has some advantages in computation.

Table 2: Security Comparison between Different Protocols

Attack Type	Ref [16]	Ref [9]	Ref [7]	Ref [2]	This Protocol
Mutual Authentication	√	√	√	√	√
Impersonation Attack	√	×	√	×	√
Replay attack	√	√	×	√	√
Asynchronous attack	×	√	×	√	√
Exhaustive Attack	√	√	√	√	√
Pursuit Attack	√	√	√	√	√
Backward-secure	√	√	√	√	√
Forward-secure	√	√	√	√	√

Explain: × indicates irresistibility;√ indicates resistance.

References

[1] J. Y. Chun, G. Noh, "Privacy-preserving RFID-based search system," *Electronics*, vol. 10, pp. 1–13, 2021.

[2] A. Dewanje, K. A. Kumar, "A new malware detection model using emerging machine learning algorithms," *International Journal of Electronics and Information Engineering*, vol. 13, no. 1, pp. 24–32, 2021.

[3] Y. P. Duan, "Lightweight RFID group tag generation protocol," *Control Engineering of China*, vol. 27, no. 4, pp. 751–757, 2020.

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

[5] L. Jing, Z. Zhou, W. Ping, "Cryptanalysis of the lmap protocol: a low-cost RFID authentication protocol," in *Proc of the 29th Chinese Control and Decision Conference (CCDC'17)*, pp. 7292– 7297, Chongqing: IEEE Press, 2017.

[6] Z. Li, G. He, D. Xu, *et al.*, "Evaluation of centralized reader anti-collision protocols for mobile RFID system based on maximum independent set: a simulation study," *IEEE Access*, vol. 8, pp. 123381–123397, 2020.

[7] D. W. Liu, J. Ling, "An improved RFID authentication protocol with backward privacy," *Computer Science*, vol. 43, no. 8, pp. 128–130, 2016.

[8] L. Liu, Y. Liu, "On the anonymity of one multiserver authenticated key agreement with offline registration centre," *International Journal of Electronics and Information Engineering*, vol. 13, no. 3, pp. 105–110, 2021.

[9] Q. Ma, X. Li, G. Li, *et al.*, "MRLIHT: mobile RFID-based localization for indoor human tracking," *Sensors*, vol. 20, no. 6, pp. 1–19, 2020.

[10] S. Maheshwari, "Detection of amplitude shift keying signals using current mode scheme," *International Journal of Electronics and Information Engineering*, vol. 11, no. 2, pp. 73–80, 2019.

[11] T. Pan, K. Zuo, T. Wang, "Design of mobile RFID high-efficiency authentication protocol," *Application Research of Computers*, vol. 40, no. 2, pp. 6, 2018.

[12] Y. J. Priyanka, A. K. Turuk, "Rfid authentication protocol for mobile readers satisfying epc-c1-gen2 standard of passive tags," in *Proc of IEEE International Conference on Technologies for Smart-City Energy Security and Power*, pp. 1–5, Bhubaneswar: IEEE Press, 2018.

[13] G. F. Shen, S. M. Gu, D. W. Liu, "Anti-counterfeit complete RFID tag grouping proof generation protocol," *International Journal of Network Security*, vol. 21, no. 6, pp. 889–896, 2019.

[14] J. Su, Z. Sheng, A. X. Liu, "Capture-aware identification of mobile RFID tags with unreliable channels," *IEEE Transactions on Mobile Computing*, vol. 14, no. 8, pp. 1182–1195, 2020.

[15] Z. Sun, S. Li, "Lightweight authentication protocol for location privacy using puf in mobile RFID system," *Journal of Frontiers of Computer Science and Technology*, vol. 13, no. 3, pp. 418–428, 2019.

[16] D. Wang, F. He, Z. Fang, "Improved secure lightweight RFID authentication protocol," in *Proc of the 3rd International Conference on Machine Learning and Machine Intelligence*, pp. 127–132, New York: Association for Computing Machinery, 2020.

[17] J. Q. Wang, Y. F. Zhang, D. W. Liu, "Provable secure for the ultra-lightweight RFID tag ownership transfer protocol in the context of IoT commerce," *International Journal of Network Security*, vol. 22, no. 1, pp. 12–23, 2020.

[18] Y. Yao, J. Su, "An efficient identification algorithm to identify mobile," *Wireless Communications and Mobile Computing*, vol. 1, pp. 1–8, 2021.

[19] L. Zheng, C. Song, N. Cao, *et al.*, "A new mutual authentication protocol in mobile RFID for smart campus," *IEEE Access*, vol. 6, pp. 60996– 61005, 2018.

Biography

Zhen-Hui Li received a master's degree in School of computer science and engineering from SUN YAT-SEN University (China) in 2007. His current research interest fields include information security and artificial intelligence.

A Study of Web Crawler Recognition Algorithms under the Background of Legal Systems

Daiwei Zhang
(Corresponding author: Daiwei Zhang)
Jilin Police College, Changchun, Jilin 130117, China
Email: fwd3647@163.com
(Received Sept. 4, 2022; Revised and Accepted July 28, 2023; First Online Oct. 8, 2023)

Abstract

With the advancement of technology, web crawlers are widely used but also pose a challenge to the legal system as they can be utilized for illegal data acquisition. Therefore, it is necessary to identify web crawlers accurately. This paper first analyzes the working principle of crawlers, elaborates on the related legal system, and explains the current legal dilemma and the necessity of crawler identification. The seagull optimization algorithm-support vector machine (SOA-SVM) algorithm was developed by analyzing web access logs and using the session length and maximum click rate as the features to distinguish between crawlers and humans while optimizing the SVM parameters with the SOA. The results showed that the SOA-SVM algorithm was more effective in recognizing crawlers than the particle swarm optimization (PSO), ant colony optimization (ACO), and other optimization algorithms, and its F1 value reached 93.16%, which showed an increase of 13.98% compared to the SVM algorithm, as well as increases of 8.91%, 7.22%, and 4.93% compared to the PSO-SVM, ACO-SVM, and GWO-SVM algorithms respectively. Compared to other crawler recognition, the SOA-SVM algorithm also performed well. The experimental results validate the reliability of the SOA-SVM algorithm for web crawler recognition and its potential for further real-world applications.

Keywords: Legal System; Recognition Algorithm; Support Vector Machine; Web Crawler

1 Introduction

A web crawler is a tool that can automatically crawl and save data from the Internet [11]. It helps individuals or enterprises to efficiently obtain required data for data analysis and search engine [7]. However, while crawling information efficiently, web crawlers may occupy a large amount of bandwidth [13] and even cause server crashes.

The behavior of some enterprises to obtain non-public data through abnormal web crawlers seriously infringes on others' information rights and property rights. Therefore, in order to strengthen the protection of data and reduce the pressure on the Internet due to the access of crawlers, identifying web crawlers is particularly important. With the development of technology, research on crawlers continue to deepen.

In order to improve the effectiveness of the search engine, Goel *et al.* [5] proposed a web crawler that can crawl in specific categories to remove irrelevant uniform resource locators (URLs) and implement URL normalization for removing duplicate URLs within a specific category.

Albdour *et al.* [1] designed a dedicated Internet of things crawler to capture malicious nodes and then detected the malicious nodes based on the data stream collected by the crawler using a machine learning approach. The comparison found that a test accuracy of 98.3% was achieved when Extra tree was used. Based on the web crawler technique, Liao *et al.* [9] crawled URLs containing traffic accident data from specified websites of the operating tunnels and processed and analyzed them to manage and control traffic risks.

Ro *et al.* [12] proposed a detection method for identifying distributed web crawlers by classifying frequently used Internet protocol (IP) addresses that request web pages in the long-tail region of the power distribution graph as crawler nodes. Through experiments on National Aeronautics and Space Administration (NASA) web traffic data, the method was found to have a false alarm rate of 0.0275%.

This paper focuses on briefly analyzing the legal system related to web crawlers and designing a support vector machine (SVM)-based recognition algorithm. The research in this paper provides a new algorithm for recognizing web crawlers, which can help websites intercept malicious crawlers, reduce website resource consumption, and improve data security.

2 Web Crawlers in the Context of the Legal System

2.1 Working Principles of Web Crawlers

Web crawlers automatically crawl data on websites, applications, etc. through programs written in advance [8]. However, under the influence of social development, the use of web crawlers has deviated from their original intention and become a tool for some enterprises to illegally obtain data. The working principles of web crawlers are shown below.

1) The URL of the initial page is used as the seed URL and put into the queue;
2) URLs are extracted from the initial resource pool to access web resources;
3) The crawled URLs are put into the pool of crawled resources;
4) New URLs are extracted from the accessed URLs and put into the pool of resources to be crawled;
5) The above process is repeated over and over until all URLs have been crawled.

The following are the main strategies that a crawler can choose from when crawling.

Breadth-first search [4]: crawl the URLs at the current level before moving to the next level;

Depth-first search [10]: crawl the URL and all its layers from the start page before crawling the next URL;

Best-first search: crawl the optimal URL by calculating the URL similarity.

2.2 Legal Regulation of Web Crawlers

When a web crawler violates the reasonable will of the crawled party and contravenes relevant laws and regulations, its crawling behavior is deemed inappropriate. Due to the problems caused by web crawlers, there has been increasing legal scrutiny on their actions. From a legal perspective, crawlers engage in illegal behavior when they exhibit the following actions.

Illegal intrusion: It is determined based on whether the crawled party agrees to access, with reference to the robots protocol. According to the Criminal Law, this type of computer information system intrusion constitutes a crime.

Illegal access: Non-normal crawlers illegally access and obtain data.

Damage or control: Abnormal web crawlers, when accessing a website in large numbers within a short period of time, can disrupt the normal operation of the network system. Moreover, crawler programs perform operations such as adding, deleting, and modifying on the system.

However, the current legislation governing network crawlers still has many deficiencies as it fails to effectively regulate abnormal crawling behavior due to the lack of provisions for corresponding liabilities and lacks a uniform standard for determining the reasonableness of crawler agreements. Moreover, there are ambiguities in defining data and personal information. Therefore, given these inadequacies in the law, identifying web crawlers is of great importance in enhancing data protection.

3 Web Crawler Recognition Algorithm

3.1 Analysis of Web Crawler Features

The access log of a website contains information such as the time of access, URL, etc., so features can be extracted from the access log to distinguish the access behavior of humans and crawlers. Take a particular access log as an example:

$$127.0.0.1 - -[03/Feb/2015:23:14:24+0800] \text{"}GET/HTTP/1.1\text{"}2002,$$

where "127.0.0.1" is the IP of the remote host, "[03/Feb/2015:23:14:24 +0800]" is the time of the access, with "+0800" representing the time zone, "GET / HTTP/1.1" specifies which request was received by the server, "200" is a status code reflecting whether the request was successful or not (with 200 often appearing to indicate a successful response), and "2" at the end denotes the total number of bytes sent to the client.

On the basis of access logs, the features selected in this paper are as follows.

Session length: In terms of time spent browsing the web, the crawler is significantly higher because the crawler runs and crawls the content of the web page without interruption. It refers to the length of time from the beginning to the end of the session.

Maximum click rate: When accessing web pages, crawlers have a higher frequency than human beings. Crawlers will crawl the next web page immediately after crawling a web page resource, while human beings will read for a certain period of time if they are interested in the content of a certain web page after opening it or go to the next page or leave if they are not interested in the content of the web page. Therefore, human beings and crawlers can be distinguished through the click rate, The calculation formula of the click rate is:

$$r_i = \frac{\triangle m_i}{\triangle t_i},$$

where i is the time window, $\triangle t_i$ is the length of the time window, and $\triangle m_i$ is the number of visits in the time window. The maximum click rate is:

$$R = max(r_i).$$

Whether access the robot.txt file: While a human accesses a web page through a browser, a crawler needs to access the robot.txt file first.

Whether the request not includes all requests: Web pages usually contain pictures, cascading style sheets, and other files. These files will be downloaded by the browser and will not be skipped when people are browsing the web page, while crawlers will not make additional requests for these files in order to improve the efficiency of crawling web pages.

Whether the request not carries cookies: When people visit web pages, they usually carry cookies from the previous page. However, crawlers do not carry any cookies when requesting web pages.

In order to facilitate the subsequent identification, these feature data need to be processed. Firstly, for the session length and maximum click rate, they are mapped to the interval of [0,1]. For the feature data without an upper limit, an empirical value is taken as the maximum value, and the formula is:

$$f(x) = \frac{\min(x, \max) - \min}{\max - \min}$$

where x is the empirical maximum. The empirical maximum of the session length is 12 h, and the empirical maximum of the maximum click rate is 50 clicks/s. For Features (3), (4) and (5), "yes" is marked as 1 and "no" is marked as 0.

3.2 Crawler Recognition Algorithm Based on Support Vector Machine

The SVM algorithm is chosen for crawler recognition in this paper. Crawler recognition can be regarded as a two-classification problem, distinguishing between crawlers and humans. The SVM algorithm achieves classification by finding the optimal classification surface, which has high accuracy and efficiency [14].

For the sample set: (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n), $y_i \in \{1, -1\}$, the goal of SVM classification is to find a classification hyperplane:

$$w^T x + b = 0,$$

where w is the coefficient and b is the intercept. To maximize the classification interval, the optimal hyperplane and its constraint can be obtained:

$$\min \frac{1}{2}||w||^2,$$
$$y_i(w^T x_i + b) \geq 1.$$

It is solved according to the Lagrange function:

$$L(w, b, a) = \frac{1}{2}||w||^2 - \sum_{i=1}^{N} a_i[y_i(w^T x_i + b) - 1],$$

where a_i is the Lagrange multiplier. When dealing with the linearly indivisible case, the SVM algorithm adds penalty factor C and kernel function K, then the dual of the original problem can be written as:

$$\begin{aligned} \max L(a) &= \sum_{i=1}^{N} a_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j y_i y_j K(x_i, x_j), \\ &\text{s.t.} \sum_{i=1}^{N} a_i y_i = 0 \text{ and } 0 \leq a_i \leq C, \end{aligned}$$

where $K(x_i, x_j)$ is the kernel function. The radial basis function (RBF) can transform the status data into high dimensions; therefore, in this paper, the RBF kernel function is used in web crawler recognition. Its expression is:

$$K(x_i, x_j) = -\exp(-g||x_i - x_j||^2),$$

where g is the kernel function parameter. In the SVM algorithm, the selection of parameters C and g will have an impact on the recognition results; therefore, this paper uses the seagull optimization algorithm (SOA) [2] to optimize these two parameters. In the SOA, the update of the seagull position, i.e., the parameter optimization process, is as follows.

Migration.

Seagulls need to avoid collision when moving from one position to another. Searching for the position where an individual will not collide with other individuals can be written as:

$$\begin{aligned} C_s &= A \times P_s(t), \\ A &= 2 - [t(\frac{2}{\max_T})], \end{aligned}$$

where $P_s(t)$ is the current position of the seagull, t is the number of iterations, and A is the motion behavior of the individual seagull in the given search space.

After avoiding collisions, individuals move toward their best neighbors:

$$\begin{aligned} M_s &= B \times [P_{bs}(t) - P_s(t)], \\ B &= 2 \times A^2 \times rand, \end{aligned}$$

where $P_{bs}(t)$ is the position of the best seagull and rand is a random number in [0,1]. After obtaining the direction of the best seagull, the distance between the individual seagull and the best seagull can be written as:

$$D_s = |C_s + M_s|.$$

Attack.

During the attack phase, seagulls use spiral motion to approach their prey. The process in the xyz plane can be described as:

$$\begin{aligned} x &= r \times cos\ k, \\ y &= r \times sin\ k, \\ z &= r \times k, \\ r &= u \times e^{kv}, \end{aligned}$$

where u and v are constants for the spiral shape, r is the radius of each turn of the spiral, and k is the random number ($0 \leq k \leq 2\pi$). The final seagull attack position, i.e., the new position after iteration, can be written as:

$$P_s(t) = (D_s \times x \times y \times z) + P_{bs}(t).$$

The process of the SOA-SVM algorithm for web crawler recognition is as follows. It is assumed that the position of each individual seagull is (C, g). The seagull position is updated according to the above formula until it reaches the maximum number of iterations to get the global optimal seagull position (C_{op}, g_{op}), which is input into the SVM algorithm to differentiate the test samples and determine whether they are crawlers or humans.

4 Results and Analysis

4.1 Experimental Setup

The experimental environment was the Windows 10 operating system with an Intel ®CoreTM i5-8300H CPU@2.3GHz. Access logs collected from the campus network outlet of Jilin Police College from March 15, 2023 to March 21, 2023 were analyzed. 496 data were marked as accessed by web crawlers and 572 data were accessed by normal users, totaling 1,925,325 records with a size of 12.94 GB. They were divided into a training set and a test set in a ratio of 7:3.

4.2 Evaluation Indicators

For the recognition effectiveness of the algorithm, Table 1 is used as a basis for evaluation.

Table 1: Crawler identification results

	Identify as a Crawler	Identify as a Human
Actually a crawler	TP	FN
Actually a human	FP	TN

According to Table 1, the evaluation indicators used are as follows:

$$\begin{aligned} Precision &= TP/(TP+FP); \\ Recall &= TP/(TP+FN); \\ F_1 &= \frac{2 \times Precision \times Recall}{Precision + Recall} \end{aligned}$$

4.3 Analysis of Results

Firstly, the SOA-SVM algorithm was compared with several other methods in terms of SVM parameter optimization:

SVM;

PSO-SVM: Optimizing SVM parameters using the particle swarm optimization (pso) algorithm [15];

ACO-SVM: Optimizing SVM parameters using the ant colony optimization (ACO) algorithm (ACO) [6];

GWO-SVM: Optimizing SVM parameters using the gray wolf optimization (GWO) algorithm [16].

The performance of different methods in crawler recognition was compared, as illustrated in Figure 1.

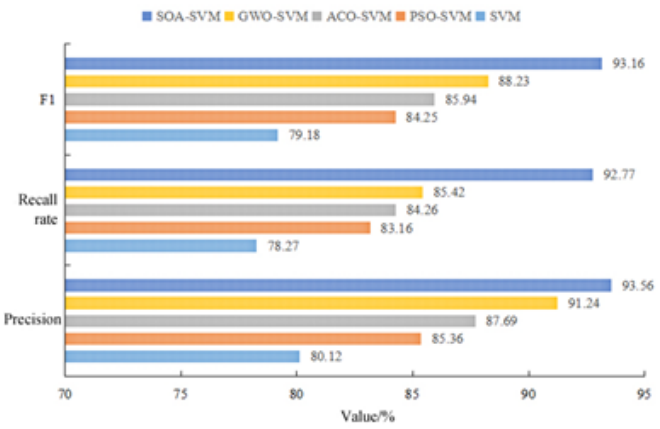


Figure 1: Performance of different methods in crawler recognition

From Figure 1, firstly, the SVM algorithm performed poorly in web crawler recognition without parameter optimization, with a precision of 80.12%, a recall rate of 78.27%, and an F1 score of 79.18%. After parameter optimization, various improved methods for SVM showed significant improvements in web crawler recognition. The PSO-SVM algorithm achieved precision and recall rates both above 80% and an F1 score of 84.25%, which was a 5.07% improvement compared to the SVM algorithm; the ACO-SVM algorithm had an F1 score of 85.94%, which was a 6.76% improvement compared to the SVM algorithm; the GWO-SVM algorithm had an F1 score of 88.23%, which was a remarkable improvement of 9.05% compared to the SVM algorithm. The SOA-SVM algorithm designed in this article achieved a precision of

93.56% in crawler recognition, which was a 13.44% increase compared to the SVM algorithm. The recall rate was 92.77%, which was a 14.5% increase compared to the SVM algorithm, and the F1 value reached 93.16%, which was a 13.98% increase compared to the SVM algorithm and respectively increased by 8.91%, 7.22%, and 4.93% compared with the PSO-SVM, ACO-SVM, and GWO-SVM algorithms. These results proved the advantages of the SOA for optimizing SVM parameters and demonstrated the reliability of the SOA-SVM algorithm in crawler recognition that can accurately distinguish between crawlers and humans. To further demonstrate the reliability of the algorithm designed in this paper for crawler identification, it was compared with the following methods.

Method 1: Determine whether it is a crawler based on the number of image visits. If the percentage of image requests in all requests is less than 10%, it will be considered as a crawler; otherwise, it will be considered as a human.

Method 2: Determine whether a request is from a crawler based on its access frequency; if more than 2 URLs per second are accessed, it is recorded as a crawler; otherwise, it will be considered as a human.

Method 3: Determine whether it is a crawler based on accessing the robot.txt file. If the robot.txt file is accessed, it will be recorded as a crawler; otherwise, it will be recorded as a human.

Method 4: The first-order discrete-time Markov chain model [3].

The results of the comparison are presented in Figure 2.

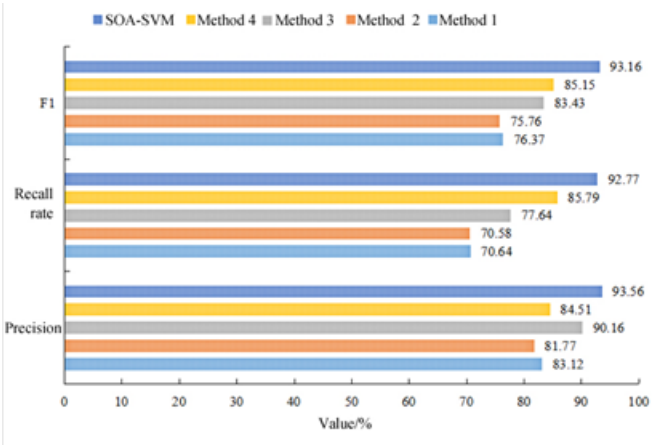


Figure 2: Comparison with other crawler recognition methods

According to Figure 2, firstly, methods 1, 2, and 3 relied solely on one crawler feature to distinguish between crawlers and humans. The precision, recall rate, and F1 values of both method 1 and method 3 were below 90%, indicating that there were some shortcomings in judging whether it is a crawler based solely on image access or access frequency. Compared with these two methods, method 3 which judges through accessing the robot.txt file had relatively better results with a precision of 90.16%, but a lower recall rate of 77.64%. The final F1 value was 83.43%. For the robot.txt file, normal and reasonable crawlers will access it while some illegal malicious crawlers will bypass it. Therefore, relying solely on this item to identify crawlers also has limitations. Then, comparing method 4 with the approach in this paper, it can be observed that method 4 had a precision of 84.51%, a recall rate of 85.79%, and an F1 score of 85.15%, all lower than the approach presented in this paper. This result indicated that the approach proposed in this paper achieved effective identification of web crawlers by utilizing five features related to web crawling.

5 Discussion

The development of crawlers and the resulting social impact has led to the emergence of new areas of legislation and the introduction of new requirements for improving and supplementing existing laws. Currently, from a legal perspective, there are several laws that can be referred to when considering network crawlers.

1) Civil Code of the People's Republic of China
If the information that the crawler crawls and then presents has characteristics such as originality, and the crawled information is a substitute for the services provided by the person being crawled, then this behavior is illegal.

2) Law of the People's Republic of China for Countering Unfair Competition
According to the "Internet Special Provision" regarding competition behavior in the internet industry, it is stipulated that operators should not hinder or damage others' normal operations through technical means, including inserting inappropriate links into others' products to redirect customers and using illegal methods to prevent users from receiving reasonable services provided by others. However, this provision does not make explicit regulations on web crawling behavior.

According to the general provisions of the Law of the People's Republic of China for Countering Unfair Competition, it can be determined whether the behavior of crawlers violates the principles of good faith and commercial ethics. If the crawler does not comply with a reasonable crawler agreement, it can be considered a violation of the Self-Discipline Convention.

According to the relevant legal regulations, the illegal use of web crawlers may constitute the following crimes.

1) The crime of infringing citizens' personal information

From February to April 2018, Ma crawled user information from applications and websites through a crawler program and sold it to others.

2) The crime of unlawful intrusion into a computer information system
The defendant used crawler software to crawl the bulletin of the vehicle administration office for license plate release information.

3) The crime of illegally obtaining computer information system data
The defendant utilized Taobao store vulnerabilities to illegally crawl Taobao user cookies and obtain user transaction order data.

4) The crime of damaging computer information system
The defendant caused a system to malfunction through a web crawler program, resulting in the deletion of a large amount of data stored in the system.

However, there are also some cases that reflect the inadequacy of current laws. In 2016-2017, Shanghai Shengpin Network Technology Company was found guilty of illegally obtaining computer information system data by using technical means to capture ByteDance's data. This is considered the first case of criminalizing "crawlers". In 2017, Yuanguang Company used crawler technology to unlawfully obtain data from Goome Company, constituting unfair competition. These two cases have similar circumstances; however, the former is a typical example of a criminal case while the latter is an instance of unfair competition disputes. This shows that there are significant differences in how courts determine cases where crawlers violate the law. Moreover, in many judicial decisions, whether or not the crime was committed for the purpose of making profit was used as a basis for sentencing, and the crawling of insects was not taken into account as a separate criminal circumstance. In addition, in the actual judgment, the subjective judgment of whether the crawler behavior is illegal is also more difficult. Therefore, in the future, it is necessary to strengthen the legal means to regulate the crawler technology to ensure the free flow of data in accordance with the law. Moreover, in many judicial decisions, whether or not the crime was committed for the purpose of making a profit has been used as a basis for sentencing, and crawling has not been taken into account as a separate criminal circumstance. Additionally, in actual judgments, it is also more difficult to subjectively determine whether crawler behavior is illegal. Therefore, in the future, it is necessary to strengthen legal measures to regulate crawler technology and ensure the lawful free flow of data.

6 Conclusion

In the context of legal regulations, this article proposes a SOA-SVM method for identifying web crawlers, using session duration, maximum click rate, etc. as features. The method was tested on actual access logs and achieved an F1 score of 93.16% in crawler identification, outperforming PSO and ACO algorithms in SVM parameter optimization. Compared to using a single feature alone, the SOA-SVM algorithm showed better performance in crawler identification and can be applied in practical scenarios.

References

[1] L. Albdour, S. Manaseer, A. Sharieh, "IoT crawler with behavior analyzer at fog layer for detecting malicious nodes," *International Journal of Communication Networks and Information Security*, vol. 12, no. 1, pp. 83-94, 2020.

[2] G. Dhiman, V. Kumar, "Seagull optimization algorithm: Theory and its applications for large scale industrial engineering problems," *Knowledge-Based Systems*, vol. 165, pp. 169-196, 2019.

[3] D. Doran, S. S. Gokhale, "An integrated method for real time and offline web robot detection," *Expert Systems*, vol. 33, no. 6, pp. 592-606, 2016.

[4] W. A. Gab-Allah, B. B. S. Tawfik, H. M. Nassar, "Performance analysis of an ontology based crawler operating in a distributed environment," *International Journal of Science Research in Science and Technology*, vol. 2, no. 3, pp. 334-339, 2016.

[5] K. Goel, J. S. Prasad, S. Hilal, "Removing duplicate URLs based on URL normalization and query parameter," *International Journal of Engineering & Technology*, vol. 7, no. 3.12, pp. 361-365, 2018.

[6] M. Hamdi, "Affirmative ant colony optimization based support vector machine for sentiment classification," *Electronics*, vol. 11, no. 7, pp. 1-11, 2022.

[7] N. L. H. Hien, T. Q. Tien, V. H. Nguyen, "Web crawler: Design and implementation for extracting article-like contents," *Cybernetics and Physics*, vol. 9, no. 3, pp. 144-151, 2020.

[8] M. K. Hossen, Y. Wang, H. A. Tariq, G. Nyame, R. Nuhoho, "Statistical analysis of extracted data from video site by using web crawler," in *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence*, pp. 41–46, 2018.

[9] Z. Liao, H. Ding, Y. Xia, F. Ma, "Web crawler based study on traffic accident data acquisition for operating tunnels," *IOP Conference Series: Materials Science and Engineering*, vol. 741, no. 1, pp. 1-7, 2020.

[10] M. A. Mabayoje, O. S. Oni, O. S. Adebayo, "A full-text website search engine powered by lucene and the depth first search algorithm," *International Journal of Computer Network & Information Security*, vol. 5, no. 3, pp. 1-12, 2012.

[11] Y. D. Pramudita, D. R. Anamisa, S. S. Putro, M. A. Rahmawanto, "Extraction system web content sports new based on web crawler multi thread," *Journal of Physics: Conference Series*, vol. 1569, no. 2, pp. 1-6, 2020.

[12] I. Ro, J. S. Han, E. G. Im, "Detection method for distributed web-crawlers: A long-tail threshold model," *Security and Communication Networks*, vol. 2018, no. 4, pp. 1-7, 2018.

[13] N. Singhal, A. Dixit, R. P. Agarwal, A. Sharma, "Regulating frequency of a migrating web crawler based on users interest," *International Journal of Engineering and Technology*, vol. 4, no. 4, pp. 246-253, 2018.

[14] A. A. Sretenović, R. Ž. Jovanović, V. M. Novaković, N. M. Nord, B. D. Živković, "Support vector machine for the prediction of heating energy use," *Thermal Science*, vol. 22, no. Suppl. 4, pp. 1171-1181, 2018.

[15] H. T. Thom, C. M. Yuan, V. Q. Tuan, "A novel perturbed particle swarm optimization-based support vector machine for fault diagnosis in power distribution systems," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 26, no. 1, pp. 518-529, 2018.

[16] T. Zhongda, "Kernel principal component analysis-based least squares support vector machine optimized by improved grey wolf optimization algorithm and application in dynamic liquid level forecasting of beam pump," *Transactions of the Institute of Measurement and Control*, vol. 42, no. 6, pp. 1135-1150, 2020.

Biography

Daiwei Zhang, born in 1973, graduated from Northeast Normal University in July 1996 with a B.S. degree. He is an associate professor and is now working in Jilin Police College, China. He is interested in Administrative Law.

Spatial Optimal Method of Crowdsourcing Allocation Algorithm Based on Tree Decomposition

Hui Xia[1*], Shufeng Zhang[2*], and Weiji Yang[3]

(Corresponding author: Weiji Yang)

**Hui Xia and Shufeng Zhang contribute equally to the article.*

Shenyang Normal University, Shenyang 110034, China[1]

Suzhou Industrial Park Institute of Service Outsourcing,Suzhou 215123, Jiangsu, China[2]

Zhejiang Chinese Medical University, HangZhou 310000, China[3]

Email: yangweiji@163.com

(Received Dec. 7, 2022; Revised and Accepted Aug. 12, 2023; First Online Oct. 6, 2023)

Abstract

With the popularity of mobile devices equipped with high-fidelity sensors and the rapid decline of wireless network fees, space crowdsourcing as a problem-solving framework is used to solve the problem of assigning location-related tasks (such as road condition reports, food distribution) to workers (people equipped with intelligent devices and willing to complete tasks). The key to studying the optimal task allocation in space crowdsourcing is to design a task allocation strategy that assigns each task to the most suitable worker to maximize the total number of tasks completed, and all workers can return to the starting point before the expected last working hours after completing the assigned tasks. Finding the optimal global allocation is a difficult problem because it is not equal to the simple accumulation of the optimal allocation of individual workers. Some workers have task dependence, so we use tree decomposition technology to divide workers into independent sets and propose a heuristic depth-first search algorithm. This algorithm can quickly update the heuristic function boundaries to efficiently prune the allocation scheme that can not be the optimal solution as soon as possible. The experimental results show that the proposed method is very effective and can be very good. Solve the problem of optimal task allocation.

Keywords: Optimal Solution Algorithm; Spatial Crowdsourcing; Task Allocation; Task Dependency; Tree Decomposition

1 Introduction

In this paper, the optimal task [6] allocation scheme in spatial crowdsourcing based on the above scenario is studied. Specifically, given the location of each worker and the worker's expected latest working time, given the location and expiration time of each task, is used to find the optimal task allocation scheme that maximizes the number of global task assignments.

Compared with the existing jobs, the main difficulty of this problem is that once the time cost and the expiration time of the task need to be considered, the local optimization may not necessarily produce the global optimal solution. The second challenge of this paper is that the reachable task scope of the worker highly depends on the starting position of the worker and the expected working time of the worker.Define the scope of the work area or the maximum number of acceptable tasks to eliminate unreachable tasks [1, 4, 5, 7].

In order to solve this problem, this paper proposes an exact solution algorithm to find the optimal allocation scheme to maximize the global task allocation number. The main idea of this paper is as follows: according to the relationship of task dependence (for example, two workers can complete a task, then there is a task dependency between them), the workers are divided into independent worker sets by tree decomposition [5]; Then, the worker set is indexed as a node into a search tree structure. Finally, a heuristic depth-first traversal algorithm is used to search the optimal solution.

At the same time, the proposed algorithm combines a variety of optimization strategies, which can shrink the upper and lower bounds of the search process quickly and prune efficiently. Compared with the iteration-based method, the proposed algorithm can obtain the optimal solution at the end of the search [13]. The main contributions of this paper are three points.

1) For the first time, the task allocation problem in space crowdsourcing with the latest working time constraint is studied. In this model, workers have different working time constraints and need to con-

sider the expiration time limit of tasks. The number of tasks acceptable to workers is no longer fixed.

2) This paper presents an accurate solution algorithm, which can effectively find the optimal task allocation scheme. The algorithm uses tree decomposition technology to segment workers without task dependence, and uses effective pruning algorithm to improve search efficiency.

3) The effect of key parameters on the efficiency of the scheme is analyzed by experiments.

2 Related Work

In the present work [3], it has been proved that the global optimal task allocation scheme in space crowdsourcing is a difficult problem for NP. Therefore, the simplest method is to use greedy algorithm to find the maximum set of effective tasks for workers in turn, and then accumulate the number of assigned tasks. The main problem of this method is that for tasks that can be accomplished by multiple workers, simple random assignment on technology to divide unrelated workers into different worker sets and index these worker sets using search tree structure. Finally, this paper uses heuristic depth-first algorithm to search the search tree constructed in Step 2.

2.1 Computing Effective Task Sets

2.1.1 Finding Reachable Tasks

Restricted by the latest working hours and the expiration time of tasks, each worker can only accomplish a small number of tasks. Therefore, first of all, we should find out the set of tasks that each worker can reach without violating the constraints. The set of tasks that workers can reach should be recorded as meeting the following two conditions.

1) $\forall s \in RS_w, c(w.1, s.1) \leq s.e$;

2) $c(w.l, s.l, w.l) = c(w.l, s.l) + c(s.l, w.l) \leq w.t$.

$C(w.l, s.l, w.l)$ is the time for workers to return to $w.l$ from $w.l$ via $s.l$. The above two conditions can ensure that a worker can reach the position of the task from his starting point before the expiration of the task, and at the same time leave enough time to return to his starting point before the worker's latest working time. From the point of view of calculation, reachable task is a circle centered on the worker's starting point, with the worker's starting point as the center and the worker's starting point as the starting point.The maximum driving distance ($w.t/2$ multiplied by a fixed speed) is a circular radius.

2.1.2 Searching for Maximum Effective Task Sets

When searching for the optimal solution, in order to speed up the search efficiency, this paper hopes that a worker can be assigned multiple simultaneous reachable tasks at one time, instead of one task at a time, and then judge whether the newly assigned tasks and the assigned tasks are reachable at the same time. Therefore, it is necessary to estimate the maximum effective task set of each worker. Given the reachable tasks of each $worker.Sets$, it can be proved that: find out the maximum effective task set of each worker (MVTS) is a NP-hard problem, and the proof process is similar to that in document [5]. However, because each worker's achievable task set is usually small, this means that this problem can be solved by efficient algorithms. Moreover, the calculation of MVTS for each worker is completely independent, so it can be calculated in parallel.

Next, this paper introduces the state transition equation of the dynamic programming algorithm to solve the maximal effective task set. By gradually increasing the size of the reachable task set, the algorithm constantly extends the worker's reachable task set, and finds out all the MVTS under the set in each iteration. Given a worker w and a set of tasks $Q \subset RS_w$, this paper defines opt (Q, s) as a task after passing through the Q set, The maximum number of tasks that can be completed at the location where the task $s.l$ is located. R is the task scheduling sequence in the Q set. Representation of s_j in sequence R by s_i. The previous task, and R' denotes the corresponding sequence of tasks for opt $(Q - \{s_j\}, s_i)$. $Opt(Q, s_j)$ can be calculated from Formula (1).

$$opt(Q, s_j) = \begin{cases} 1 & \text{if } |Q| = 1 \\ \max_{s_j \in Q, s_i \neq s_j} & \text{otherwise} \end{cases} \quad (1)$$

Wherein

$$\sigma_{i,j} = \begin{cases} 1 & \text{if } t(s_j, l) \leq s_j.e, t(s_j, l) + c(s_j.l, w.l) \leq w.t \\ 0 & \text{otherwise} \end{cases}$$

$\sigma_{i,j} = 1$ indicates that the task s_j can still be completed at the end of adding the task to the sequence R', and that the worker s_j can return to the starting point before the latest expected time.

When Q only includes one task s_i, the problem is very simple. When $opt(s_i, s_i) = 1$ and $|Q| > 1$, it is necessary to search Q to check all possibilities of the effective task set s_i and find them so as to maximize $qpt(Q, s_j)$. The time complexity of the MVTS set Q_w in Formula (1) is $O(n^3 \cdot 2^n)$, while the space complexity is $O(n \cdot 2^n)$.

2.2 Segmentation of Worker Set

The main challenge in finding the optimal solution lies in the large search space. When enumerating all possible effective task sets of all workers, the time complexity increase exponentially with theincrease of the number of workers.

Definition 1. *(Task Dependence). Given two workers, WJ and their respective achievable task sets RS_w, if RS_w mutual independence; otherwise, there is task dependence between them.*

For example, in Figure 1, workers are only dependent on and exist task dependence, but not related to other workers. Obviously, if there is no task dependence among all workers, the optimal task allocation scheme only needs to simply add up the optimal scheme of each worker. Therefore, in order to reduce the computational cost of finding the optimal solution, it is necessary to use task dependence as much as possible to divide workers into unrelated sets.In each independent set, the optimal allocation scheme in the set is found.

2.2.1 Graph Decomposition

Definition 2. *(Task Dependency). Given two workers w_i, w_j and their reachable task sets RS_{w_i}, RS_{w_j}, if $RS_{w_i} \cap RS_{w_j} = \phi$, they are independent of each other; Otherwise, there is task dependency between them.*

Figure 1(a) shows the worker dependency diagram in the example.

3 Problem Statement and Preliminaries

Definition 3. *(Space task) A spatial task can be defined by a binary group $s =< s.e, s.l >$. Where, $s.l$ is the position of the task, $s.e$ is the expiration time of the task, $s.l$ is a point (x, y) in a two-dimensional plane space. In spatial crowdsourcing, the task can be completed only when the worker actually reaches the specified position $s.l$ of task s. At the same time, considering the expiration time of the task, The worker can complete the task only when he arrives at the $s.l$ location before the task s expiration time $s.e$. Under the non redundant task allocation mode considered in this paper, the server will only assign one task to a single worker. Like the previous work [2, 17, 21], this paper assumes that the worker will immediately go to the next task after completing one task, and the time required to complete the task itself is negligible.*

Definition 4. *(worker) Workers generally refer to people who carry mobile equipment and voluntarily complete space tasks Workers can be represented by the binary group $s = (w.l, w.t)$, where $w.l$ is the worker's starting position, $w.t$ is the expected latest working time of the worker, and the worker needs to return to the starting position no later than the latest working time.*

The worker's working mode is divided into online mode and offline mode. When the worker is in online mode, he can accept space tasks. Once the worker is in online mode, he will send a task request to the server. The request contains the worker's location $w.l$. The latest expected working time of the worker $w.t$. The server will consider all workers and tasks acquired in a very short time interval at the same time, and then make the global optimal task allocation. Finally, the task sequence assigned to each worker is returned.

Definition 5. *(task sequence) Given any worker w and the task set s_w assigned to w, the task sequence of s_w is expressed as $R(s_w)$, representing the time sequence in which workers access each task. The time $t_{w,R}(s_i.l)$ when workers arrive at each task can be defined by Formula (??):*

$$t_{w,R}(s_i.l) = \begin{cases} t(s_{i-1}.l) + c(s_{i-1}.l, s_i.l) & if\ i \neq 1 \\ c(w.l, s_1.l) & if\ l = 1 \end{cases} \quad (2)$$

In Formula (2), $C(a, b)$ represents the travel time from task a to task b. Under the condition that w and R are not ambiguous, $t(s.l)$ is used to represent the time when worker w arrives at task s_i's location $s.l$, and $t(w.l)$ is used to represent the time when worker w returns to its original location. The time when a worker returns to the starting point after completing all tasks in the task set S_w is defined as:

$$t(w.l) = t(s_{|s|}.l) + c(s_{|s|}.l, w.l). \quad (3)$$

Because the driving speed is not within the scope of the main factors considered in this paper, for simplicity, this paper assumes that all workers have the same driving speed. Therefore, the travel time cost of workers can be defined by Euclidean distance at two locations. However, the method proposed in this paper does not rely on this assumption, and can be used when workers have different speeds.

Definition 6. *(valid task set VTS) When and only when the following conditions are true, the task set S_w is called the effective task set of worker w:*

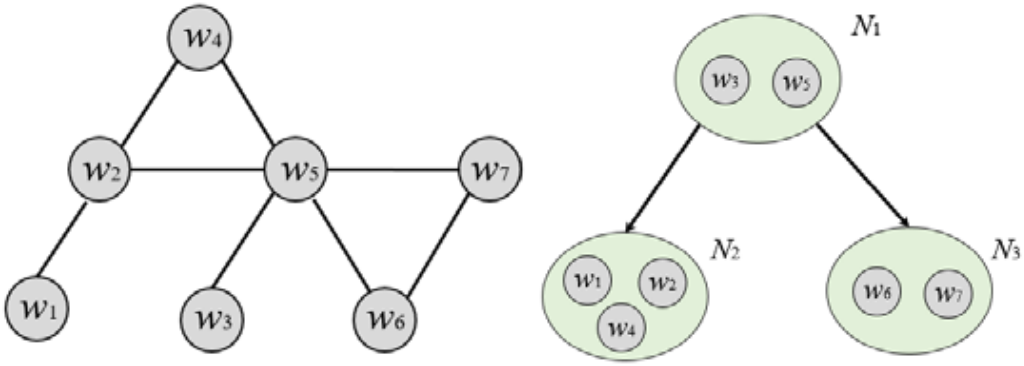
1) *All tasks in S_w can be completed before their expiration time, that is, $\forall S_i \in S_w$, $t(S_i.l) \leq S_i$;*

2) *Worker w can return to the starting point no later than the latest working time after completing all tasks in S_w, that is, $t(w.l) \leq w.t$.*

Definition 7. *(MVTS) If any superset of the effective task set S_w is not the effective task set, then S_w is called the maximum effective task set.*

Definition 8. *(Space Task Assignment) Given worker set W and task set S, spatial task assignment A consists of a series of workers, VTS tuples, such as $A = \{< w_1, VTS(w_1) >, < w_2, VTS(w_2) >, \cdots, < w_{|S|}, VTS(w_{|S|}) >\}$. Let $A.S$ represent the task set assigned to all workers, that is, $A.S = \cup_{w \in W} S_w$. The problems raised in this paper are summarized as follows:*

Problem definition. *Given a worker set W, a task set S, and a space crowdsourcing task allocation problem with the latest working time constraints, the goal is to find a global optimal task allocation scheme A_{Opt}, so that $\forall A_i \in A, |A_i.S| \leq |A_{Opt}.S|$, where A represents all space task allocation schemes.*



(a) (b)

Figure 1: Worker dependency graph partition (a)Worker Dependency graph, (b) Constructed search tree.

4 Algorithm Analysis

4.1 Introductions of Algorithm

The purpose of graph decomposition is to divide irrelevant workers into different sets, and at the same time ensure that the used partition set (hereinafter referred to as cut set) can divide the connected graph as evenly as possible (see Algorithm 1 for decomposition algorithm). In each step of decomposition, take a point and connect to it as a cut set (Row 9), and try to use the cut set to cut the connected graph (Rows 10 and 11), and record the size of the cut set (the cut set contains multiple workers) and the sum of the number of workers which is in the maximum connected subgraph after the graph is cut (Row 12). In each step, select the smallest cut set (Lines 11∼14) to segment the original graph, and recursively call the algorithm on the connected subgraph after segmentation until the number of workers in the connected subgraph is less than the threshold (Lines 3∼5).

Because the optimal graph decomposition problem itself is NP hard, this scheme does not solve the optimal cut set in each iteration of tree decomposition, but attempts to take each node in the dependency graph and its surrounding nodes as the cut set in turn, and records the best cut set that has been found (Lines 7 to 15) As shown in Figure 2, when w_3 and w_5 connected with w_3 are used as cut sets, other workers are divided into two relatively average groups, and the resulting point cut set sequence C is $[\{w_3, w_5\}, \{w_1, w_2, w_4\}, \{w_6, w_7\}]$. At this time, the cut set size is 2, w_1, w_2, and w_4, and the maximum size of the connected subgraph is 3, and the sum of the two is 5. It can be found that the sum of the cut set size and the maximum size of the connected subgraph is not smaller than 5 in other segmentation methods. Therefore, Algorithm 1 will preferentially use w_3 and w_5 as cut sets.

Given the worker set W and the task set S, we first calculate the maximum task subsequence Q_{W_i} (the second and third lines) for each worker W_i, and S_i is the reachable task set, and then establish the corresponding task dependency graph G (the fourth line) For each con-

Algorithm 1 Tree decomposition algorithm of graph

1: Input: worker dependency graph G, sequence C of point cut set (worker set), sequence I generated by point cut set (initialized to 0, representing the sequence number of the first point cut set), and global variable index, representing the current point cut set sequence number cur;
2: Output: point cut set sequence C and sequence I generated by each point cut set.
3: TreeDecomposition(G,C,I,index,cur);
4: Best_h=infinity, Best_S=empty set;
5: **if** $|G|$ <threshold **then**
6: $C[cur] \leftarrow$ nodes in G;
7: **end if**
8: **for** each node $w_i \in G$ **do**
9: $Set(V_i) \leftarrow$ node connected to w_i;
10: $Set(V_i')$‘$leftarrow$ node not connected to w_i;
11: $G' \leftarrow$ build graph of $Set(V_i)$;
12: $h \leftarrow |Set(V_i)| + \max|SubGraph(G')|$;
13: **if** $h < Best\_h$ then **then**
14: $Best\_h \leftarrow h$;
15: $Best\_S \leftarrow Set(V_i)$;
16: **end if**
17: **end for**
18: $C[cur] \leftarrow Best\_S$;
19: $G'' \leftarrow$ build graph for node not in Best_S;
20: **for** each sub graph $G_s \in G''$ do **do**
21: Sub_index=++index;
22: $I[cur] \leftarrow Sub\_index$;
23: $TreeDecomposition(G_s, C, index, Sub\_index)$;
24: **end for**

nected subgraph $g \in G$ in the task dependency graph, use the tree decomposition algorithm to divide the workers into different worker sets (Row 8), and then build the sequence of worker sets into a search tree structure (Row 9). Finally, use the depth first search algorithm to find the optimal task allocation scheme for the search tree established in the previous step. The initial heuristic function value can be calculated by the greedy algorithm. Because different connected subgraphs of G are not related to each other, the final task allocation scheme only needs to add up the schemes of different connected subgraphs (Line 8) .

Algorithm 2 Search framework

1: Input: worker set W, task set S;
2: Output: Optimal task allocation scheme Opt
3: $Solve(W, S)$
4: **for** each $w_i \in W$ **do**
5: $Q_w = MVTS(w_i, S_i)$;
6: $G \leftarrow$ build WDG;
7: **end for**
8: **for** each connected component $g \in G$ **do**
9: $X_g \leftarrow$ TreeDecomposition() of g;
10: $N_g \leftarrow$ Build search tree;
11: $Opt \leftarrow Opt + DFSearch(N_g, S, W_N, LB(N_g))$;
12: **end for**
13: Return Opt;

Algorithm 3 Heuristic search algorithm

1: Input: the current node serial number N, the unassigned task set S, the workers set WN that has not been searched in the Nth subtree, and the heuristic function value h used to cut the search space;
2: Output: Optimal task allocation scheme Opt
3: $DFSearch(N, S, W_N, h)$
4: $Opt \leftarrow 0$;
5: $UB(N) \leftarrow$ Calculate upper bound of sub tree rooted at N;
6: **if** $UB(N) < h$ then **then**
7: return 0;
8: **end if**
9: **if** $W_N \neq \phi$ then **then**
10: **for** each worker $w_i \in W_N$ **do**
11: **for** each MVT set $Q \in MVTS(w_i, S)$ **do**
12: $Opt \leftarrow \max\{DFSearch(N, S-Q, W_{N-w}, h - |Q|) + |Q|, Opt\}$
13: $h \leftarrow Opt$;
14: **end for**
15: **end for**
16: **else**
17: **for** each child node N_i of N **do**
18: $Opt+ = DFSearch(N_i, S, W_N, h)$;
19: **end for**
20: **end if**
21: return Opt;

The depth first algorithm is described in detail below 3. The four key parameters of the search process are as follows: (1) the root node of the N^{th} subtree; (2) Unassigned task set S; (3) The worker set W_N that has not been searched in the N^{th} subtree; (4) The heuristic function value h used to prune the search space represents the minimum number of tasks to be allocated without pruning the subtree.

Algorithm 3 will recursively call itself to search for all task allocation schemes, and will exclude the allocation schemes that will not become the optimal solution. Therefore, when the algorithm exits, the algorithm can obtain the allocation scheme that maximizes the number of global tasks allocated.

4.2 Upper Bound of Estimation

The upper bound of the number of tasks that can be completed by node N is recorded as UB(N), which represents the maximum number of tasks that can be completed by all workers in the subtree with node N as the root. The basic method for estimating the upper bound is to accumulate the maximum effective task set of all workers in the subtree with node N as the root, and the maximum effective task set with the largest number of workers. The calculation formula is as follows:

$$UB(N) = \sum_{i=1}^{|W|}(|\max R_{w_i}|). \tag{4}$$

W in the formula represents all workers of the current subtree; $\max R_{w_i}$ represents the maximum effective task set of worker w_i, and the set with the largest number of set elements. For example, when the search algorithm starts to search in Figure 3, $UB(N_3) = |\max R_6| + |\max R_7| = 1 + 3 = 4$, and the value of $|\max R|$ can be obtained by looking up Table 2.

For all task allocation schemes A (including the optimal task allocation scheme), the number of tasks that each worker can complete will not exceed $|\max R|$, so the following inequality is true:

$$|A.S| = |\cup_{w \in W} S_w| \leq \sum_{w \in W} |S_w| \leq \sum_{w \in W} |\max R_w| = UB(N).$$

4.3 Heuristic Function

In order to prune the scheme that is not expected to be the optimal solution as soon as possible, the algorithm will calculate the heuristic lower bound h and transfer it to the recursive function as a parameter H indicates the minimum number of tasks to be completed by the subtree with node N as the root Only when it is not less than h, it is possible to produce a more promising solution than the currently searched optimal case Easy to get: When the upper bound (the maximum number of tasks that can be allocated) of a subtree is less than the lower bound, you can safely exclude this scheme.

The following describes how to estimate the heuristic function value of the subtree with node N as the root.

Assuming that node N contains m sub nodes, such as $N_1, N_2, \cdots, N_m$, the depth first search algorithm will be used to search each sub tree in turn to find a better solution than the currently found optimal solution Opt. The heuristic function value is updated by the following formula:

$$h' = h - \sum_{j=1}^{i-1} Opt(N_j) - \sum_{j=i+1}^{m} UB(N_j). \quad (5)$$

In Formula (3), h represents the minimum number of tasks to be allocated to all sub nodes of node $\sum_{j=1}^{i-1} Opt(N_j)$ represents the sum of the maximum number of tasks that can be allocated to the traversed subtree, and $\sum_{j=i+1}^{m} Opt(UB(N_j))$ represents the sum of the estimated upper bounds of the unserved subtree. That is, the number of tasks that the current subtree N_i needs to complete at least is equal to the heuristic function value h. Subtract the subtree $(N_1 - N_{i-1})$ that has been searched before the current subtree to determine the number of tasks that can be completed, and then subtract the estimated upper bound UB(N) of the subtree $(N_{i+1} - N_m)$ that has not been searched after the current subtree.

4.4 Optimization Strategy

Three optimization strategies are introduced here to further reduce the search cost:

1) UB(V) optimization: before calling the depth first search algorithm, start from the leaf node of the search tree, from the bottom to the top, and use the greedy algorithm to find the upper bound of the tasks that can be assigned to the subtree with each node N_i as the root (each worker only takes the maximum effective task set, regardless of the task being assigned), which can limit the heuristic function to a smaller extent;

2) Single worker search optimization: sort the maximum effective task set ($|MVTS|$) of each worker from large to small, and search the set with large base first. Because if the search for the larger effective task set that is more likely to produce the optimal solution is completed, the smaller effective task set is easier to prune;

3) Search sub tree order optimization: all sub nodes of a node are sorted from small to large according to the number of workers contained in the sub tree whose sub node is the root. Because small subtrees can be searched quickly, the heuristic function value h can be updated quickly, which will become more compact and have better pruning effect for large subtrees.

5 Experiments and Results

5.1 Experimental Setup

Because of the lack of benchmark experimental data in the space crowdsourcing field, this paper uses the simulation data set to generate experimental data. The rules are as follows:

1) First, use uniform distribution on a two-dimensional plane of 100×100, and randomly generate 100 point coordinates to represent the position of workers; Then, change the average number of tasks (T/W) of each worker, randomly generate points representing tasks around each worker, and the value range of T/W is [9–12,14] (experimental verification: when conducting experiments on a larger scale, the depth of search usually exceeds 20, and the search time increases exponentially, exceeding the allowable range of the experimental machine configuration);

2) Secondly, given any worker and the tasks around him, the expiration time of the task is defined as follows: starting from the current position of the worker, the greedy algorithm is used to select the tasks closest to the current position of the worker in turn. For the task sequence calculated by the greedy algorithm, the total travel time t is calculated and used as the upper bound of the task expiration time; Then define a range $[e^l, e^u](0 < e^l < e^u < 1)$. The expiration time of tasks is defined as $[e^l.t, e^u.t]$. This paper uses five groups of task expiration ranges [0.2, 0.3], [0.3, 0.4], [0.4, 0.5], [0.5, 0.6] and [0.6, 0.7];

3) Finally, the expected latest working time of workers is defined as follows: the distance t_w between the last task and the worker's starting point is obtained by adding the worker's traveling time t calculated by the greedy algorithm as the upper bound of the expected latest working time of workers; Then define a range $[d^s, d^t](0 < d^s < d^t < 1)$. The expected latest working time of workers is defined as $[d^s.t_w, d^t.t_w]$. This paper uses five groups of this range [0.2, 0.3], [0.3, 0.4], [0.4, 0.5], [0.5, 0.6], and [0.6, 0.7].

Basically, the task expiration time range and the expected latest working time range of workers determine the percentage of tasks that workers can complete For the results of each parameter change, 50 groups of experiments were carried out The reported results are the average results of 50 groups of experiments All experiments were conducted on a machine with Core i5-2400, 3.1G HZ CPU and 8GB RAM.

5.2 Experimental Result

This paper evaluates the performance of worker partitioning stage and the impact of task partitioning results on search, and compares random tree construction algorithm (RTA) with balanced tree construction algorithm (BTA)

proposed in this paper RTA algorithm randomly selects a worker cluster as the current node of the search tree. BTA algorithm always selects a better worker cluster as the current node when constructing the current search node This section compares two dimensions: (1) Search depth: search the maximum number of workers enumerated from the root node to the leaf node using depth first; (2) Search time: CPU time spent searching the optimal allocation scheme using the constructed search tree.

Figure 3 shows the influence of the average number of workers' tasks T/W, the task expiration time coefficient el, and the worker's latest working time coefficient ds on the search depth Figure 4 shows the effect of the above parameters on search time As shown in Figure 2(a), the search depth of the two tree construction algorithms increases with the increase of T/W. However, the BTA algorithm can produce a more balanced tree, which makes it more efficient than RTA, as shown in Figure 4(a).

As shown in Figure 2(b): when the task expiration time coefficient is small, the number of reachable tasks for each worker is also small, and the BTA algorithm and RTA algorithm have no difference in the ability to construct a search tree; However, with the increase of $[e^l, e^u]$, the number of accessible tasks for workers also increases rapidly, and the advantages of BTA algorithm become more obvious As shown in Figure 3(b), although the search time increases exponentially with the increase of task expiration time coefficient, the performance of BTA algorithm is an order of magnitude higher than that of RTA.

Based on the search tree constructed by BTA algorithm, this section compares the performance of three different search algorithms: (1) depth first search (DFS) without any optimization; (2) Use sorting based search to optimize DFS+W (depth first search+worker sort) (the effective task set of a single worker is in the order of large to small and the number of workers in the subtree is in the order of small to large); (3) On the basis of (2), the algorithm DFS+W&U [8, 12, 16, 19] (depth first search+worker sort and upper bound estimate) is added For the final task allocation quantity, this paper compares the task allocation quantity of the algorithm proposed in this paper with that of two basic methods, namely, greedy algorithm (GA for short) and iterative greedy algorithm (IGA for short) [20] GA calculates the maximum number of tasks that can be assigned to the worker in the unassigned task set for each worker in turn until all workers have been assigned or the task set to be assigned is empty IGA performs task allocation and task scheduling iteratively until no better solution can be found within 1000 steps, and finally selects the best allocation as the result.

Figure 4 shows the efficiency of different search algorithms It can be seen from the figure that the number of reachable tasks for each worker is very small, and the task optimization strategy will not achieve significant effect, no matter the average number of tasks, the task expiration time coefficient and the latest working time coefficient of workers are small However, with the increase of the above three parameters, the number of reachable tasks for workers will increase, and the problem will become more complex.

The sorting based search algorithm DFS+W first searches for the scheme that can be completed quickly, and quickly modifies the value of the heuristic function, so as to prune the later scheme that requires a lot of search time as early as possible Since the ordinate in Figure 4 increases exponentially, it is known that the performance of DFS+W algorithm is improved by constant coefficients compared with DFS algorithm [15, 18]. The DFS+W&U algorithm has already calculated a static upper bound for searching each node in advance in the pre calculation process, and combined with the search upper bound dynamically calculated in each step, the heuristic function value is more tightly restricted It can be seen from Figure 4 that DFS+W&U implements more efficient pruning, and with the problem scale growing, the search performance is improved by at least one order of magnitude compared with DFS.

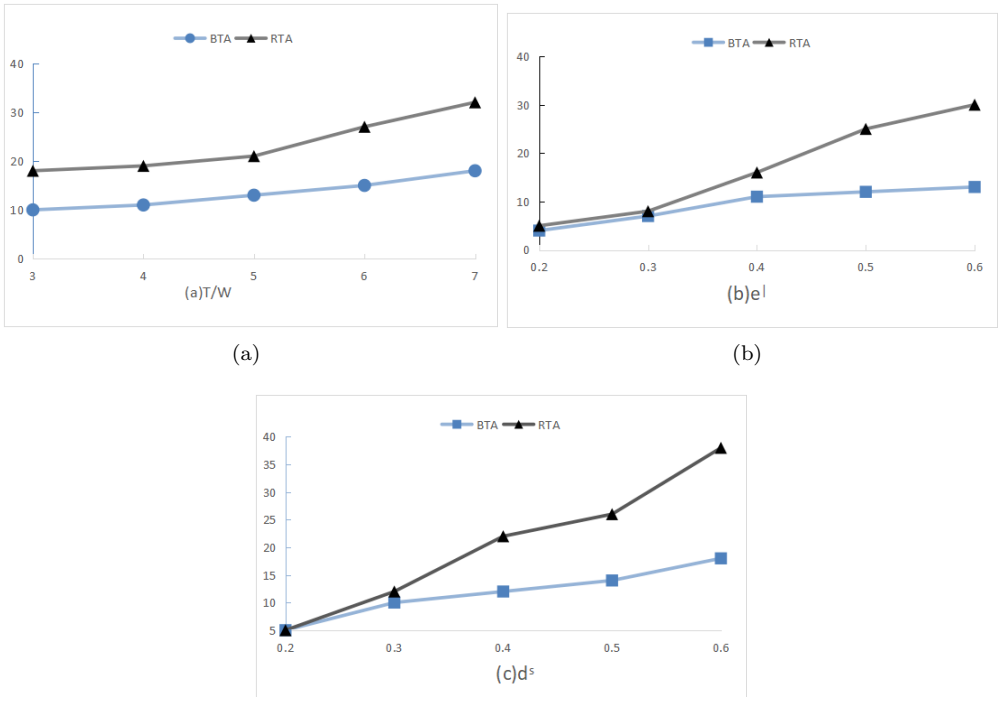
Table 1 shows the search results of GA, IGA and OPT search algorithms as the number of workers' tied tasks increases The OPT algorithm is the optimal case under the assumption. Compared with GA, the larger the data size is, the greater the difference between the number of tasks that can be allocated by OPT and the GA algorithm is. Table 2 and Table 3 show that IGA algorithm has less than 20 differences with OPT under the conditions of el¡0.5 and ds¡0.5, and has obvious advantages over GA But in more complex cases, even if the local optimization is achieved, there is still a gap of about 10% between the OPT algorithm and the global optimization algorithm.

Table 1: Effect of T/W on theTotal Number of Assigned Tasks, $e^l = d^s = 0.6$

T/W	GA	IGA	OPT
3	97	106	121
4	125	136	167
5	176	189	223
6	222	243	284
7	255	282	327

Table 2: Effect of el on the Total Number of Assigned Tasks, T/W=5, $d^s = 1$

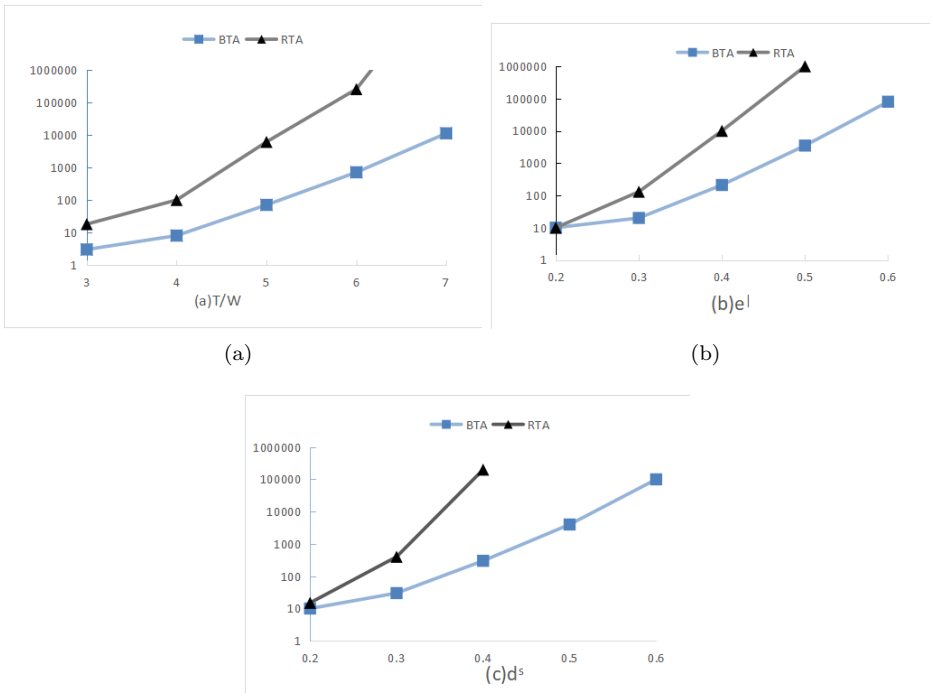
e^l	GA	IGA	OPT
0.2	113	120	124
0.3	145	154	168
0.4	174	194	216
0.5	220	243	279
0.6	263	285	328



(a) (b)

(c)

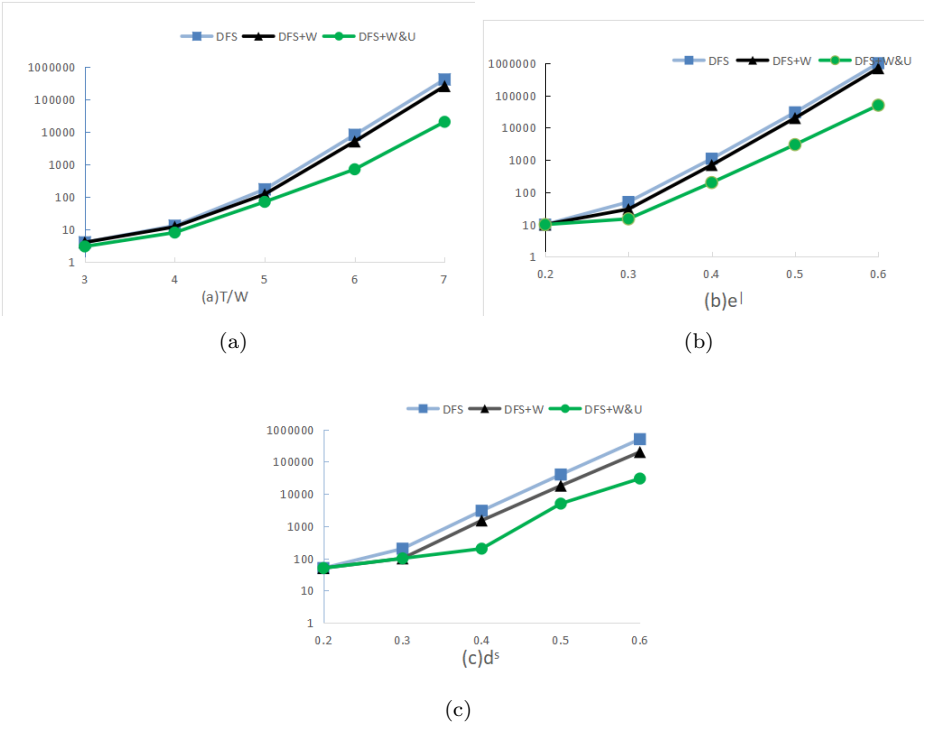
Figure 2: Effect of T/W, e^l and d^s on the depth of constructed search trees



(a) (b)

(c)

Figure 3: Effect of T/W, e^l and d^s on the Search Time



(a) (b)

(c)

Figure 4: Effect of T/W, e^l and d^s on different search strategies

Table 3: Effect of d^s on the Total Number of Assigned Tasks, T/W=5, $e^l = 1$

d^s	GA	IGA	OPT
0.2	140	143	148
0.3	165	172	183
0.4	218	226	242
0.5	259	273	295
0.6	298	317	342

6 Discussion and Conclusions

Tasks in spatial crowdsourcing have specific location requirements, and only when workers actually travel to the designated location can they complete the task. This paper studies the task assignment problem with the latest working time constraint for workers This paper creatively proposes a tree decomposition method, which divides workers without task dependency into mutually independent worker sets, and uses the best effort search tree construction algorithm to construct a balanced search tree as much as possible Finally, this paper designs a depth first search algorithm, which combines the optimization strategy to quickly tighten the upper and lower bounds, and can effectively cut the scheme that is not likely to be the optimal solution The experimental results show that the optimal assignment algorithm proposed in this paper has more advantages in complex task assignment scenarios.

The next research direction is to mine the parallelizable part of this algorithm, and use parallel algorithms and distributed algorithms to achieve more efficient allocation schemes, so as to facilitate the promotion of this algorithm to the practical application scenarios with larger and more complex data.

Acknowledgments

This work is supported by Scientific Study Project for Institutes of Higher Learning,Ministry of Education,Liaoning Province (LQN201720),Natural Science Foundation of LaioNing Province, China(20170540819) and 2023 ZheJiang Traditional Chinese Medicine Science and Technology Program (Fund No. 2023ZF010).The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] A. A. Alabbadi, M. F. Abulkhair, "Multi-objective task scheduling optimization in spatial crowdsourcing," *Algorithms*, vol. 14, no. 3, pp. 77, 2021.

[2] A. Ballatore, T. J. Verhagen, Z. Li, *et al.*, "This city is not a bin: Crowdmapping the distribution of urban litter," *Journal of Industrial Ecology*, vol. 26, no. 1, 2022.

[3] N. Bhaskar, "Optimal processing of nearest-neighbor user queries in crowdsourcing based on the whale op-

timization algorithm," *Soft Computing*, vol. 24, no. 3, 2020.

[4] S. S. Bhatti, Y. Chang, X. Gao, *et al.*, "Affinitive diversity-aware task allocation in spatial crowdsourcing," in *IEEE International Conference on Web Services (ICWS'20)*, 2020.

[5] S. S. Bhatti, J. Fan, K. Wang, X. Gao, F. Wu and G. Chen, "An approximation algorithm for bounded task assignment problem in spatial crowdsourcing," *IEEE Transactions on Mobile Computing*, vol. 20, no. 8, pp. 2536-2549, 2021.

[6] J. Jung, B. Kim, *Method and system for crowdsourcing content based on geofencing*, KR20190073032A, South Korea, 2020.

[7] A. R. Kurup, G. P. Sajeev, J. Swaminathan, "Aggregating reliable submissions in crowdsourcing systems," *IEEE Access*, 2021.

[8] M. Li, J. Wang, L. Zheng, *et al.* , "Privacy-preserving batch-based task assignment in spatial crowdsourcing with untrusted server," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pp. 947–956, 2021.

[9] Y. Li, Chang L , Li L , *et al.* , "TASC-MADM: Task assignment in spatial crowdsourcing based on multi-attribute decision-making," *Security and Communication Networks*, 2021.

[10] X. Liu, J. Fu, Y. Chen, *et al.* , "Trust-aware sensing quality estimation for team crowdsourcing in social IoT," *Computer Networks*, vol. 184, no. 6, 2021.

[11] Z. Liu, Li K , Zhou X , *et al.* , "Multi-stage complex task assignment in spatial crowdsourcing," *Information Sciences*, vol. 586, pp. 119-139, 2022.

[12] L. Qian, G. Liu, F. Zhu, Z. Li, Y. Wang and A. Liu, "Enhancing user experience of task assignment in spatial crowdsourcing: A self-adaptive batching approach," *IEEE Access*, vol. 7, pp. 132324-132332, 2019.

[13] W. A. de O. Soler, M. O. Santos, K. Akartunali, "Decomposition based heuristics for a 1ot sizing and schedu1ing problem on mu1tip1e heterogeneous production 1ines with perishab1e products," *Pesquisa Operacional*, vol. 41. e240377.. 2021.

[14] Y. Sun, M. Liu, L. Huang, *et al.* , "An embedding-based deterministic policy gradient model for spatial crowdsourcing applications," in *IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD'21)*, 2021.

[15] Y. Sun, W. Tan, "Combining spatial optimization and multi-agent temporal difference learning for task assignment in uncertain crowdsourcing," *Information systems frontiers*, vol. 22, pp. 1447-1465, 2020.

[16] Q. Tao, Y. Tong, Z. Zhou, *et al.*, "Differentially private online task assignment in spatial crowdsourcing: A tree-based approach," in *International Conference on Data Engineering*, IEEE, 2020.

[17] Y. Tian, B. Song, T. Ma, A. Al-Dhelaan and M. Al-Dhelaan, "Bi-tier differential privacy for precise auction-based people-centric IoT service," *IEEE Access*, vol. 9, pp. 55036-55044, 2021.

[18] Y. X. Tong, Y. Yuan, Y. R. Cheng, L. Chen, G. R. Wang, "Survey on spatiotemporal crowdsourced data management techniques," *Journal of Software*, vol. 28, no. 1, pp. 35-38, 2017.

[19] K. Xiong, Y. Dong, Z. Guo, *et al.*, "Exploring the ranking, classifications and evolution mechanisms of research fronts: A method based on multiattribute decision making and clustering," *International Journal of Information Technology & Decision Making*, vol. 22, no. 1, pp. 157-185, 2023.

[20] C. Zhang, Y. Guo, P. Lin, *et al.*, "Location prediction-based task assignment in spatial crowdsourcing," *Journal of Nanjing University (Natural Science)*, 2018.

[21] S. Zhang, T. Zhang, S. Z. Li, *et al.*, "Geo-indistinguishable mechanisms for spatial crowdsourcing via multi-objective evolutionary optimization," *arXiv e-prints*, 2022.

Biography

Hui Xia is currently an associate professor in Software College of Shenyang Normal University. He received received the B.S. and M.S. degree from XiDian University, China in 2003 and 2006, respectively.He has authored or coauthored more than twenty journal and conference papers. His current Acknowledgments research interests include data mining, privacy preserving and network security.

Shufeng Zhang, Master of Engineering, Associate Professor, Senior Engineer, graduated from Soochow University and currently works at Suzhou Industrial Park Institute Of Services Outsourcing, has over ten years of experience in project development and enterprise training.research interests include computer software design, development, and artificial intelligence theory research.

Weiji Yang works in Zhejiang TCM university, got bachelor's degree of computer and science in 2005, received double master's degrees of engineering and medicine in 2009 and 2014 respectively, the main research area is artificial intelligence, digital medical image processing and analysis, and smart health care, etc.

An Improved Time-Varying Collaborative Filtering Algorithm Based on Global Nearest Neighbor

Xuqi Wang, Weichao Zhang, and Qianchang Xu
(Corresponding author: Xuqi Wang)

Department of Electronic and Information Engineering, Xijing University
Xi′an 710123, China
Email: xuqi.wang@163.com

(Received Nov. 21, 2022; Revised and Accepted July 25, 2023; First Online Oct. 6, 2023)

Abstract

Collaborative filtering has widely been applied to collaborative filtering recommendation systems to filter a tremendous amount of information and screen the information that may be of interest to the target users through the evaluation and feedback of everyone. In contrast, traditional collaborative filtering does not consider the impact of project type and time on user interest changes. A modified time-varying collaborative filtering algorithm (TVCFA) is presented to address the problem of information overload. It is based on a content recommendation algorithm and reduces the static score error of different user features of the similarity expression. It uses the weight values of time and periods to describe the dynamic characteristics of the user-item score. The algorithm's feasibility is verified by simulation using the public data set. It can improve its recommendation model by 6.13% and 2.69%, respectively, compared with the UCF and ICF models in the global nearest neighbor. The experimental results show that the improved dynamic, collaborative filtering algorithm can improve the accuracy and track the dynamic characteristics of users.

Keywords: Collaborative Filtering; Dynamic Characteristics; Information Overload; Similarity Expression

1 Introduction

Recommender system is widely used in social life, its purpose is to recommend affordable, personalized, high matching products. Because users sometimes do not know exactly what they want, the recommendation system will also consider psychological factors.

It is not just a concept of decision-making theory, but a typical method of using collective wisdom. Collaborative filtering (CF) generally discovers a small part of the mass of users that are similar to your taste. In CF, these users become neighbors, and then organize a sorted directory based on other things they like to recommend to you. The recommendation system must develop and maintain a user model or user profile to store preference information, so that personalized recommendations can be made. CF recommendation algorithm originated from information filtering, and is applied to the classification of email and news. The earliest sales company received a large number of emails, but could not reasonably classify them and find out which emails would be of interest to them. This filtering algorithm is different from simple text filtering. In addition to some keywords, it also makes it possible to analyze some content manually.For example, in music appreciation. At the same time, new features can be summarized from some fixed filtering information.

However, the advantages and disadvantages coexist. The following disadvantages exist when using CF algorithms to implement information filtering: First, even if users have special needs and interests in information, they cannot recommend it or filter it if they have never made relevant evaluations. Second, the CF algorithm analyzed from the user group does not have high accuracy for new users. This is because the latest evaluation or the first input information may not be the most interesting and demanding. Generally, it is obtained through explicit and implicit methods.

"The light bulb is on because you turned on the light", "I washed the dishes because my brother washed them last time", people often use explanations when communicating contract reasoning, distinguishing four explanations: functional, causal, willing and scientific.

The recommendation system provides a set of solutions (such as products), so users may need to explain why the proposed solutions are beneficial to them.

This explanation is interpreted as how to explain. The recommendation system includes CF recommendation, content-based recommendation, knowledge-based recommendation, and hybrid recommendation methods. The CF recommendation system is that if users have the same

preferences in the past, people can predict that they will have similar preferences in the future. For example, if user A and user B have similar purchasing experiences, and user A has purchased an air purifier, but user B does not know the brand of this air purifier at present, then user B will be recommended to this brand of air purifier. Because this recommendation logic contains implicit coordination between two users, it is also called CF (Collaborative Filering).

The nearest neighbor recommendation based on items is widely used. Large e-commerce websites with millions of users and items are scanning potential neighbors. Due to the huge amount of computing, it is difficult to achieve real-time computing.

Item based recommendations can also be calculated in real time when the scoring matrix is large (sarwar et al. 2001). As the name implies, it is to use the similarity of items to calculate the predicted value [15].

CF algorithms are divided into two different types due to the difference in principle: item-based method and user-based method [11]. There are many recommendation algorithms, such as content-based recommendation, model-based real-time recommendation and CF recommendation. This kind of combined application recommendation is called Hybrid Recommendation. The combination of content recommendation and CF recommendation is the most studied and applied. The simplest way is to use the content based method and CF recommendation method to generate a recommendation prediction result, and then use some method to combine the results. The feature of content based recommendation is that it does not need to know any information about the item, and its advantage is that the system does not need to pay a huge price to update the item information. Content based recommendation directly recommends items to users according to their favorite items. Content is the description of the item. The content representation of the item generally maintains a detailed list of each item's features, such as attribute sets, feature sets, and item records. Knowledge-based recommendation system is divided into constraint based recommendation and instance based recommendation. It requires users to specify requirements, and then the system tries to provide solutions. Content based recommendation generally uses the relevant keywords that appear in the document, and uses different methods to convert the document content to the keyword list. Knowledge based recommendation is a specific type of recommendation system. It uses domain ontology to express semantic knowledge and increase the associated information between projects; Through the different effects of junction point, edge, depth and density in domain ontology on similarity calculation, the algorithm combines the concept of mutual information correlation in information theory to improve the similarity calculation formula and improve the calculation accuracy. When people use some commercial websites, they will pop up an explicit scoring collection window, because most recommendation systems use CF algorithms. However, buying a house, a mobile phone, or a car is not as frequent as buying clothes or books, so the effect will be poor because of sparse ratings (burke 2000). The CF system is not applicable at this time. The description of a certain residential area is still in the period when it was newly built five years ago. For current house buyers, their description is not appropriate, and content based recommendations are therefore not applicable. Knowledge based recommendations are specific to users such as cars and houses. For example, a house with three bedrooms and two living rooms is located between the 5-20th floor, the car color is white, and the budget is no more than 200000 yuan. It does not filter items based on the ratings of individual users. It is a conversational system with strong interaction. There are many recommendation algorithms, such as content-based recommendation, model-based real-time recommendation and CF recommendation. This kind of combined application recommendation is called combined recommendation (HR) [11, 17]. The combination of content recommendation and CF recommendation is the most studied and applied. The simplest way is to use the content based method and CF recommendation method to generate a recommendation prediction result, and then use some method to combine the results. The hybrid recommendation methods are divided into integrated hybrid design, parallel hybrid design and pipeline hybrid design:

1) The overall design is to integrate the centralized recommendation strategy into one algorithm to achieve hybrid design.

2) The parallel hybrid design runs the parallel hybrid recommendation systems independently of each other, generates recommendation lists respectively, and then combines them into the final recommendation set.

3) The pipeline hybrid design connects multiple recommendation systems according to the pipeline architecture, take the output of the previous recommendation system as the input of the latter recommendation system [3, 5].

Many CF algorithms only considers the static behavior data of ideal users; however, they ignore the characteristic differences of user and the time-varying characteristics of item ratings [4, 7, 9], and its accuracy and timeliness can no longer meet the requirements. To this end, this paper introduces similarity expressions, combines existing collaborative filtering algorithms, constructs improved similarities and weights them with time point and period parameters, and presents an improved time-varying collaborative filtering algorithm that can characterize different user characteristics. Using the public data set, simulation results show that the average recommendation error of the algorithm is significantly lower than that based on items and users, and has better accuracy and effectiveness.

2 Construct Similarity Expression

Whether in Information retrieval (IR) or Recommendation system (RS), the general idea is to conduct reasonable data processing first, and then calculate the similarity to get the results.

For example, in a recommendation system, it is first necessary to obtain the data of users or items, and construct the embedding of users or items by isomorphic and reasonable representation functions, and then calculate the similarity between users or items. Because the feature extraction of text information is relatively easy, the similarity based recommendation method has been widely used in the field of text recommendation. The user preference document is constructed based on historical data, the similarity expression between recommended items and user preference document is constructed, and the most similar items are recommended to users.

Word Frequency Inverse Document Frequency (TF-IDF) is a constant weighting technique for information retrieval and text mining. This technique is a statistical method to evaluate the importance of a word to a file set or a file in a corpus. The importance of a word increases proportionally with the number of times it appears in the file, but decreases inversely with the frequency of its appearance in the corpus. The features of recommended project documents and user preference documents are represented by keywords, the weight of each feature value is determined by using TF-IDF, the more the exact keyword appears in a single document, the higher the keyword weight, but the more a single keyword appears in multiple documents, the weight of the keyword decreases for any document.

Let N be the number of documents contained in the document set, and the number of documents containing the keyword k_i in the document set is n_i. $f_{i,j}$ is the number of times the keyword k_i appears in the document d_i. The word frequent $TF_{i,j}$ of k_i in the document d_j is defined as

$$TF_{ij} = \frac{f_{ij}}{\max_z f_{zj}} \tag{1}$$

where Z is a keyword that appears in the document d_j. Inverse frequency IDF_i of k_i in the document set is as

$$IDF_i = \log \frac{N}{n^i} \tag{2}$$

The k-dimensional vectors $d_j = (w_{1j}, w_{2j}, \cdots, w_{kj})$ and $d_c = (w_{1c}, w_{2c}, \cdots, w_{kc})$ are used to represent the item document and the configuration document of user C, respectively, k is the number of keywords, and each vector component is calculated as

$$w_{ij} = TF_{ij} \cdot IDF_i = \frac{f_{ij}}{\max_z f_{zj}} \cdot \log \frac{N}{n_i} \tag{3}$$

The cosine distance is often used to calculate the similarity $sim\,(c, d_j)$ between project documents and user configuration documents. It is calculated as

$$sim\,(c, d_j) = \cos\,(d_c, d_j) = \frac{\sum_{i=1}^{m} w_{ic} w_{ij}}{\sqrt{\sum_{i=1}^{m} {w_{ic}}^2} \sqrt{\sum_{i=1}^{m} {w_{ij}}^2}} \tag{4}$$

The content recommendation algorithm can use similarity parameters for recommendation. However, the phenomenon of synonyms and polysemous words will still cause a significant error in the similarity calculation, which needs further improvement.

3 Adding User Features

Characteristic differences must exist between different users. However, the interests and hobbies of the same type of user groups have certain similarity, so the similarity function can be combined with the traditional collaborative filtering algorithm to calculate the similarity of user groups with similar characteristics. Improve the accuracy of recommendations. User characteristics can be selected from any aspect. Here, four of the most common descriptions are selected.

Age groups. The user characteristics of different age groups are obviously different. Using Peng Dewei's method of classification [1], the age groups are divided into six stages: less than 6 years, 7 to 11 years, 12 to 15 years, 16 to 22 years, 23 to 30 years, 31 to 40 years, 41 to 50 years, 51 to 60 years, more than 60 years, recorded as 0, 1, 2, 3, 4, 5, 6.

Gender. Gender is an important parameter in distinguishing user characteristics. In many choices, gender difference has a great impact on users' interests and behaviors. Men and women are recorded as M, F.

Occupation. According to a large number of literature research conclusions, users of different occupations understand the same problem at different levels. Based on this conclusion, it can be considered that other occupations reflect different life experiences, and people with the same occupation are more likely to have the same way of understanding things. Depending on the occupation, the occupational characteristics can be recorded as $0, 1, 2, 3, \cdots, n$.

Education. People with different educational backgrounds have experienced different levels of education popularization, from primary school, junior high school, high school (technical secondary school), college, undergraduate, master's and doctoral degrees, recorded as 0, 1, 2, 3, 4, 5, 6.

Establish a feature data table based on the user feature data. The data table contains the user?s age, gender, occupation, and educational background data. User

similarity $\mathrm{sim}\,1(u,v)$ based on user feature data can be obtained.

$$\mathrm{sim}\,1(u,v) = \sum \mathrm{similar}\left(U_{un}, U_m\right)/k \tag{5}$$

Equation (5) reflects the ratio of the number of identical features of users u, v, U_{um} is the m-th characteristic attribute of user u, U_{un} is the n-th characteristic attribute of user v, and k is the total number of user characteristics.

To reduce the deviation of the scoring scale of different users, an improved cosine similarity measurement method is used to subtract the average score of users. Assume that the item set rated by user u and v is I_{uv}, I_u and I_v denote the set of items rated by user u and user v, respectively, and the similarity between user u and user v is $\mathrm{sim}\,2(u,v)$ denoted as

$$\mathrm{sim}\,2(u,v) = \frac{\sum_{c\in I_u}\left(R_{u,c}-\bar{R}_u\right)\left(R_{v,c}-\bar{R}_v\right)}{\sqrt{\sum_{c\in I_u}\left(R_{u,c}-\bar{R}_u\right)^2}\cdot\sqrt{\sum_{c\in I_v}\left(R_{v,c}-\bar{R}_v\right)^2}}$$

where, $R_{u,c}$ is the rating of user u on item c, $\bar{R}_u$ and $\bar{R}_v$ are the average rating between users u v on item c, respectively.

The user similarity parameter $\mathrm{sim}\,1(u,v)$ is introduced into the modified cosine similarity formula $\mathrm{sim}\,2(u,v)$ to obtain the end-user similarity.

4 Collaborative Recommendation Algorithm

The traditional CF algorithms only consider the static project rating data and ignore the dynamic characteristics of the project and the users, and the calculation results are flawed [2,8,12,19,20]. Therefore, the influence of different times on behavior characteristics should be considered in traditional algorithms [6]. The hybrid recommendation method combines multiple recommendation technologies to make up for each other's shortcomings, so as to obtain better recommendation results. Different from traditional hybrid recommendation technologies (such as weighted fusion, hybrid recommendation, cascading recommendation), this paper adopts collaborative training strategy to build a hybrid model of item based collaborative filtering recommendation (algorithm 1) and item content based recommendation (algorithm 2) when building a hybrid recommendation system. After the improvement in similarity based on attributes, the time weight and the weight of time period weight should be introduced into the final prediction score. The time weight reflects the difference in the rating of the user's item with the drift of interest. For different items, the earlier ones that appear on the same day have a smaller weight, and the later ones that appear on the same day have a larger weight to solve the phenomenon of score drift in a large time span. The weight of the time period reflects the change characteristics of the user's daily repetitive items, and the weight of the change of the characteristic parameters of different items should be different in different time periods.

By above the analysis, a method is adopted to introduce weights for time and time periods to improve the accuracy and timeliness of the recommendation algorithm.

4.1 Time Weight

The weight of time reflects the gradual trend of the past user ratings of the project. In practice, exponential time fitting is generally used to highlight the weight of users' latest project ratings and reduce the proportion of early projects. The time function is set as:

$$f\left(t_{ni}\right) = 1/\left(1+e^{-t_{ni}}\right) \tag{6}$$

where, t_ni is the time difference between the time when user n rated item i and the specific date. From the formula 7, the function $f\left(t_{ni}\right)$ is a monotonically increasing function, and its value increases with the increase of time t, but does not exceed 1. The closer the time is, the greater the time weight is, which reflects the changing status of user project ratings. In other words, all data contribute to the recommended project, and the most recent data contribute the most. The old data reflects the user's previous preferences.

4.2 Time Period Weight

According to the periodic characteristics of time, it can be considered that the changes of the user's item are cyclical, and the impact on the feature score fluctuates in different periods. First, construct the time period data table, record the time period as $0,1,2,\cdots,23$, and calculate the time period in which all ratings are based on the time stamp. q_avg is the mean value of all feature ratings, $\{q_{total0}, q_{total1}, \cdots, q_{total23}\}$ is the average rating of each time period. Set the time period function as follows:

$$q_n = q_{avg} - q_{total} \tag{7}$$

where, q_n is the time period rating weight of the time period of item n, $n \in \{0,1,2,\cdots,23\}$.

In the collaborative recommendation algorithm for final prediction ratings, time weight and time period weight are introduced, user characteristics are added when calculating user similarity, which reflects the difference between long-term and daily changes in user interest preferences of different characteristics. The formula of the CF algorithm is expressed in Equation (8).

$$p_{u,i} = \overline{R}_u + \frac{\sum_{a=1}^{n}\left(R_{u,i}-\overline{R}_a\right) sim(u,a) f\left(t_{ni}\right)}{\sum_{a=1}^{n}\left|sim\left(u,a\right)\right| f\left(t_{ni}\right)} + q_n \tag{8}$$

where, $p_{u,i}$ for the CF ratings of user u for item i.

According to the analysis listed above, the specific implementation steps of the TVCFA algorithm and the flow diagram (Figure 1) can be concluded as follows:

Step 1. Enter the user rating matrix R_{mn}, and the number of elements in the recommendation set is N.

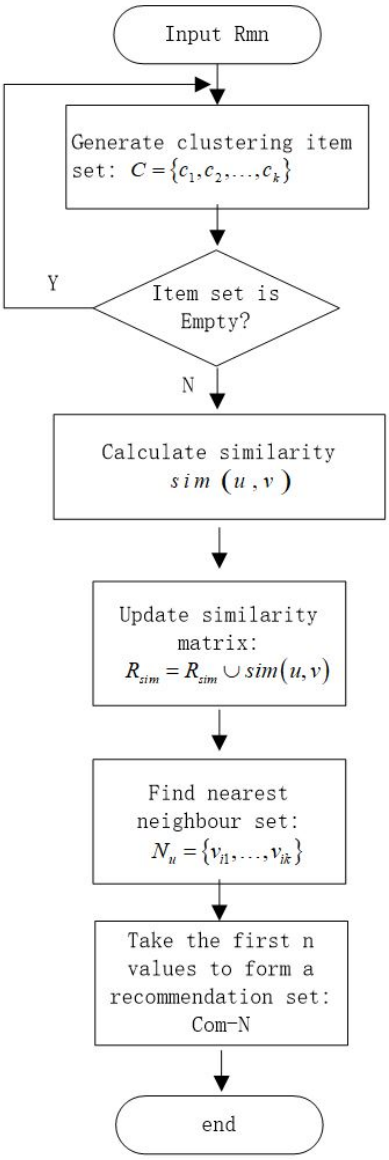


Figure 1: Algorithm flow chart based on time period weight.

Step 2. Generate the clustering item set: $C = \{c_1, c_2, \cdots, c_k\}$.

Step 3. For any user u and v, the overlap factor is used to modify the local similarity calculation, and the formula is used to calculate K local similarities $sim_j(u,v)\,(j=1,\cdots,k)$.

Step 4. Calculate the global similarity according to the global similarity measurement formula between users u and V.

Step 5. Update the similarity matrix R_{sim}, that is $R_{sim} = R_{sim} \cup sim(u,v)$.

Step 6. For each user u find the nearest neighbor set $N_u = \{v_{i1}, v_{i2}, \cdots, v_{ik}\}, u \notin N_u$, and $sim(u, v_{i1}) \geq \ldots \geq sim(u, v_{ik})$.

Step 7. Use the prediction score formula of target user u for item i to calculate the prediction score of ungraded item i, sort the prediction scores in ascending order, and take the items corresponding to the first n values to form a recommendation set Com-N.

5 Simulation

The project uses the MoviesLens dataset and GroupLens, which stores movie scores, including 1 million pieces of scoring data from 6k users on 4k movies. User data consists of five parts of user, including ID, gender, age, occupation ID and zip code. The public data set was used for simulation verification. The data included basic information such as age, gender, occupational characteristics, area code of 943 users, and their 100,000 rating data for 1682 movie items. During the experiment, the data set was divided into a training set and a test set according to the ratio of the 7:3 ratio [14].

The average error is introduced to measure the performance of CF algorithm. MAE is a matrix commonly used in collaborative filtering CF, which is usually used to evaluate the deviation between the predicted value and the actual value [16]. The smaller the deviation, the higher the prediction accuracy and the higher the recommendation quality. If the user rating set is represented by $\{p_1, p_2, \cdots, p_n\}$ and the corresponding actual user rating set is represented by $\{q_1, q_2, \cdots, q_n\}$, average error MAE expression is as follows:

$$MAE = \frac{\sum_{i=1}^{N} |p_i - q_i|}{N} \tag{9}$$

The number of the nearest neighbors in the experiment starts from 5 with an interval of 5 and increases to 40. The experimental results are shown in table 1. The results of the algorithm and the traditional algorithm based on the project-based ICF recommendation algorithm, the user-based UCF CF recommendation algorithm's MAE value are shown in Table 1 and Figure 2.

Table 1: MAE between improved and traditional algorithms

nearest neighbors	MAE		
	UCF	ICF	Improved
5	0.8621	0.8236	0.8015
10	0.8453	0.8143	0.7727
15	0.8356	0.7858	0.7638
20	0.8242	0.7739	0.7339
25	0.8138	0.7615	0.7291
30	0.8025	0.7562	0.7017
35	0.7952	0.7628	0.6907
40	0.7864	0.7536	0.6836

When the number of nearest neighbors is 5, it can be seen from Table 1 that the algorithm in the article is significantly superior to the user based collaborative filtering algorithm. From Table 1, it is seen that the MAE of the improved algorithm is 10.5% lower than UCF and 5.7% lower than ICF. According to relevant literature, the smaller the MAE value, the higher the accuracy of the recommendation. Therefore, the accuracy of the algorithm recommendation is better than UCF and ICF.

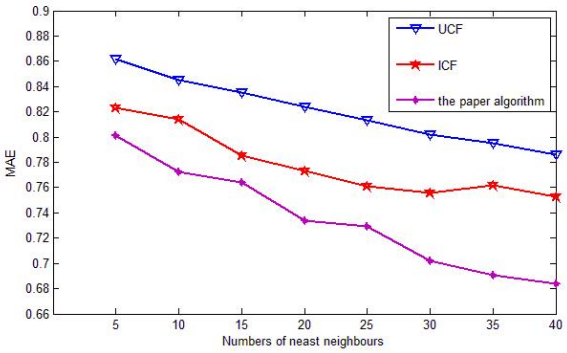


Figure 2: MAE contrast diagram of different algorithms.

In order to verify the recommendation accuracy of the algorithm proposed in this article, the algorithm proposed in this article (TVCFA), item based (ICF), and user based (UCF) recommendation algorithms are compared and analyzed under different K-nearest neighbor numbers. The experimental data are shown below. The results are shown in Figure 2.

Next, we compare the improved algorithm with the classical item-based algorithm [10, 18]. The parameters used are for different users and different project groups. Obviously, as seen from the results of Figure 3, Figure 4. Our new algorithm is able to boost the prediction accuracy for all configurations. Compared with other models, TVCFA can more accurately obtain the calculation results of the learning situation similarity by improving the time-varying CF algorithm [13], and realize the improvement of the CF algorithm, and, the global nearest neighbor method is introduced, and the improved prediction scoring formula is used for calculation to provide more accurate personalized recommendation.

1) We have conducted experimented on the MovieLens movie recommender dataset (http://www.grouplens.org/node/73). It collected by GroupLens research, and was bootstrapped from the EachMovie data set. It has made three sub-datasets available: one with 100K timestamped user ratings of movies, another with 1M ratings, and a third containing 10M ratings and 100K timestamped records of users applying tags to movies.

   The experimental results on the MovieLens data set of results are shown in Figure 3.

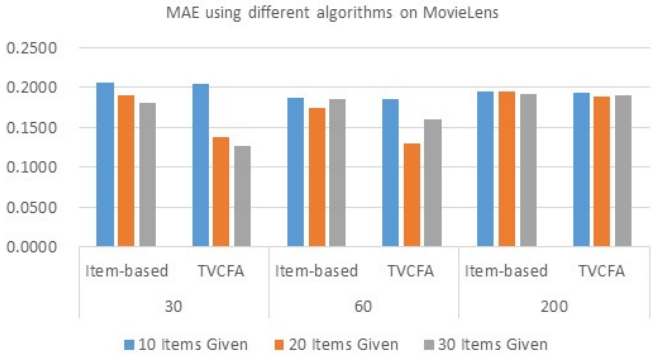


Figure 3: MAE uses different algorithms to compare results on MovieLens dataset.

2) We have conducted experimented on the GroupLens dataset. GroupLens consisted of 1,000,209 ratings for 3900 movies by 6040 users. Recommender system (https: // grouplens.org/datasets/movielens/). The experimental results on the GroupLens data set of results are shown in Figure 4.

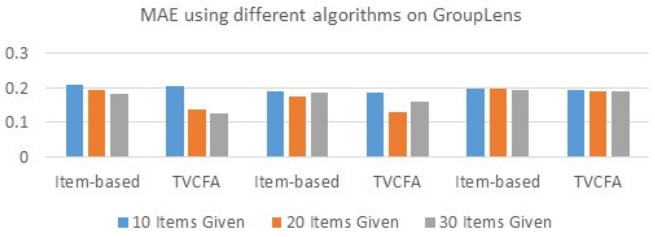


Figure 4: MAE uses different algorithms to compare results on GroupLens dataset.

The experimental results show that the proposed recommendation algorithm can effectively improve the similarity between users and significantly improve the accuracy of recommendation. In the case of massive data, it can better improve the recommendation quality. The disadvantages of this algorithm are: in the collaborative recommendation algorithm of final prediction score, if appropriate, the choice of time weighted and time weighted weights can obtain closer scores; Otherwise, the effect of

the recommendation algorithm is not good, and the accuracy of the score of the prediction knowledge points needs to be improved.

6 Conclusion

Compared with the weighted fusion hybrid recommendation, which needs to constantly adjust the weight of each recommendation result, the difficulty in sorting hybrid recommendation, and the staged process of cascade recommendation, the hybrid recommendation method based on collaborative training makes full use of the user's rating information and item content description information in each iteration of training, realizes the fusion of the two recommendation views, and achieves a better hybrid recommendation effect. A novel time-varying CF algorithm (TVCFA) is proposed by introducing user characteristic parameters and time-varying parameters that are not considered by traditional CF algorithms, the improved time-varying CF algorithm can characterize the characteristics of different users and the dynamic characteristics at different times. In the conventional algorithm, the time weight function is deployed to accurately predict the user's purchase interest. According to the evaluation time sequence, the user's recent interest change trend can be judged. It is verified by simulation using public data, compared with traditional algorithms based on items and users, this algorithm has a significant improvement in accuracy and timeliness. The future work will combine CF algorithms with convolutional neural network. Such as the determination of better convolution parameters, the calculation of similarity with other users through CF algorithm, the "cold start" of the recommendation system.

Acknowledgement

This work was supported by the second batch of the special fund for high-level talents of Xijing University in 2021(XJ21B14), Shanxi Tiandi Coal Machinery Equipment Co., Ltd. (2018-TD-MS040), Key Research and Development Plan Project of Shaanxi Provincial Science & Technology Department (Program No. 2018 ZD XM –NY -014).

References

[1] M. Alam, O. Goni, A. Shameem, S. Islam, N. K. Datta, S. Ahmed, G. Moazzam, "An approach for the normalization of short message service to detect shorter form of words and find out actual word," *International Journal of Electronics and Information Engineering*, vol. 13, no. 3, pp. 111-118, 2021.

[2] F. Cacheda, V. Carneiro, D. Fernández, V. Formoso, "Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems," *ACM Transactions on the Web*, vol. 5, no. 1, pp. 1-33, 2011.

[3] K. H. Chen, P. P. Han, J. Wu, "User clustering based social network recommendation," *Chinese Journal of Computers*, vol. 36, no. 2, pp. 349-359, 2013.

[4] Z. Cui, X. Xu, X. U. E. Fei, X. Cai, Y. Cao, W. Zhang, J. Chen, "Personalized recommendation system based on collaborative filtering for IoT scenarios," *IEEE Transactions on Services Computing*, vol. 13, no. 4, pp. 685-695, 2020.

[5] J. Feng, X. Fengs, N. Zhang, J. Peng, "An improved collaborative filtering method based on similarity," *PLOS ONE*, vol. 13, no. 10, pp. e0206629, 2018.

[6] T. Ha, S. Lee, "Item-network-based collaborative filtering: A personalized recommendation method based on a user's item network," *Information Processing and Management*, vol. 53, no. 5, pp. 1171-1184, 2019.

[7] B. Hong, M. Yu, "A collaborative filtering algorithm based on correlation coefficient," *Neural Computing and Applications*, vol. 31, no. 12, pp. 8317-8326, 2019.

[8] D. Jia, F. Zhang, S. Liu, "A robust collaborative filtering recommendation algorithm based on multidimensional trust model," *Journal of Software*, vol. 8, no. 1, pp. 11-18, 2013.

[9] H. N. Kim, I. Ha, K. S. Lee, G. S. Jo, A. El-Saddik, "Collaborative user modeling for enhanced content filtering in recommender systems," *Decision Support Systems*, vol. 51, no. 4, pp. 772-781, 2011.

[10] G. Li, Z. Zhang, L. Wang, Q. Chen, J. Pan, "One-class collaborative filtering based on rating prediction and ranking prediction," *Knowledge-Based Systems*, vol. 124, pp. 46-54, 2017.

[11] Z. J. Li, "Interior design recommendation technology based on collaborative filtering algorithm," *Modern Electronics Technique*, vol. 43, no. 13, pp. 176-179, 2020.

[12] M. Peng, G. Zeng, Z. Sun, J. Huang, H. Wang, G. Tian, "Personalized APP recommendation based on APP permissions," *World Wide Web*, vol. 21, pp. 89-104, 2018.

[13] C. Rana, S. K. Jain, "A study of the dynamic features of recommender systems," *Artificial Intelligence Review*, vol. 43, pp. 141-153, 2015.

[14] L. Ren, W. Wang, "An SVM-based collaborative filtering approach for Top-N web services recommendation," *Future Generation Computer Systems*, vol. 78, no. 2, pp. 531-543, 2018.

[15] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th International Conference on the World Wide Web*, ACM Press, pp. 285–295, 2001.

[16] R. Shaw, D. K. Agrawal, B. K. Patra, "An effective similarity measure for improving performance of user based collaborative filtering," in *IEEE EUROCON 2021-19th International Conference on Smart Technologies*, pp. 209-215, 2021.

[17] D. Wang, X. Wang, S. Yin, "A new recursive neural network and center loss for expression recognition," *International Journal of Electronics and Information Engineering*, vol. 13, no. 3, pp. 97-104, 2021.

[18] J. Yang, B. L. Liu, Z. X. Zhao, "A collaborative filtering recommendation algorithm with improved similarity calculation," *Proceedings of the 2018 Second International Conference of Sensor Network and Computer Engineering (ICSNCE'18)*, pp. 158-161, 2018.

[19] F. Zhang, T. Gong, V. E. Lee, G. Zhao, C. Rong, G. Qu, "Fast algorithms to evaluate collaborative filtering recommender systems," *Knowledge-Based Systems*, vol. 96, pp. 96-103, 2016.

[20] H. R. Zhang, F. Min, Z. H. Zhang, S. Wang, "Efficient collaborative filtering recommendations with multi-channel feature vectors," *International Journal of Machine Learning and Cybernetics*, vol. 10, pp. 1165-1172, 2019.

Biography

Xuqi Wang received the B.S. degree in computer science and technology from Xidian University, Shaanxi, China, in 1999, the M.S. degree in software engineering from Xi'an Jiaotong University, Shaanxi, in 2005, and the Ph.D. degree in communication and information system from the School of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing, China, in 2020.He is currently an Associate Professor with the Department of Information and Engineering, Xijing University, Xi'an, China. His research interests include wireless sensor networks and arterials intelligence.

Weichao Zhang received the B.S. degree in computer science and technology from North China University of Water Resources and Electric Power, Currently, he pursuing a Master's degree in Computer Technology at Xijing University.

Qianchang Xu received the B.S. degree in Mechanical manufacturing and automation from Guangdong Institute of Technology, Currently, he pursuing a Master's degree in Computer Technology at Xijing University.

Research on Bayesian-based Network Situational Awareness Algorithm

Zhiyong Luo, Shuyi Wang, Haifeng Xu, and Weiwei Song
(Corresponding author: Zhiyong Luo)
School of Computer Science and Technology, Harbin University of Science and Technology
Harbin 150080, China
Email: luozhiyongemail@sina.com
(Received Nov. 29, 2022; Revised and Accepted Aug. 21, 2023; First Online Oct. 6, 2023)

Abstract

A Bayesian network security situational awareness technique based on Bayesian networks is proposed to address the current problem of the inability to analyze and predict the security situation of established networks comprehensively and accurately. The security elements affecting the current network security situation are obtained by using the log information of the system, the operation information of the network devices, and the alert information and logs of the protection tools. Calculate the vulnerability attack probability using factors such as vulnerability value, attack cost, and attack benefit. Subsequently, the extracted security elements are inserted into a Bayesian network model to obtain the current security posture values, evaluate the current security posture, adapt the network environment with security countermeasures, and predict future network security trends by results.

Keywords: Bayesian Network; Network Security; Network Situational Awareness; Situational Indicators; Situational Prediction

1 Introduction

With the progress of science and technology, computer networks have become an indispensable part of people's lives. However, they face various cyber attacks while changing people's production and lifestyle. To cope with the increasingly complex and covert network threats, various detection techniques have emerged, such as vulnerability detection techniques, malicious code detection techniques, and intrusion detection techniques *etc.* An intrusion refers to any malicious activity that violates the confidentiality, integrity, or availability of data and equipment [1]. These technologies try to detect possible security problems in the network from different perspectives and build a security line of defense for the network. However, with the continuous development and progress of the network, many computer networks are now facing a high level of network threats that are difficult for security administrators to detect but continue to occur. These threats present a trend of scale, proceduralization, and concealment.

In view of the above, the research and application of Network Security Situation Awareness [16] (NSSA, for short) has received more attention from researchers. Different from traditional security measures, NSSA can identify the behavior of various activities in the network, understand the intent and assess the impact from a macro perspective, then provide reasonable decision support and predict the trend of network security.

As a cyber risk analysis-based algorithm, there are precedents for its use in both industry and government. For example, Riptide Networks in the US and Toshiba in Japan are applying the algorithm to monitor security threats in their networks; Telstar Technologies has developed its own security situational awareness algorithm to enable enterprises to monitor and predict cyber risks in real time and take defensive measures. The UK's National Cyber Security Centre (NCSC) uses the algorithm to detect and analyze attack threats nationwide; in addition to the UN's growing focus on cyber security, the UN Cyber Security Working Group uses the algorithm to conduct nationwide cyber threat alerts.

At present, inadequate assessment functions for network situational awareness [10]. Nikoloudakis *et al.* [9], proposed a machine learning-based situational awareness framework to enable situational awareness. Zhu *et al.* [18], changed the existing quantitative assessment technique of hierarchical threat situations by introducing time parameters in basic probability assignment. Zhang *et al.* [14], constructed an LSTM-DT network security situation assessment model, focusing on the time series problem of network security situation assessment. Hu *et al.* [2] proposed an attack prediction algorithm for a dynamic Bayesian attack graph and a security situation quantification algorithm based on attack prediction, combined with CVSS to dynamically predict possible attack paths and probabilities *etc.* Tao *et al.* [12], proposed a new unit situation awareness method based on autoencoder and

simple storage to achieve accurate and efficient situation awareness. Kou *et al.* [5], proposed a security posture assessment method based on attack intent identification, which simplifies the intrusion path based on the intruder to achieve the status assessment. Zhao *et al.* [17], for network security situational awareness in the big data environment, first, established a network security situational awareness index system, selected and quantified index factors, and then calculated the situational values to construct a network security situational awareness system. For the selection and quantification of index factors, multiple sources of data in the big data environment are selected, and a parallel imputation algorithm based on the attribute importance matrix is proposed to reduce the data source and data attributes. For the calculation of context, the traditional wavelet neural network learning method is easy to fall into local minima, the wavelet neural network parameters are optimized by the particle swarm algorithm, and then the particle swarm optimization-based wavelet neural network is applied to calculate the situation values. Lei *et al.* [6], in order to defend against mobile targets in the network and calculate the cost and benefit, used the attack graph to establish a hierarchical network resource graph. Proposed a network moving target defense effectiveness evaluation method based on variable point detection combined with the variable point detection method, which can effectively improve the efficiency of the network resource graph construction. Sun *et al.* [11], proposed a probabilistic approach and implemented ZePro, a prototype system for zero-day attack path identification. Bayesian networks can calculate the probability of an object instance being infected by taking intrusion evidence as input. Zhang *et al.* [15], proposed a hierarchical cybersecurity situational awareness data fusion method in a cloud computing environment to establish a hierarchical model. The hierarchical cybersecurity situational data are collected and processed in parallel through cloud computing technology. Husak *et al.* [4], proposed that many prediction methods in cybersecurity use models to represent and predict the future state of an attack or security situation. Although there is a clear division in the use cases of models (attack prediction more often uses discrete models, while predicting cybersecurity situations mainly uses continuous models), these two main use cases often complement each other as well as overlap in many cases. Second, many new approaches based on data mining and machine learning have significantly changed the state of research in cybersecurity prediction. Data mining addresses the reliance on manually provided predictive models, while machine learning challenges model-based approaches. Hu *et al.* [3], proposed a support vector machine and adaptive weight-based network security posture assessment model, training the collected samples using Libsvm and adaptive weight strategy by designing network data feature terms and indicator values and collecting some network data for predictive analysis. López-Cuevas *et al.* [7], proposed a new visualization method to track and identify in real time when a person is in a risk-prone state. The model can provide decision-makers with a visual description of individual or group physiological behavior; through it, decision-makers can infer whether further assistance is needed if a risky situation exists. Wang *et al.* [13], calculated the reachable probability of each node by Bayesian theory, describes the probability of single-step attack occurrence, dynamically predicts the potential risk in the network, and proposes an improved intrusion prediction algorithm based on attack graphs, which simplifies the connection between alert evidence, as well as attack behavior and improves the accuracy of prediction.

The above research results are advanced for network situational awareness. However, attack graphs do not provide information about the probability of vulnerabilities or information about the severity of vulnerabilities, and Bayesian attack graphs are not a complete model for mitigating network risks because they do not inherently consider the set of security countermeasures to address vulnerabilities. To quickly and accurately reflect the current security posture of the network and adjust the network using security countermeasures, this paper uses a Bayesian network approach to sense, analyze, and predict the network posture.

2 Bayesian Network Situational Awareness Model

2.1 Network Security Situational Awareness

Network situational awareness refers to a holistic situation in which the internal system state, external behavioral state, and internal user state of the network are in balance with each other. The goal is to enable commanders on both sides of a military game to be informed of the other side's military behavior and to make military judgments in their favor.

In 1988, Endsley proposed that "situational awareness is the understanding and perception of factors or events and the prediction of the future state of development that affect the environment in a certain time and space." The process of achieving this definition is divided into three levels: acquisition of situational elements, understanding of the situational situation, and prediction of the situational situation, as shown in Figure 1.

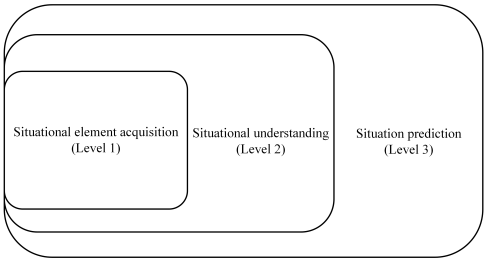


Figure 1: Endsley situational awareness process

In 2000, Tim Bass first proposed the concept of network security situational awareness by integrating situational awareness and network security technologies, while he also proposed a five-layer framework model for network security situational awareness based on multi-sensor data in one of his papers, namely: data collection, security event object extraction, situational extraction, threat assessment, and resource management, which indicated a more detailed research direction for later researchers, as shown in Figure 2.

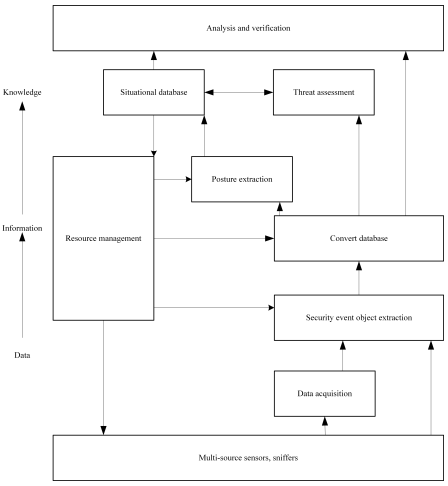


Figure 2: Tim Bass network security situational awareness model

Situation assessment is the key part of the process. The so-called security posture assessment refers to the establishment of a suitable mathematical model based on the construction of security indicators. The security events generated by the summary, filtering, and correlation of the analysis of devices to assess the extent of security threats to the network system, and the analysis of the stage of network attacks to achieve a comprehensive grasp of the overall security posture of the network.

2.2 Bayesian Network Model and Related Definitions

Definition 1. *Bayesian network model (BNM) is a directed acyclic graph that can be expressed as* $BNM = (S, A, E, \tau, P)$*, The specific definitions are as follows.*

1) *S is the set of resource attribute nodes whose state takes the value of true or false, denoted as $S = 1$ or $True$, as well as $S = 0$ or $False$. The nodes are divided into three categories, S_{start} is the initiating node of the network attack, $S_{transition}$ is the process node of the attack, and S_{target} is the target node of the attacker.*

2) *A denotes the set of atomic attacks, representing the attacker's attack behavior on the node vulnerability. The current attack behavior can be either present or absent, which is respectively represented as $a_i = 1$ or $a_i = 0$.*

3) *E is the set of directed edges in a Bayesian network, and the edges connecting the nodes are represented by a set of ordered pairs, denoted as τ, which reflect the causal relationships between attribute nodes in the attack behavior, where $E \in (S_{pre}, S_{\text{past}})$ denotes the edges of the attacker.*

4) *τ represents the conjunction or disjunction relationship between multiple edges pointing to a node, and its possible values are expressed as $\{AND, OR\}$, if $\tau = AND$, it means that the relationship between the incoming edge and node s_j is AND, the product rule shown in Equation (1) is used; if the relationship between the incoming edge and s_j is or ($\tau = OR$), then Equation (2) is used to calculate.*

$$\Pr(s_j \mid Pa[s_j]) = \begin{cases} 0, \exists s_i \in Pa[s_j], s_i = 0 \\ \Pr\left(\bigcap_{s_i=1} e_i\right) = \\ \quad \prod_{s_i=1} TP(e_i), otherwise \end{cases} \tag{1}$$

$$\Pr(s_j \mid Pa[s_j]) = \begin{cases} 0, \forall s_i \in Pa[s_j], s_i = 0 \\ \Pr\left(\bigcup_{s_i=1} e_i\right) = \\ 1 - \prod_{s_i=1} [1 - TP(e_i)], \ otherwise \end{cases} \tag{2}$$

5) *P denotes the reachable probability of a node in a Bayesian network, and there exists a Conditional Probability Table (CPT) for each state node, which shows the probability of a node given the state of its parent node, and P denotes its CPT set. Suppose the set of S_i's parents in the target network is S_j , then P contains the conditional probability $P_{S_i|S_j} = P_B(S_i \mid S_j)$.*

In the directed acyclic graph BNM, there can be N attribute nodes. During the process of network security situational awareness, each attribute node can represent a factor that influences the situational factors, such as network vulnerabilities, external anonymous attacks, alarm information and so on. In the directed edge (S_i, S_j), S_i exists as the parent of S_j, and the set of all parents of S_i can be represented by $P_B(S_i)$. In the Bayesian network, there is a conditional independence requirement for each node, and any node S_i is conditionally independent of all nodes in the non-$S_i$descendant node set $A(S_i)$. As shown in Equation (3).

$$P(S_i/A(S_i)) = P(S_i/P_B(S_i)) \tag{3}$$

In the Bayesian network model, the conditional probability table is denoted by P. Given a set of parent nodes, each attribute is assumed to be conditionally independent of its non-child descendants, the joint probability distribution among all attributes is shown in Equation (4).

$$P(S_1, \ldots, S_n) = \prod_{i=1}^{n} P(S_i/P_B(S_i)) \tag{4}$$

2.3 Bayesian Network Quantification

2.3.1 Vulnerability Value

Large networks today consist of numerous hosts, each of which contains several vulnerabilities. As a result, determining the probability of exploiting each vulnerability by relying solely on expert knowledge can be a tedious, time-consuming, and error-prone task.

The value of a vulnerability is related to the ease of exploitation and impact of the vulnerability of the attribute node, and is generally quantified using the vulnerability scoring system (CVSS, common vulnerability scoring system) provided by the National Vulnerability Database (NVD). CVSS provides complete scoring parameters and an open scoring framework that combines dynamic evaluation and dependencies of vulnerabilities among attribute nodes to quantify the ease of vulnerability exploitation.

CVSS provides a set of metrics, namely: Base, Temporal and Environmental, to quantitatively assess the severity of existing security vulnerabilities. Base is used to describe the intrinsic characteristics of a vulnerability using two attributes: (1) Exploitability. (2) Impact; Temporal quantifies the vulnerability characteristics that vary with events; Environmental captures the vulnerability characteristics associated with a particular IT environment. Multiple values can be assigned to each CVSS metric. In this paper, metrics are selected from the Base and Temporal of the CVSS to calculate the probability of attack vulnerability. As a result, the vulnerability exploitation probability calculated at the time of evaluation is more accurate and closer to the actual situation. Equation (5) gives the equation for calculating the score representing the value of vulnerability, which is calculated as:

$$\text{Score} = \begin{cases} \text{Min}(1.08(\text{Exp}+\text{Impact}), 10), & \text{Scope}=C \\ \text{Min}(\text{Exp} + \text{Impact}, 10), & \text{Scope} = U \end{cases} \quad (5)$$

Here, *Impact* indicates the vulnerability impact degree, *Exp* indicates the exploitability of the vulnerability, 10 indicates that the maximum value of score is 10, and other constant coefficient values are set by CVSS according to the security policy. Where *Exp* is calculated as shown in Equation (6).

$$Exp = 2 \times AV \times AC \times AU \quad (6)$$

Here AV is the access vector, AC is the access complexity, and AU is the authentication instance. *Impact* is calculated as shown in Equation (7).

$$\text{Impact} = \begin{cases} 7.52(ISC - 0.029) - 3.25 \\ \left((ISC - 0.02)^{15}\right), & \text{Scope} = C \\ 6.45ISC, & \text{Scope} = U \end{cases} \quad (7)$$

The ISC represents the intermediate constant and is calculated as in Equation (8).

$$ISC = 1 - ((1 - C) * (1 - I) * (1 - A)) \quad (8)$$

C,I, and A denote Confidentiality (C), Integrity (I), and Availability (A) in the CVSS scoring system.

The CVSS metric scoring system is shown in Table 1.

The vulnerability value indicates the possibility of an attacker exploiting a vulnerability. For vulnerability vi, the vulnerability value is expressed by $Value(v_i)$, and the size is related to the scoring process described above. Since the CVSS standard vulnerability score takes values in the range of [0,10], the $Value(v_i)$ is calculated as Equation (9) for the convenience of the post-order calculation.

$$\text{value}\ (v_i) = \frac{\text{Score}}{10 * 100\%} \quad (9)$$

2.3.2 Attack Benefit

Definition 2. *Atomic attack gain (AProfit) is the gain obtained by implementing an atomic attack, and this paper quantifies the attack gain in terms of resource loss.*

Definition 3. *Resource Loss (RL): Indicates the loss suffered by a resource after an atomic attack, which is defined in this paper to describe the resource loss in three aspects: Attack threat degree, resource importance, and resource security attributes.*

Definition 4. *Attack threat (AT): indicates the damage caused to the target resource by the attacker's executed attack. The attack threat metric is shown in Table 2.*

C, I and A are usually used to represent the security attributes of resources, and different attacks cause different damage to the three indicators in the security attributes. Using (L_C, L_I, L_A) indicates the bias to the three indicators respectively, and $L_C + L_I + L_A = 1$, (1,0,0) means that the attack is implemented against the confidentiality of the security attributes.

Definition 5. *Resource Importance (RI): indicates the importance of the target node in the network, expressed in three levels: high, mid and low, and its quantitative criteria are shown in Table 3.*

Resource importance has different biases (R_C, R_I, R_A) for the three indicators of resource security attributes, and $R_C + R_I + R_A = 1$, whose quantitative criteria are shown in Table 4.

Combining the above table, it can be concluded that the calculation formula of attack revenue is Equation (10).

$$\text{AProfti}\,(e_{i\to j}) = RL\,(e_{i\to j}) \quad (10)$$

2.3.3 Attack Cost

Definition 6. *Atomic Attack Cost (ACost): The price that an attacker needs to pay for an atomic attack. Usually, the greater the impact on the network after the target node is compromised, the more critical the position of the target node in the network is considered, and the greater the possibility of an attacker launching an attack on the target node. The more likely it is to be discovered. The*

Table 1: CVSS metrics scoring system

Base	Influencing factors	Metric value	Score
Exp	AV	Network (N)	0.85
		Adjacent (A)	0.62
		Local (L)	0.55
		Physical (P)	0.20
Exp	AC	Low(L)	0.77
		High(H)	0.44
Exp	AU	No authentication required (N)	0.70
		Single authentication (S)	0.56
		Multiple authentication (M)	0.45
Impact	C, I, A	High(H)	0.56
		Low(L)	0.22
		None (N)	0.00
Scope	S	Unchanged (U)	
		Changed(C)	

Table 2: Attack threat metrics

Level	Attack classification	Threat level
L1	Information leakage	0.3
L2	Remote Login	0.5
L3	Obtain User privileges	0.8
L4	Obtain Root privileges	1.0

Table 3: Quantifying the importance of resources

Importance	Value	Resource description
high	1/2	Hosts where critical information is stored, such as Database Server
mid	1/3	Mail Server, FTP Server, Web Server
low	1/6	General hosts

criticality corresponding to the target node is calculated by Equation (11).

$$M\left(v_i\right)=\frac{\text{Impact}\left(v_j\right)}{\sum_{i=1}^{N}\text{Impact}\left(v_i\right)} \tag{11}$$

Among them, N represents the number of target nodes in the network.

The complexity ($\delta\left(e_{i\to j}\right)$) of each attack performed by the attacker is different, which also affects the cost of this attack. The quantification of the attack complexity is shown in Table 5.

The more times an attacker attacks a node, the more attacking experience the attacker has, and when he attacks the node again, the cost will be reduced accordingly.

Combined with the above data, the attack cost formula is given as Equation (12).

$$A\,\text{Cost}\left(e_{i\to j}\right)=\eta\left(e_{i\to j}\right)\times\delta\left(e_{i\to j}\right)\times M\left(e_{i\to f}\right) \tag{12}$$

2.3.4 Attack Probability

Combining the above definition, the probability of launching an attack on its child nodes from the current attribute node can be calculated when an attack occurs, which is referred to as the attack probability. It is expressed by $P(A_i)$, and its calculation formula is Equation (13).

$$P\left(A_j\right)=\min\left(\frac{\text{value}\ \left(v_i\right)*\ \text{Aprofit}\ \left(A_j\right)}{A\,\text{Cos}\,t\left(A_j\right)},1\right) \tag{13}$$

Definition 7. *Attack path selection probability $P(AP_i)$. To calculate the probability of different attack paths being compromised by attackers, the product of the state transfer probabilities of all nodes on a path AP_i is multiplied and its product is the attack path AP_i selection probability, denoted as $P(AP_i)$, which is calculated using Equation (14).*

$$P\left(AP_i\right)=\prod P_{ij} \tag{14}$$

3 Bayesian-based Network Situational Awareness Model

3.1 Bayesian Network Model Generation

Attack graph is a tool for modeling network security vulnerabilities and their interactions. It is widely used in various fields of network security, including risk assessment, because it can be used to describe the path of attackers exploiting vulnerabilities to damage network security. Each path consists of one or more vulnerabilities in a chain, some of which are prerequisites for others to be exploited.

To generate an attack graph for a given network, information about existing vulnerabilities, network topology, and host connectivity is required. On the basis of previous research results [8], this paper finds common vulnerabilities on hosts according to network vulnerability scanners (such as Nessus, OpenVAS or Retina) or online vulnerability repositories (such as the National Vulnerability Database (NVD) and MITRE's common vulnerabilities)

Table 4: Security attribute partial weight quantization

Application scenarios	Classification	Emphasis	Value (RC, RI, RA)
General hosts	R1	None	(1/3, 1/3, 1/3)
Information system	R2	Integrity	(1/4, 1/2, 1/4)
Data storage systems such as Database Server	R3	Confidentiality	(1/2, 1/4, 1/4)
Mail Server, Web Server	R4	Availability	(1/4, 1/4, 1/2)

Table 5: Attack complexity metrics

Attack complexity description	Value	Score
Code information (SI)	Complete	0.10
	Part	0.30
	None	0.70
Attack code platform (SP)	Normal	0.15
	Special	0.35
	Particular	0.60
Attack operation requirements (OR)	Tools	0.10
	Scripts	0.25
	Manuals	0.45
	Groups	0.70
Information collection requirements (IR)	None	0.00
	Normal	0.20
	Configuration	0.55
	Critical	0.80

searching existing loopholes. The host connectivity and topology of the network can be determined based on the knowledge of the network security administrator or using network tools such as Namp. With this information, tools such as MulVAL can be used to automatically generate attack graphs.

The proposed algorithm faces the problems of low data volume and high dimensionality. Bayesian network models require large data sets, and when the number of nodes increases, the complexity of the modeling of the node relationships performed increases, leading to difficulties in the interpretation and visualization of the model. Researchers need to synthesize large data sets using data augmentation techniques, reduce the number of features by feature selection for large amounts of data, and enhance the visualization of the graph by reducing the number of nodes and edges of the attack graph as much as possible through coefficient shrinkage and feature extraction techniques.

The Bayesian based situational awareness model can be generated by Algorithm 1.

Algorithm 1 Bayesian network model BNM generation algorithm

1: Input: Attack graph $AG = (S, A, E, \tau)$, P is the prior probability of the attribute node;
2: Onput: $BNM = (S, A, E, \tau, P)$;
3: Initialize each parameter of BNM to null;
4: Copy the individual parameters S, A, E, τ of AG;
5: **for** E_i in BNM **do**
6: Calculate the value of Exp using Equation (6).
7: **end for**
8: **for** Each S_j in BNM **do**
9: **if** $j = 1$ **then**
10: $P_1(S_1 = true) = p$
11: $P_1(S_1 = false) = 1 - p$
12: **else**
13: Calculate the prior probability P_j of the attribute node S_j using Equation (4), and fill in the parameters P of BNM;
14: **end if**
15: **end for**

3.2 Quantification of Security Posture

After generating the BNM according to Algorithm 1, the network is recalibrated by calculating the posterior probability.

Traditional network security prediction methods are

difficult to accurately identify new types of malicious attacks, and the data also has the problem of large scale and high latitude, which requires the use of more efficient algorithms for analysis and processing. Researchers establish adaptive network security models based on efficient algorithms such as random forests to improve the efficiency and accuracy of processing and analyzing data.

The BNM is a directed acyclic graph, and the data set $D = (d_1, d_2, \ldots, d_n)$ represents the observations of n variables $(x_1, x_2, \ldots, x_n)$, assuming that θ_{BNM} is the parameter value corresponding to the BNM . $P(BNM)$ denotes the prior knowledge of the nodes in the Bayesian network model, and when the model is initially completed, θ_{BNM} is denoted using $P(BNM_\theta)$. Its correction function is shown in Equation (15).

$$P(BNM, D) = \log_a P(D \mid BNM) + \log_a P(BNM) \quad (15)$$

$P(BNM)$ is generally assumed to be uniformly distributed. $P(D \mid BNM)$ is the marginal likelihood function, which can be expanded as Equation (16).

$$P(D \mid BNM) = \int P(D \mid BNM, \theta_{\mathrm{BNM}}) P\left(\theta_{\mathrm{BNM}} \mid BNM\right) d\theta_{\mathrm{BNM}} \quad (16)$$

After the above processing, the greedy search algorithm is selected and the search for a network structure that satisfies the requirements begins.

Each time a directed edge is selected from the model and added, the evaluation value is calculated using the above formula, and if the evaluation value becomes larger, the directed edge is added, otherwise it continues to the next step. In calculating the posterior probability, if the variables are not discrete, the probability density function $p(x \mid c) \sim N\left(\mu_{c,i}, \sigma_{c,i}^2\right)$ can be used, where $\mu_{c,i}$ and $\sigma_{c,i}^2$ are the mean and variance of the values taken by the class c samples on the attribute i respectively, the function can be given by Equation (17).

$$p(x \mid c) = \frac{1}{\sqrt{2\pi}\sigma_{c,i}} \exp\left(-\frac{(x_i - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) \quad (17)$$

The discreteization of continuous data in the sample makes multiple continuous intervals of data mapped into different categories, effectively reducing the difficulty of calculation, presenting the changing pattern of data more intuitively, conforming to the computer's processing method facilitating funny calculation and storage, better understanding and analysis of data, and improving the efficiency of security situational awareness algorithms.

Using Equation (18), the network situational values for each phase of the current network can be calculated and used for situational assessment and analysis.

$$V = \sum_{i=1}^{k} \frac{\mathrm{Aprofit}\left(e_i\right) - A\,\mathrm{Cost}\left(e_i\right)}{k} \quad (18)$$

Network situational awareness needs to be real-time, requiring the collection, processing, and analysis of large amounts of real-time data to discover and track events. The data obtained from Algorithm 1 is used to predict network behavior, make accurate judgments about the network situation, reduce errors and uncertainties, and monitor the security status of the system in real time.

Algorithm 2 Security posture quantification algorithm

1: Input: Security data obtained from Bayesian network security situational awareness model;
2: Onput: Network security posture value;
3: Step1: Combining the current network structure information and calling Algorithm 1 to obtain the Bayesian network model BNM;
4: Step2: For the continuous type data in the sample, it is discretized using Equation (15), Equation (16), Equation (17) and formed into a new data set with the non-discrete type data;
5: Step3: Using Equation (10) and Equation (12) to calculate the cost of attack and the benefit of attack for the current network;
6: Step4: The cyber security posture values for each phase can be calculated using Equation (18);
7: End

3.3 Risk Mitigation

Definition 8. *Security Countermeasure (SC) is a risk mitigation measure that can be implemented on a vulnerability to further reduce or eliminate residual risk by reducing the exploitability of the affected vulnerability, and in this way, prevent attackers from reaching their goal of compromising IT assets.*

The status of the SC_i defined as A Boolean variable $True, False$. $True$ indicates that the security countermeasure has been implemented, while $False$ indicates that the countermeasure has not been implemented as part of the security risk link plan.

Security countermeasures can be implemented proactively to mitigate the impact of known vulnerabilities on the network, by predicting in advance the set of vulnerabilities in the test network, the likelihood of attacks on these high-risk vulnerabilities and the potential attack paths. Then the network can be better prepared to prevent or mitigate any damage caused by these vulnerabilities.

The algorithm can detect network attacks through abnormal changes in network potential, carry out the detection of packets with abnormal characteristics in real-time to detect network intrusion and analyze the trend of changes in network potential. The node's vulnerability utilization and vulnerability value are combined with the real-time network state to conduct a comprehensive analysis of the attacker's attack behavior, determine the attack path so as to develop a defense method to reduce or stop the threat of harm to the network facilities.

4 Experimental Simulation and Numerical Analysis

4.1 Experimental Simulation

To explain the validity of the relevant models and algorithms proposed in this paper further, simulation experiments are conducted by deploying the experimental scenario shown in Figure 3. (1) DMZ Zone includes Mail Server, DNS Server and Web Server and is accessible to the public through fire-

Table 6: Description of the vulnerability information

Hosts	CVE ID	Vulnerability Information	Resource status number	Threat level	Resource importance	Security attribute categories
Mail Server	CVE-2014-7287	Mail Injection	S1	L4	mid	R4
	CVE-2008-3060	Error alert information leakage	S2	L1	low	R1
DNS Server	CVE 2008-1447	DNS cache poisoning	S3	L3	mid	R2
Web Server	CVE-2009-1535	IIS vulnerability in WebDAV service	S4	L3	low	R2
Gateway Server	CVE-2008-0166	OpenSSL uses predictable random numbers	S5	L1	low	R1
FTP Server	CVE-2014-1443	Buffer overflow	S6	L3	mid	R4
	CVE-2013-4465	Upload dangerous files	S7	L2	mid	R4
Administrative Server	CVE-2008-4050	Create and read arbitrary registry values	S8	L1	low	R2
Local Desktops	CVE-2015-8622	Cross-site scripting	S9	L2	low	R1
SQL Server	CVE-2018-12942	SQL Injection	S10	L4	high	R3
	CVE-2012-2592	Cross-site scripting	S11	L1	high	R3

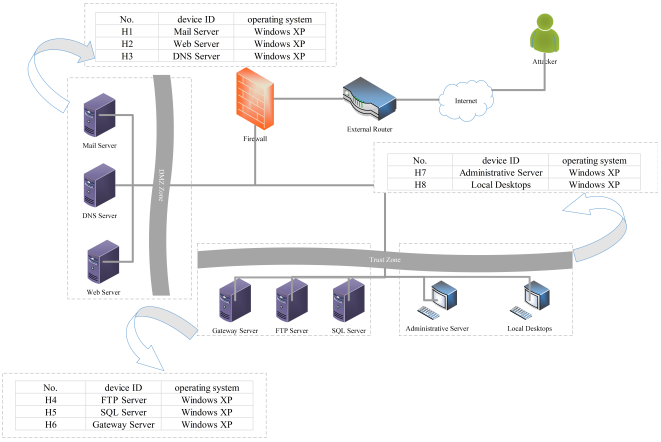


Figure 3: Experimental network topology

wall; (2) Trusted Zone includes SQL Server, Gateway Server, Administrative Server and some local computers, all access from external sources to the trusted zone and all communication with external parties is restricted. The DMZ tri-homed firewall has a pre-defined policy installed to ensure that it is separated from the trusted zone. According to the policy, the Web server is allowed to send SQL queries to the SQL server. In addition, remote desktop services for all local computers including the Administrative Server enable employees to communicate with remote sites via wired or wireless media. Remote connections are monitored by the SSHD protocol installed in the Gateway server.

For the test network, the Nessus vulnerability scanner was used to scan into different areas and the vulnerability information detected is listed in Table 6. The SQL database server has a critical role in most networks and can therefore be considered as a target for attackers.

4.2 Attack Probability Calculation

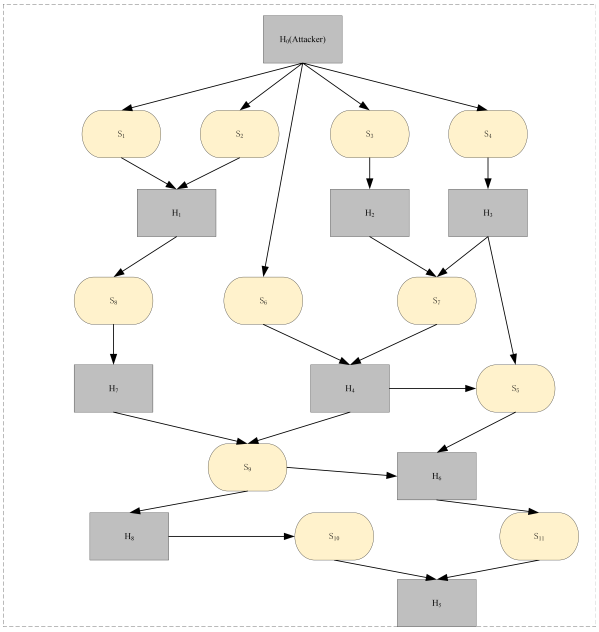


Figure 4: Schematic diagram of Bayesian network

In order to obtain the attack probability for each vulnerability, the attack cost and the attack benefit are calculated first. The attack cost is calculated according to Equation (12), the gain per attack is calculated using Equation (10), and then the probability of atomic attack can be obtained by combining Equation (13), and the calculation results are shown in Table 7. Then the network configuration, network connectivity information, vulnerabilities, and the correlation between vulnerabilities can be inputted together into the input.p file of Mulval tool to generate the attack graph model of the test network. As shown in Figure 4.

Table 7: Probability table

No.	Vulnerability value	Atomic attack probability (P(Aj))
S0		
S1	39%	0.52
S2	33%	0.42
S3	30%	0.39
S4	37%	0.47
S5	26%	0.34
S6	46%	0.66
S7	32%	0.41
S8	28%	0.38
S9	25%	0.30
S10	41%	0.58
S11	36%	0.45

Combining Table 7, we can get the probability of each vulnerability attack, and then combining with Equation (14) we can get the attack path selection probability. The obtained results are shown in Table 8.

Table 8: Attack path selection probability table

No.	Attack Path	Attack path selection probability / (P(APi))
AP1	$S_0 \to S_1 \to S_8 \to S_9 \to S_{10}$	0.0344
AP2	$S_0 \to S_1 \to S_8 \to S_9 \to S_{11}$	0.0267
AP3	$S_0 \to S_2 \to S_8 \to S_9 \to S_{10}$	0.0421
AP4	$S_0 \to S_2 \to S_8 \to S_9 \to S_{11}$	0.0215
AP5	$S_0 \to S_6 \to S_9 \to S_{10}$	0.1148
AP6	$S_0 \to S_6 \to S_9 \to S_{11}$	0.0891
AP7	$S_0 \to S_6 \to S_5 \to S_{10}$	0.1302
AP8	$S_0 \to S_6 \to S_5 \to S_{11}$	0.1010
AP9	$S_0 \to S_3 \to S_7 \to S_9 \to S_{10}$	0.0278
AP10	$S_0 \to S_3 \to S_7 \to S_9 \to S_{11}$	0.0216
AP11	$S_0 \to S_3 \to S_7 \to S_5 \to S_{10}$	0.0315
AP12	$S_0 \to S_3 \to S_7 \to S_5 \to S_{11}$	0.0245
AP13	$S_0 \to S_4 \to S_7 \to S_9 \to S_{10}$	0.0335
AP14	$S_0 \to S_4 \to S_7 \to S_9 \to S_{11}$	0.0260
AP15	$S_0 \to S_4 \to S_7 \to S_5 \to S_{10}$	0.0380
AP16	$S_0 \to S_4 \to S_7 \to S_5 \to S_{11}$	0.0295
AP17	$S_0 \to S_4 \to S_5 \to S_{10}$	0.0927
AP18	$S_0 \to S_4 \to S_5 \to S_{11}$	0.0719

Security countermeasures can be deployed in advance to make an impact on the vulnerabilities, reduce the probability of vulnerability attacks, and play a role in protecting network security after predicting the vulnerabilities with high attack probability and their attack paths. The proposed security countermeasures for the test network are shown in Table 9.

After the processing of security countermeasures, the network is tested again and the network posture values are recalculated using Algorithm 2.

4.3 Numerical Analysis

After analyzing Table 8, we found that the attack paths $AP_5, AP_6, AP_6, AP_7, AP_{17}, AP_{18}$ have higher intrusion probability, and attackers are more likely to choose these paths to attack the system, especially the probability of attack path AP_7 reaches 0.1302. At this time, the attacker can cause extremely

Table 9: Security measures and their coverage

Security Countermeasures (SC)	Coverage	Coverage Ratio
Disable WebDav	CVE-2009-1535	0.61
Filtering External Traffics	CVE-2008-0166	0.55
Query Restriction	CVE-2018-12942	0.28
	CVE-2012-2592	
Add Network IDS	CVE-2014-7287	0.37
	CVE-2008-3060	
Limit Access To DNS Server	CVE 2008-1447	0.50

serious damage to the current network, but the network risk changes when we predict these attacks in advance through network situational awareness and use preventive measures in advance then the probability of attack for each path is shown in Figure 5.

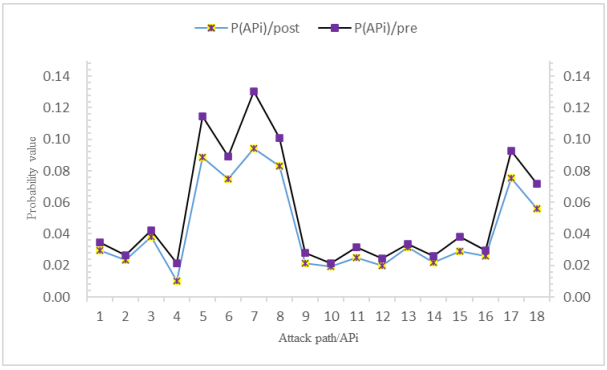


Figure 5: Comparison of attack path probability before and after situational awareness

It is obvious from the above figure that when we get the attacker's attack type in advance and take precautions, the attack probability of each path changes. At the same time, the attacker's attack cost per attack will increase and the attack benefit decreases. At this point, we use Algorithm 2 to calculate the change of network security posture values over time in both states, and generate the results of the time-series-based posture index assessment in days and compare them. The results are shown in Figure 6.

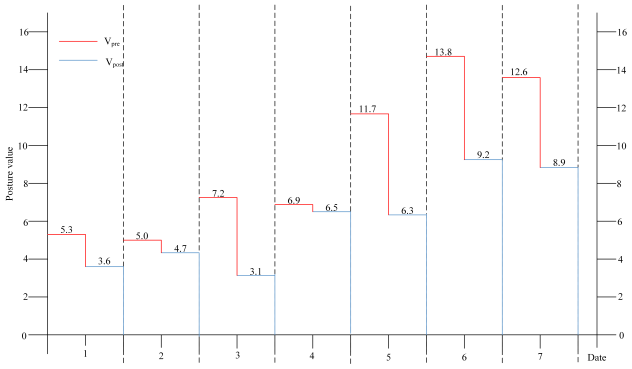


Figure 6: Comparison of cyber security posture values

By comparing the posture values in Figure 6, we can clearly see that the changes in the network posture on days 5, 6, and 7 are larger, while the rest of the time is more moderate. When we perceive the current network's posture in advance and then

measure the posture value, we can compare and find that the posture value in the test network has changed and the security level of the current network has been improved significantly after the early response. Therefore, we can conclude that not only does he target of the attacker not eliminate the risk posed by the attacker, but also can be analyzed the source of network risk more accurately and be responded when it is clearly identified, that the network level is always kept in an acceptable range.

4.4 Comparison of Methods

Different models have different degrees of protection for the network. In order to verify the superiority of the model in this paper, for different models, their respective degrees of protection for the network are compared in the same network environment, and the specific procedure is as follows.

A comparison of the data obtained by the method proposed in literature [16] and literature [18] respectively with the experimental method in this paper is given in the topological environment shown in Figure 3.

The literature [16] is also a network situational awareness model that uses Libsvm and adaptive weighting strategy to train the collected samples by designing the network data feature terms and indicator values, and collects some network data for prediction analysis; while the literature [18] uses Bayesian theory to calculate the reachable probability of each node, describe the probability of single-step attack occurrence, and dynamically predict the potential risk in the network. Combined with the attack graph, an improved attack graph-based intrusion prediction algorithm is proposed to simplify the connection between alert evidence and attack behavior and improve the accuracy of prediction. However, the former does not take into account the cost and benefit of vulnerability attacks, and the latter has too single vulnerability indicators, so both do not truly reflect the real situation of vulnerabilities exploited in the network. Figure 7 gives a comparison of the posture values of the three models within a week in the same network environment. Figure 8 then gives the overall defense level of the network during these seven days.

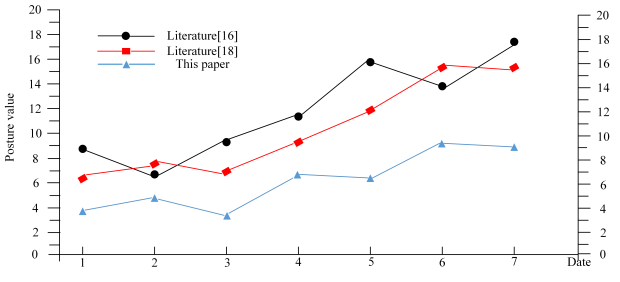


Figure 7: Comparison of security posture values

The analysis of Figure 7 shows that for the three models, the network potential values are higher on the 5th, 6th, and 7th days of the week, but the model in this paper always maintains a more stable state during the seven days, compared with the other two models, although there are still fluctuations, but the fluctuations are smaller and always lower than the other two models.

Analysis of Figure 8 shows that the overall defense level of all three models is relatively high, with the minimum defense

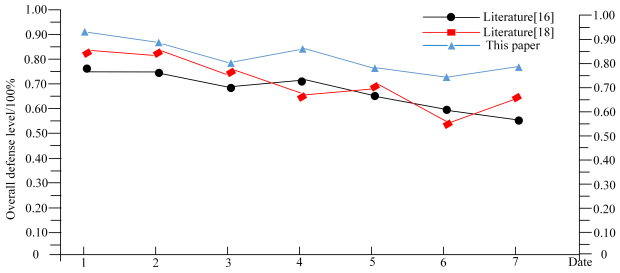


Figure 8: Overall Defense Level Comparison

level above 55%, but the model in this paper is significantly better than the models in the other two papers. The overall defense level of this paper is above 70%, and the reason for this is that this paper has pre-knowledge of the attacker's target, and the probability of the attack, and has made deployment of these attacks in the network in advance, so predicting the attacker's target in advance can keep the network defense level at a more secure level.

5 Conclusion

Bayesian security situational awareness algorithm has a wide range of applications in network monitoring and management system, through the analysis of network traffic, protocols, attack methods, etc. to quickly detect intrusions and update the situational information at any time, quickly check out abnormal situations such as DDos attacks or discover the network vulnerabilities and other security risks so as to take corresponding measures, can calculate the host device weights to optimize the network topology and extend the network lifecycle. In the actual network monitoring and management system to apply the current algorithm, we should first collect the security elements of the network security posture through the system log information, equipment protection tools, network monitoring equipment and routers, etc., process the data and then make corrections and optimization, and finally pay attention to the maintenance and upgrading of the algorithm to ensure the accuracy and reliability of the algorithm.

It is a considerable challenge to analyze and predict the network posture of a given network environment so that the network security level is always kept in an acceptable range. In this paper, a Bayesian network-based network security situational awareness method is proposed to address this problem. By obtaining the situational indicators affecting the network environment from system logs and protection tools, calculating the vulnerability attack probability as well as the attack path selection probability using indicators such as vulnerability value, attack cost, and attack benefit, processing the vulnerabilities with higher attack probability as well as the attack path to reduce the degree of network victimization, and calculating the security situational values of the network before and after processing by the security situational quantification algorithm for comparison; finally, the superiority of the method in this paper is verified by comparing with the methods in other literature.

The algorithm application scenario is very broad, based on historical data and real-time data to predict the future network security threats of the current environment, timely detection of network attacks and abnormal activities, and then

take security strategies. It can be expected that with the development of Bayesian network security posture-based algorithms will continue to develop and improve, for example, in the future will be combined with big data analysis, machine learning, etc, and continue to innovate algorithms to improve network security.

Acknowledgments

This study was supported by Heilongjiang Provincial Natural Science Foundation of China:LH2021F030. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] N. Gupta, V. Jindal, P. Bedi, "A survey on intrusion detection and prevention systems," *SN Computer Science*, vol. 4, no. 5, 2023.

[2] H. Hu, H. Zhang, and Y. Liu, "Quantitative method for network security situation based on attack prediction," *Security and Communication Networks*, p. 2017, 2017.

[3] L. Hu, L. Q. Zhou, J. Deng, R. Li, Z. W. Zhao, "Evaluation model of network security situation based on support vector machine and self-adaptive weight," *Computer Systems and Applications*, vol. 27, no. 7, pp. 188-192, 2018.

[4] M. Husák and J. Komárková and E. Bou-Harb, "Survey of attack projection, prediction, and forecasting in cyber security," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 640-660, 2018.

[5] G. Kou, S. Wang, and G. Tang, "Research on key technologies of network security situational awareness for attack tracking prediction," *Chinese Journal of Electronics*, vol. 28, no. 1, pp. 162–171, 2019.

[6] C. Lei, D. H. Ma, and H. Q. Zhang, "Performance assessment ap- proach based on change-point detection for network moving target de-fense," *Journal on Communications*, vol. 38, no. 1, pp. 126–140, 2017.

[7] A. López-Cuevas and M. A. Medina-Pérez and R. Monroy, "A visualisation approach for real-time risk situation awareness," *IEEE Transactions on Affective Computing*, vol. 9, no. 3, pp. 372-382, 2017.

[8] Z. Luo, X. Yang, J. Liu, and R. Xu, "Network intrusion intention analysis model based on bayesian attack graph," *Journal on Communications*, vol. 41, no. 9, pp. 160–169, 2020.

[9] Y. Nikoloudakis, I. Kefaloukos, and S. Klados, "Towards a machine learning based situational awareness framework for cybersecurity: An sdn implementation," *Sensors*, vol. 21, no. 14, p. 4939, 2021.

[10] X. Pang, Y. Su, W. Tao, Z. Hu, and H. Chen, "A network security situational awareness system for text classification," *Journal of Physics: Conference Series*, vol. 2358, no. 1, 2022.

[11] X. Sun, J. Dai, and P. Liu, "Using bayesian networks for probabilistic identification of zero-day attack paths," *IEEE Transactions on Information Forensics and Security*, no. 1-1, 2018.

[12] X. Tao, Z. Liu, and C. Yang, "An efficient network security situation assessment method based on ae and pmu," *Wireless Communications and Mobile Computing*, p. 2021, 2021.

[13] H. Wang, S. K. Lu, and Y. C. Wang, "Intrusion prediction algorithm based on correlation attack graph," *Computer Engineering*, vol. 9, no. 3, pp. 79–89, 2017.

[14] H. Zhang, C. Kang, and Y. Xiao, "Research on network security situation awareness based on the lstm-dt model," *Sensors*, vol. 21, no. 14, p. 4788, 2021.

[15] H. Zhang, K. Kang, and W. Bai, "Hierarchical network security situation awareness data fusion method in cloud computing environment," vol. 23, no. 1, pp. 237–251, 2023.

[16] J. Zhang, H. Feng, B. Liu, and D. Zhao, "Survey of technology in network security situation awareness," *Sensors*, vol. 23, no. 5, pp. 2608–2608, 2023.

[17] D. Zhao and J. Liu, "Study on network security situation awareness based on particle swarm optimization algorithm," *Computers and Industrial Engineering*, vol. 125, pp. 764–775, 2018.

[18] Y. Zhu and Z. Du, "Research on the key technologies of network security-oriented situation prediction," *Scientific Programming*, p. 2021, 2021.

Biography

Luo Zhiyong biography.Luo Zhiyong, male, was born in Shandong, China in July 1978.He is a professor at the School of Computer Science and Technology, Harbin University of Science and Technology. Research direction: computer network and information security, network optimization.

Wang Shuyi biography. Wang Shuyi, female, was born in Heilongjiang, China in June 2000. She is a postgraduate student in the School of Computer Science and Technology, Harbin University of Science and Technology. Research direction: network security.

Xu Haifeng biography.Xu Haifeng, male, was born in March 1999 in Jilin , China. He is a postgraduate student in the School of Computer Science and Technology, Harbin University of Science and Technology, Research direction: network security.

Song Weiwei biography.Song Weiwei, male, was born in May 1998 in Shanxi , China. He is a postgraduate student in the School of Computer Science and Technology,Harbin University of Science and Technology, Research direction: network security

Blockchain Technology: The Prevention of Corporate Financial Statement Forgery and Frau

Ke Li, Tianwen Chen, and Weimin Yang
(Corresponding author: Weimin Yang)

School of Economics and Trade Management, Yancheng Polytechnic College
Jiangsu Vocational Institute of Commerce
No. 180, Longmian Avenue, Jiangning District, Nanjing, Jiangsu 210000, China
Email: ya67431@163.com

(Received Sept. 4, 2022; Revised and Accepted July 28, 2023; First Online Oct. 8, 2023)

Abstract

Preventing financial statement forgery and fraud is essential for maintaining the stability of the economic market. In this paper, financial statements and fraud were introduced briefly. Then, blockchain technology was used to prevent financial statement forgery and fraud. Furthermore, simulation experiments were carried out within a laboratory environment to evaluate the efficacy of the proposed preventive measures. The experimental results demonstrated that conventional financial statements could be successfully chained, whereas anomalous financial statements were effectively identified and prevented from being chained. As the quantity of financial data slated for chaining increased after adopting the preventive measures, the average server throughput increased first and then stabilized. Simultaneously, the average latency initially remained consistent before experiencing a subsequent increase. Incorporating blockchain technology within these preventive measures effectively thwarts any attempts at financial data compromise or tampering.

Keywords: Blockchain; Financial Statement; Forgery and Fraud; Preventive Measure

1 Introduction

Owing to various interest-driven motives and malpractices, instances of forgery and fraud are recurrent in financial statements [15]. The presence of deceptive financial statements not only undermines the credibility and reputation of enterprises but also exerts a profound impact on the overall market stability. Consequently, the effective prevention of financial statement forgery and fraud has emerged as an urgent imperative for both enterprises and regulatory bodies [19]. Through the practice of recording financial report data on a blockchain and facilitating real-time sharing and verification, it is feasible to enhance data transparency and credibility, thereby diminishing the likelihood of data tampering [3, 16].

Liu *et al.* [14] constructed a heterogeneous graph transformation network designed for detecting anomalies in smart contracts, aiming to identify financial fraud on the Ether platform. Hyvarinen *et al.* [7] devised and assessed a prototype blockchain-based system, aiming to eliminate instances of double-taxation-related tax fraud and concurrently enhance dividend flow transparency. Meanwhile, Kabra *et al.* [8] proposed MudraChain, a framework aimed at automating check clearing operations, in which a blockchain network orchestrates clearing processes. The obtained results were subjected to rigorous evaluation using state-of-the-art methodologies to highlight the efficacy of the proposed framework.

This paper briefly introduced financial statements and financial statement fraud. Then, blockchain technique was employed to prevent financial statement forgery and fraud. Finally, the preventive measures were tested using simulation experiments in a laboratory.

2 Financial Statements and Fraud

Financial statements can help stakeholders comprehensively understand an enterprise's financial status through a series of financial data [18]. The balance sheet serves as a reflection of an enterprise's assets, liabilities, and owner's equity during a specific time frame, which unveils the asset composition, liability extent, and net asset position of the enterprise and offers stakeholders a benchmark for assessing solvency and risk tolerance [21]. Meanwhile, the income statement records an enterprise's sales revenue, costs, and profits within the corresponding time frame. It illustrates the operational performance and profitability of an enterprise, providing a foundation for evaluating the sustainability of the enterprise's business mode. Additionally, the cash flow statement scrutinizes

the inflow and outflow of cash for an enterprise during the corresponding period, which unveils an enterprise's cash liquidity, along with the cash-related income and expenditures stemming from operational activities [11].

It enables stakeholders to gauge the enterprise's capacity for cash management and solvency. Undoubtedly, financial statements constitute the bedrock for decision-making among investors, creditors, and regulatory bodies. Furthermore, they serve as a crucial reference point for an enterprise's internal business analysis and management strategies. Consequently, any incidence of fraudulent financial statements not only impacts the enterprise directly but also negatively affects stakeholders and the market [20]. Fraudulent activities within financial reporting encompass a spectrum of tactics, including but not limited to, inflating income, concealing liabilities, and overstating assets. The reasons for fraudulent behavior are also diverse, including top executives manipulating financial data for personal gain, companies falsifying information to enhance their investment appeal due to market competition pressure, and inadequate regulation or internal control within the company as well as negligence on the part of auditing agencies providing opportunities for fraudulent activities. In the face of the challenge posed by financial reporting fraud, extant preventive measures exhibit certain limitations, such as insufficiency in the independence and professional competence of auditing organizations, internal control defects, and weak supervision force [12].

3 Blockchain-based Financial Statement Fraud Prevention Measures

Blockchain technology, a decentralized ledger system, interconnects transactional data through blocks and employs cryptographic algorithms for encryption and verification, achieving a decentralized trust mechanism [1, 2, 4, 13, 17]. The inherent attributes of blockchain, notably its resistance to tampering, transparency, and decentralization, neatly align with the requisites for countering fraudulent activities in financial reporting [5]. Its resistance to tampering ensures the validity of financial data stored in the blockchain. Simultaneously, the transparency inherent to blockchain empowers anyone to peruse and authenticate financial data. Furthermore, the decentralized nature of blockchain technology facilitates the dispersed storage of financial data, thereby curbing the likelihood of manipulation and forgery during auditing processes.

Figure 1 shows the blockchain-based financial statement monitoring process, and the specific steps are shown below.

1) The business department sends the financial data generated by the enterprise in conducting business activities to the supervisory authority in the form of a digitally signed document.

2) The supervisory authority checks the signed documents and returns them to the business department if they do not pass the check [10].

3) If the financial data passes the verification, it will be broadcasted to all nodes in the blockchain network, and then each node checks the financial data through the consensus algorithm. When the majority of nodes pass the verification, that is to say, after the consensus is reached, the financial data will be deposited into the newly generated block and backed up in other nodes. The consensus algorithm used in this paper is the practical byzantine fault tolerance (PBFT) algorithm [9], which will first select a master node in the blockchain network when it carries out consensus, and the rest are slave nodes. The formula of master node selection is:

$$\begin{cases} p = v \bmod |n| \\ \theta_1 = \alpha \cdot \theta \\ \theta_{GC} = \begin{cases} \frac{\theta}{n} - c, & \text{node cooperation} \\ \frac{\theta_1}{n}, & \text{node non-cooperation} \end{cases} \end{cases} \tag{1}$$

where p is the master node number, v is the view number, $|n|$ is the total number of nodes in the view, θ is the total service fee paid by the user to the blockchain network, α is the coefficient used for revenue lure, θ_1 is the service fee used for revenue lure, n is the number of nodes in the blockchain, c is the cost of cooperation for each node, θ_{GC} is the incentive value that each node can obtain after the consensus is successful.

After electing the master node, the master node verifies the financial data and broadcasts it to the whole network after passing, and the rest of the nodes also verify the financial data and feedback the verification results to the other nodes in the blockchain. When more than two thirds of the nodes pass the verification, then the consensus is successful, and each node deposits the uploaded financial data into the local blockchain ledger, completing the safe chaining of the financial data; otherwise, it returns to the business department to re-upload.

4 Experimental Analysis

4.1 Experimental Environment

Simulation experiments were carried out using servers in a laboratory environment to evaluate the efficacy of blockchain-based monitoring mechanisms for financial statements. The requisite blockchain network was provided by the virtual machines on the Ethernet platform [6]. For the sake of simulation, parameter settings were unified, specifically employing a single-core i5 CPU, an operating frequency of 2.5 GHz, and 4 G of memory. Ten nodes were provided by the virtual machines. The servers deployed within the laboratory to serve as both

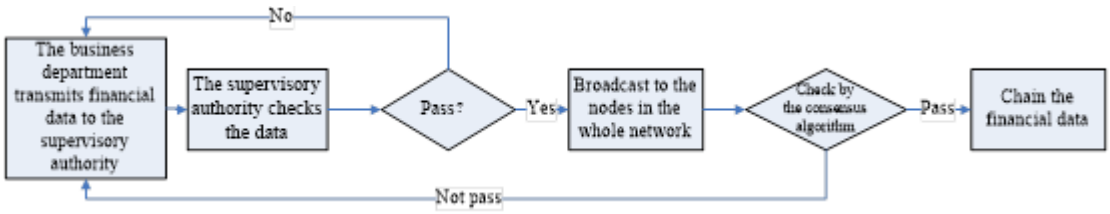


Figure 1: Blockchain-based financial statement monitoring process

the business and regulatory departments were equipped with quad-core i7 CPU, 16 G of memory, and 1,024 G hard disks.

4.2 Experimental Projects

1) Functional testing of the uplinking of financial statements
Two financial reports were entered into the server that acts as the business department, one being a normal financial report and the other an abnormal financial report. Some of the content of the two financial reports are shown in Table 1, in which the abnormal financial report was changed from the normal financial report. Except for the order number, only the amount of the order was changed. The two financial reports were uploaded to the server of the business department and applied for chaining, after which the two financial reports were queried by the order number.

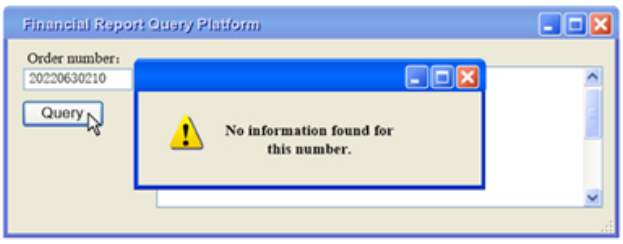
2) Chaining efficiency test of financial statements
The financial data size was set to 500 MB, 1,000 MB, 1,500 MB, 2,000 MB, and 2,500 MB, respectively, to test the average throughput and delay of the financial statement chaining under different financial data sizes.

3) Security testing of financial statements after chaining
Two aspects were considered when testing the security. On the one hand, a third-party server was used to play the role of an unfamiliar user to crack the encrypted summary information of the chained financial data with brute force, and the size of the financial data corresponding to the summary information to cracked was 5 MB, 10 MB, 15 MB, 20 MB, and 25 MB. The time spent in cracking was set to 30 minutes. On the other hand, the chained financial data was modified in the business department server to simulate the situation of tampering with the financial data after the server is successfully attacked, and then the modified financial report was queried.

4.3 Experimental Results

After the normal and abnormal financial reports were uploaded in the server of the business department and applied for chaining, they were queried according to their numbers, and the query results are shown in Figure 2. For normal financial reports, the detailed content of the order can be queried, while for abnormal financial reports, the detailed content cannot be queried, and only a feedback of "no information of this number" can be obtained, which indicates that the abnormal financial reports have not been successfully chained.

(a) The query result of a normal financial report



(b) The query result of an abnormal financial report

Figure 2: Query results of normal and abnormal financial reports after chaining operation

The chaining efficiency of financial reports under different financial data sizes is shown in Figure 3. The average throughput of the whole chaining process gradually rose with the increase of financial data and then remained stable, while the average delay first remained stable and then rose. The reason for such changes is as follows. The increase in the amount of financial data made the average throughput to increase before reaching the upper limit of the broadband, and the average delay remained unchanged. After the upper limit was reached, the accumulation of financial data occurred in the transmission, leading to an increase in the average delay.

The extent of the third-party server's cracking of summary information under different financial data sizes is shown in Figure 4. It can be seen that the complete-

Table 1: Selected elements of normal and abnormal financial reporting

	Normal financial reporting	Abnormal financial reporting
Order number	20220630211	20220630210
Order address	No. xx, xx District, Nanjing, Jiangsu Province, China	No. xx, xx District, Nanjing, Jiangsu Province, China
Order content	100 bags of wheat flour	100 bags of wheat flour
Order amount	4,290 yuan	4,310 yuan

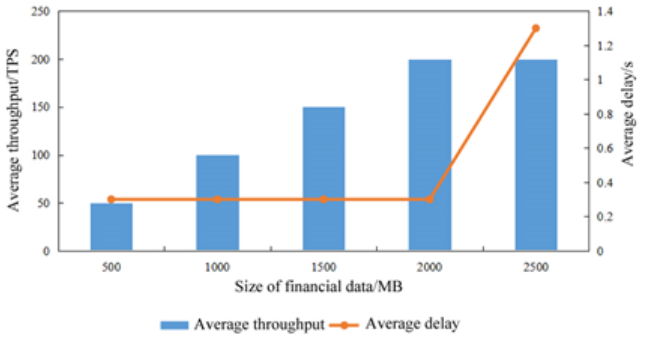


Figure 3: Chaining efficiency of financial statements under different financial data sizes

ness of even 5 MB summary information after brute-force cracking was only 15.3%, and the completeness of brute-force cracking decreased with the increase of the financial data size.

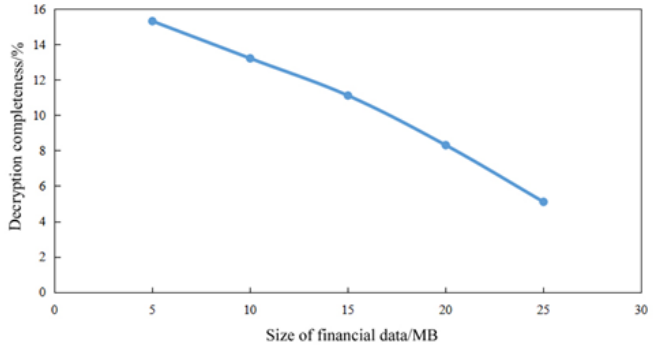


Figure 4: Extent to which the third-party servers cracks summary information under different financial data sizes

Taking the normal financial report in Table 1 as an example, the amount in the report was modified in the server of the business department, after which this financial report was queried in the blockchain platform. The modified financial data and the query results are shown in Figure 5. As can be seen from Figure 5, the order amount was changed from the original "4290" to "4112", but the query result from the blockchain platform was consistent with the original financial data.

5 Conclusion

The article provides a brief introduction to financial statements and fraudulent activities, proposes the use of blockchain technology for preventing financial statement fraud, and concludes with simulation experiments conducted in the laboratory to test the proposed preventive measures. The outcomes of these experiments are as follows.

1) Normal financial statements could be chained and queried normally, whereas abnormal financial statements could not be queried, indicating their inability to be chained.

2) In terms of the preventive measures, as the volume of financial data to be chained increased, the average server throughput demonstrated an initial increase followed by stability, and the average delay remained stable initially and subsequently increased.

3) The level of brute-force decryption attempts on financial data summary information by the third-party server diminished with the increase of the data size.

4) The financial statements could be queried in the blockchain platform even after financial data modifications were made within the business department's server.

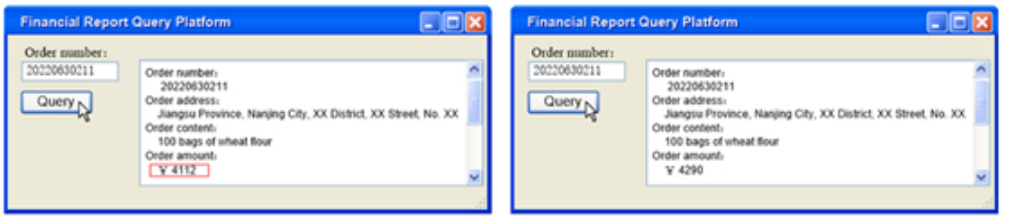
References

[1] P. Y. Chang, M.S. Hwang, C.C. Yang, "A blockchain-based traceable certification system", in *Security with Intelligent Computing and Big-data Services*, pp. 363-369, 2018.

[2] L. Chen, Q. Fu, Y. Mu, L. Zeng, F. Rezaeibagha, M. S. Hwang, "Blockchain-based random auditor committee for integrity verification", *Future Generation Computer Systems*, vol. 131, pp. 183-193, 2022.

[3] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on the secure financial surveillance blockchain systems", *International Journal of Network Security*, vol. 22, no. 4, pp. 708-716, 2020.

[4] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on blockchain technologies in bidding systems", *International Journal of Network Security*, vol. 22, no. 6, pp. 897-904, 2020.



(1) Financial data after modifiction (2) The query result of financial data after modifcation

Figure 5: The modification to the financial data and the query result after the modification

[5] P. Dangayach, "Pharmaceutical supply chain tracking system based on blockchain technology and radio frequency identification tags," *International Journal of Business Research*, vol. 19, no. 4, pp. 37-44, 2019.

[6] H. G. Driver, T. Hartley, E. M. Price, *et al.*, "Genomics4RD: An integrated platform to share Canadian deep-phenotype and multiomic data for international rare disease gene discovery," *Human Mutation*, vol. 43, no. 6, pp. 800-811, 2022.

[7] H. Hyvarinen, M. Risius, G. Friis, "A blockchain-based approach towards overcoming financial fraud in public sector services," *Wirtschaftsinformatik*, vol. 59, no. 6, pp. 441-456, 2017.

[8] N. Kabra, P. Bhattacharya, S. Tanwar, S. Tyagi, "MudraChain : Blockchain-based framework for automated cheque clearance in financial institutions," *Future Generation Computer Systems*, vol. 102, pp. 574-587, 2020.

[9] S. S. Kamble, A. Gunasekaran, M. Goswami, J. Manda, "A systematic perspective on the applications of big data analytics in healthcare management," *International Journal of Healthcare Management*, vol. 12, no. 3, pp. 226-240, 2019.

[10] L. Koh, A. Dolgui, J. Sarkis, "Blockchain in transport and logistics – paradigms and transitions," *International Journal of Production Research*, vol. 58, no. 7, pp. 2054-2062, 2020.

[11] L. Kong, "A fast encryption method for enterprise financial data based on blockchain," *Web Intelligence and Agent Systems*, vol. 20, no. 2, pp. 133-141, 2022.

[12] M. Li, S. Shao, Q. Ye, G. Xu, G. Q. Huang, "Blockchain-enabled logistics finance execution platform for capital-constrained E-commerce retail," *Robotics and Computer-Integrated Manufacturing*, vol. 65, pp. 1-14, 2020.

[13] C. Y. Lin, L. C. Huang, Y. H. Chen, M. S. Hwang, "Research on security and performance of blockchain with innovation architecture technology", *International Journal of Network Security*, vol. 23, no. 1, pp. 1-8, 2021.

[14] L. Liu, W. T. Tsai, M. Z. A. Bhuiyan, H. Peng, M. Liu, "Blockchain-enabled fraud discovery through abnormal smart contract detection on Ethereum," *Future Generations Computer Systems*, vol. 128, pp. 158-166, 2022.

[15] D. E. O'Leary, "Open information enterprise transactions: business intelligence and wash and spoof transactions in blockchain and social commerce," *International Journal of Intelligent Systems in Accounting, Finance & Management*, vol. 25, no. 3, pp. 148-158, 2018.

[16] M. T. Oladejo, L. Jack, "Fraud prevention and detection in a blockchain technology environment: Challenges posed to forensic accountants," *International Journal of Economics and Accounting*, vol. 9, no. 4, pp. 315-335, 2020.

[17] Z. P. Peng, M. L. Chiang, I. C. Lin, C. C. Yang, M. S. Hwang, "CBP2P: Cooperative electronic bank payment systems based on blockchain technology", *Journal of Internet Technology*, vol. 23, no. 4, pp. 683-692, 2022.

[18] F. R. da Silva, R. V. G. Teixeira, "The use of the blockchain protocol by public administration as an accomplishment of efficiency in the public service," *Journal of Public Administration and Governance*, vol. 8, no. 3, pp. 333-343, 2018.

[19] P. Treleaven, R. G. Brown, D. Yang, "Blockchain technology in finance," *Computer*, vol. 50, no. 9, pp. 14-17, 2017.

[20] H. Wu, N. Su, C. Ma, P. Liao, D. Li, "A privacy protection solution based on NLPCA for blockchain supply chain financial system," *International Journal of Financial Engineering*, vol. 07, no. 3, pp. 1-22, 2020.

[21] C. D. Zhu, "Fraud detections for online businesses: A perspective from blockchain technology," *Financial Innovation*, vol. 2, no. 1, pp. 256-265, 2016.

Biography

Ke Li, born in 1993, graduated from University of Reading in December 2017 with a master degree.He is pursuing a Ph.D in Accounting and Finance at INTI International University. He is an associate researcher and is now working in Yancheng Polytechnic College, China. He is interested in blockchain technology,big data and financial management.

Tianwen Chen, born in 1992, graduated from Shanghai University in June 2015 with a master degree. She is an

associate researcher and is now working in Yancheng Polytechnic College, China. She is interested in blockchain technology,big data and supply chain management.

Weimin Yang, born in 1989, graduated from KyungHee University in February 2018 with a Ph.D degree. He is an associate researcher and is now working in Jiangsu Vocational Institute of Commerce, China. He is interested in blockchain technology,big data and electronic commerce.

A Fine-Grained and Dynamic Access Control Model for Smart Home in Cloud Environment

Pengshou Xie, Pengyun Zhang, Tao Feng, Minghu Zhang, Xiaoye Li, and Linge Qi
(Corresponding author: Pengyun Zhang)

School of Computer and Communications, Lanzhou University of Technology
No. 36 Peng Jia-ping road, Lanzhou, Gansu 730050, China
Email: 2324327226@qq.com

(Received May 19, 2023; Revised and Accepted Sept. 22, 2023; First Online Oct. 8, 2023)

Abstract

Numerous smart home cloud management platforms use the role-based access control model to manage users' permissions, which leads to the problem of coarse granularity and insufficient dynamics of permission assignment. To this end, an attribute-based fine-grained dynamic access control model is proposed in this paper, which focuses on the scenario in which users use smart terminals to control smart home devices remotely through the "cloud." Considering that the user's identity, role, historical behavior, and request environment are constantly changing, and users face different risks when controlling different devices, the corresponding trust and risk value calculation methods are proposed in this paper, which are used for real-time evaluation of user control requests. Based on the evaluation results, the request credibility is calculated, and the user is dynamically authorized to control the device with the corresponding security level based on the request credibility. Further, attribute-based encryption is used to implement fine-grained permission assignment. The simulation results show that the proposed model achieves fine-grained and dynamic access control with high efficiency and security.

Keywords:Access Control; Cloud Computing; Dynamicity; Fine-Grained; Smart Home

1 Introduction

The rapid and continuous innovation in information technology has accelerated the development of the Internet of Things (IoT) and cloud computing. A significant volume of data produced by IoT is stored and processed via cloud computing. Smart homes are among the most prevalent IoT areas, with more and more smart home devices connecting to the Internet through the cloud [10]. Users can use smartphones, tablets, or smart speakers to remotely control home devices through the cloud platform, which increases the flexibility and convenience of how users control devices but also raises a series of security issues [2,20]. Such as the problem of secure sharing of device data in the cloud, the problem of fine-grained assignment of device permissions among different home users, and the security problem caused by the increasing number of cyberattacks against smart home devices. Access control is a security mechanism that determines whether users can access resources or services within the system.It can effectively address the issues mentioned above [14].

However, the access control problem in smart homes is significantly different from the traditional domain. First, the same device may be used by different users, e.g., smart locks. Second, there are often complex social relationships between occupants, which introduces a new pattern of threats, such as an annoying child trying to control the smart light in a sibling's room or a current or former partner trying to abuse one or all occupants. Another key feature of smart home devices is that most lack screens and keypads, making it easy for users to control the devices using their hands and making access control more challenging. These features suggest that, with restricted users and resources, traditional access control mechanisms are unsuitable for developing IoT. A fine-grained dynamic access control model is urgently required for smart home systems in cloud environments [6].

To this end, a fine-grained dynamic access control model for smart homes is proposed in this paper, which focuses on the security problems that exist in the scenario where users use smart terminals such as cell phones to control smart home devices remotely through the smart home cloud platform. The proposed model is based on the attribute-based access control (ABAC) model, which dynamically authorizes permissions for users by calculating the request credibility and achieves fine-grained permission assignment by using ciphertext policy attribute-based encryption technology.

This paper is organized as follows: in Section 2, the related work of other researchers is presented. In Section 3, the details of how the fine-grained and dynamic access control is implemented in this paper are described. In Section 4, the experimental scheme and the analysis of

the results are given. Section 5 concludes the paper.

2 Related Works

A smart home system in a cloud environment includes physical devices and Internet connections; these physical devices connected to the cloud connect and communicate with each other through the Internet and provide intelligent services to users based on predefined policies or user control commands [17]. With the help of various smart terminals and mobile communication technologies, users can easily control and monitor these home devices, such as controlling lighting, checking indoor temperature or air quality, and controlling the closing of doors and windows. However, due to the complexity and variability of the smart home environment and the uncertainty of the environment in which users send control commands, smart home systems in the cloud environment also face many security issues. For example, illegal users impersonate users to control smart home devices or view private data stored in the cloud through users' missing smart terminals. Since smart home devices are closely related to users' lives, these security issues may threaten users' life and property safety [11]. For this reason, many researchers have studied the problem of smart home access control.

In literature [8], Dutta *et al.* proposed an access control model to capture the physical context of a device through intrusion detection techniques to change the device policy and implement access decisions dynamically. In [19], Sikder *et al.* address the problem of conflicting user requirements in a shared smart home system with multiple users and devices in an automated and configurable approach based on the different priorities of the real user requirements.

In [23], Zhang *et al.* integrate social attributes into game theory and dynamically adjust access control policies to achieve adaptive fine-grained division of access control under the social Internet of Things. In [7], Dong *et al.* propose a multi-attribute decision-based access control model. This model starts with a risk assessment of multi-attribute factors such as environment, resources, and tasks in access control so that the access rights of users can be dynamically adjusted.

In literature [21], Wang *et al.* proposed an access control scheme for attribute-based encryption based on access structure tree pruning, and the scheme reduces the computational overhead in the encryption and decryption process of the attribute authorization center.

In literature [22], Xie *et al.* proposed an adaptive access control model based on XACML, which can satisfy various access control conditions and dynamically and adaptively adjust the access control policy, but the model is only suitable for the Internet of Vehicles environment.

In [5], Ameer *et al.* proposed two hybrid models $HyBAC_{RC}$ and $HyBAC_{AC}$ with a role-centric approach and an attribute-centric approach, respectively. The $HyBAC_{RC}$ model is based on their previously proposed $EGRBAC$ model [3], which encapsulates relatively static attributes in access decisions in user roles and device roles, using the user $HyBAC_{AC}$ is based on their previously proposed $HABAC$ model [4], where, $HyBAC_{AC}$ has the ability to capture rapidly changing attributes and further restrict the privileges available to each user. Among them, $HyBAC_{AC}$ has the point that the authorization process is simple and $HyBAC_{AC}$ model has finer granularity in privilege assignment. However, the access evaluation and authorization process is more complex.

In addition, many researchers have proposed blockchain-based access control schemes [1, 13, 16]. Nakamura *et al.* [15] proposed a decentralized capability-based access control scheme, which utilizes Ethernet smart contract technology to achieve finer-grained and more flexible capability delegation and also ensures the consistency of the delegation information with the information stored in the token. In [12], Liao *et al.* extend the attribute-based access control model by applying blockchain technology to the smart home ABAC model, and the proposed scheme solves the risk of unauthorized access due to third-party centralized management.

In summary, much research has been conducted by researchers on the access control problem for smart homes. The proposed access control schemes can be divided into two types based on the traditional access control model and blockchain technology. The traditional access model-based scheme has the advantages of low computational overhead and simple deployment but coarse granularity of permission assignment; the blockchain-based scheme achieves fine-grained access control but with high computational and storage overhead. The existing research does not consider the issue of device hazard level, which makes the proposed model not applicable to the smart home scenario. Therefore, it is necessary to investigate the access control problem related to smart homes.

3 The Fine-grained and Dynamic Access Control Model

In this section, the basic architecture of the proposed model is presented first, and the definition of the relevant attributes is elaborated later, followed by a detailed implementation of fine-grained and dynamic access control. Finally, the process of executing access decisions by the proposed model is elaborated.

3.1 The Basic Architecture of The Access Control Model

The proposed model in this paper is divided into four modules, which are the attribute resolution module, trust and risk value calculation module, policy matching module, and access decision module. The basic architecture of the model is shown in Figure 1.

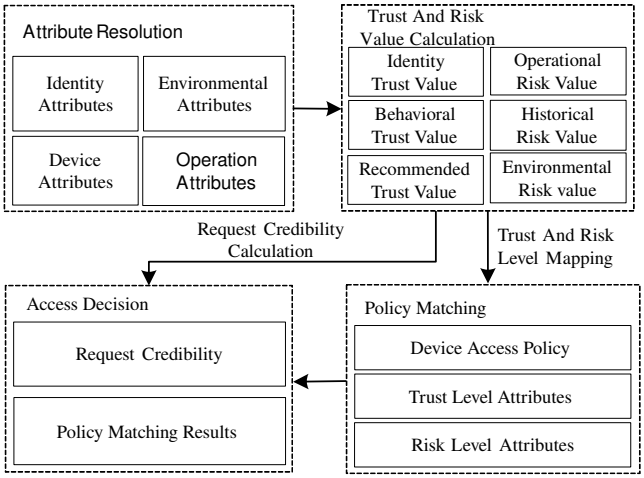


Figure 1: The basic architecture of a fine-grained dynamic access control model

The main function of the attribute parsing module is to parse the attributes of the user request to generate the set of attributes corresponding to the user's manipulation request. The trust and risk value calculation module calculates the trust and risk values of the user request based on the parsing results of the attribute parsing module and the predefined recommended trust values and risk factors of the user in the database. The policy matching module matches the trust and risk level attributes of the user with the access policy of the device. The trust and risk level attributes are implemented through predefined level mapping relationships. The main function of the access decision module is to perform access decisions on user requests based on the results of request trust calculation and policy matching to decide whether to allow the user to make the request.

3.2 Related Definitions

The access control model proposed in this paper is based on the ABAC model, so this paper uses attributes to describe the user's manipulation requests. The relevant definitions of the model are as follows:

Definition 1. *User Request(UR)*

$$UR \rightarrow (U, D, E, A)$$

A user's request UR is represented by a quaternion, which is the user U of the smart home device, the device D requested to be operated, and the context environment E when the operation request is initiated. The action A for the smart home device.

Definition 2. *Attributes*

$$Attributes \rightarrow UA, DA, EA, AA$$

UA: *A set of attributes owned by the user. Used to describe the identity of the user, such as family role, age, occupation, etc.*

DA: *The set of attributes owned by the device, is used to identify the main characteristics of the device. Attributes such as the name of the device, the location of the device, etc.*

EA: *The set of environment attributes that the user has when initiating an operation request. Used to identify relevant risks and provide a basis for subsequent access decisions.*

AA: *The set of operations that can be performed on the device, such as opening, closing, or performing a function on the device, or modifying or deleting data on the device.*

Definition 3. *Attributes of Request(AR)*

$$AR \rightarrow ATTR(U), ATTR(D), ATTR(E), ATTR(A)$$

Among them:

$$\mathrm{ATTR}(U) = UA_1 \cap UA_2 \cap UA_3 \cap \ldots \cap UA_n$$
$$\mathrm{ATTR}(D) = DA_1 \cap DA_2 \cap DA_3 \cap \ldots \cap DA_n$$
$$\mathrm{ATTR}(E) = EA_1 \cap EA_2 \cap EA_3 \cap \ldots \cap EA_n$$
$$\mathrm{ATTR}(A) = AA_1 \cap AA_2 \cap AA_3 \cap \ldots \cap AA_n$$

For each attribute corresponding to a request, it is a triple <attribute, value, association>, value is the specific value of the attribute, and the association is the relationship between the attribute and the value.

For example, for a user request *UR (U1, D1, E1, A1)*. It may have the following attributes:

$$\mathrm{ATT}(U_1) = \{ \text{"role"} = \text{"father"}, \text{"age"} = \text{"31"} \}$$
$$\mathrm{ATT}(D_1) = \{ \text{"type"} = \text{"door"}, \text{"danger"} = \text{"5"} \}$$
$$\mathrm{ATT}(E_1) = \{ \text{"location"} = \text{"school"}, \text{"net"} = \text{"5G"} \}$$
$$\mathrm{ATT}(A_1) = \{ \text{"action"} = \text{"open"} \}$$

In addition, different smart home devices have different danger levels, which are shown in the attribute set $ATT(D)$ of smart home devices by defining the "danger" attribute. However, even for the same home device, the danger level is different when different users use some of its functions. For example, a child operating a gas stove and his parents operating a gas stove have different danger levels.

For this reason, this paper proposes the concept of the danger matrix, which specifies the danger factor when different users operate the device. The higher the danger factor, the greater the impact on the user's life safety, which makes the proposed scheme more suitable for the complex and changing environment of the smart home system. For example, the danger level of three people, grandfather, mother, and son, operating the gas stove and sweeping robot can be defined as the danger factor matrix shown in Table 1. The danger factor is used to calculate the risk value of the user's operation.

3.3 Trust And Risk Calculation

The trust and risk value calculation is the basis of the access control model proposed in this paper to achieve

Table 1: Examples of danger factors for different users operating different devices

	gas stove	sweeping robot
grandfather	0.7	0.5
mother	0.3	0.2
son	0.9	0.6

dynamic and fine-grained access control. In this section, the method of trust value calculation is first described. In detail, the user trust value consists of three components: identity trust value, behavior trust value, and recommendation trust value.

3.3.1 Identity Trust Value Calculation

The identity trust value is determined by the user's identity information when initiating the operation request, which mainly includes user ID, password, credentials, and role attributes. In the scheme of this paper, the base trust value is defined as T_A. The calculation process is as follows:

Firstly, $ATTR(U)$from the set of user request attributes AR in which the basic identity attributes for trust evaluation are selected. Since the basic identity attributes here are represented as key-value pairs that cannot be computed mathematically, it is necessary to quantify each basic identity. In this paper, the attributes of the user's current request are compared with the attributes of the user's. The matching result is that the attribute is assigned a value of 1 if true. Otherwise, it is assigned a value of 0, as shown in Equation (1).

$$UA_i = \begin{cases} 1, & UA_i \&\& \text{ history } (UA_i) = TRUE \\ 0, & UA_i \&\& \text{ history } (UA_i) = FALSE \end{cases} \quad (1)$$

Finally, considering that different attributes have different weights, the user's identity trust value T_A is calculated by Equation (2).

$$T_A = UA_1 \cdot w_1 + UA_2 \cdot w_2 + \cdots + UA_m \cdot w_{\mathrm{m}} \quad (2)$$

In Equation (2), this paper assigns different weight proportions to the identity attributes of users, which are predefined by the system. The significance of this is that for a particularly complex IoT system like a smart home, defining different weight ratios for different device permissions can improve the flexibility and security of the system. The sum of the weight proportions w is 1.

3.3.2 Behavioral Trust Value Calculation

In the model proposed in this paper, the behavioral trust value of a user is defined by the number of normal and abnormal behaviors of the user, which is defined by T_B, and a slow-increasing and fast-decreasing strategy is used to make the model proposed in this paper have better dynamics. The behavioral trust value of users is calculated as shown in Equation (3).

$$T_B = \frac{\sum_{x=1}^{n} b_x}{\sum_{x=1}^{n} |b_x|}, \quad b_x = \left\{ \begin{array}{ll} 1, & b_x \text{ is normal} \\ -2, & b_x \text{ is abnormal} \end{array} \right\} \quad (3)$$

In Equation (3), the user's xth access behavior is defined by b_x.

3.3.3 Recommended Trust Value Calculation

A user's recommended trust value T_R is equal to the average of all other users' recommended trust values for that user. For example, the recommended trust value of user i can be calculated by equation (4).

$$T_R = \frac{\sum_{j=1, j\neq i}^{n} r_{j,i}}{n-1} \quad (4)$$

In Equation (4), the recommended trust value of the user j to the user i is defined by $r_{j,i}$.

3.3.4 Risk Value Calculation

In a smart home environment, there are various device types, each with different danger levels. Correspondingly, different users have different risk factors when controlling different devices. Therefore, the user's risk value when operating the device should be calculated to better protect the user's life and properties.

In this paper, the user's risk value consists of three parts, which are Environment risk R_E, operation risk R_O, and history risk R_H. Specifically, the risk of the user U_i request to operate device D_j can be defined by Equation (5).

$$R_{i,j} = R_E * k_1 + R_o * k_2 + R_H * k_3 \quad (5)$$

In Equation (5), k_1, k_2, and k_3 are the relevant weight proportions, and their sum is 1.

The environment risk value R_E is calculated from the user request with the environmental related attributes as defined in Equation (6).

$$R_E = EA_1 \cdot \alpha_1 + EA_2 \cdot \alpha_2 + \cdots + EA_n \cdot \alpha_{\mathrm{n}} \quad (6)$$

In Equation (6), α_1, α_2 .....α_n are the relevant weight proportions and their sum is 1. The calculation of EA_n is similar to Equation (1).

The operational risk value C_r is defined by the user's danger factor when operating the device.

Let $\mu_{i,j}$ be the danger factor when the user U_i operating the device D_j, then $R_o = \mu_{i,j}$.

R_H is the history risk value of the user. In particular, according to the distance between the time when the history risk was generated and the time when the current user request was generated. Different weights are assigned to history risk values for different periods, meaning a "far

small, near big" strategy is used. The detailed calculation method is shown in Equation (7).

$$R_H = R^1_{1,1} \cdot \beta_1 + R^2_{1,1} \cdot \beta_2 + \cdots + R^m_{1,1} \cdot \beta_m \quad (7)$$

In Equation (7), the formula for calculating the history risk value when user 1 operates device 1 is given, where m is the distance from the time when the user's history risk value was generated. The larger m is, the further it is from the present, the corresponding value of β_m is smaller. β_m is the weight, which sums to 1.

3.4 Dual Access Control

In subsection 3.3 of this paper, the calculation of the relevant trust and risk values when a user initiates a control request is described, and they are the basis for implementing access control in the model proposed in this paper. There are two decision processes for user control requests: dynamic access control and fine-grained access control, respectively. The steps for implementing these two access controls are described in detail in this subsection.

3.4.1 Dynamic Access Control Based on Request Credibility Calculation

Dynamic access control is implemented by calculating the request credibility, which is also the first access control of user requests in the model proposed in this paper.

In Section 3.3, the procedure for calculating the trust and risk values of a user's request is given, but in the model proposed in this paper, whether a user is allowed this access request does not rely on either of these two alone, but rather combines them to achieve accurate identification of the credibility of each user's access request. In Equation (8), the specific method for calculating the request credibility T_{req} is given.

$$T\left(req_x\right) = T_A * \left(T_A * t_1 + T_B * t_2 + T_R * t_3 - R_{i,j}\right) \quad (8)$$

The credibility calculation of user i at the xth access request, requesting control of device j, is given in Equation (8), where t_1, t_2, and t_3 are the weight, and their sum is 1.

It is worth noting that in this paper, the user's identity trust value is multiplied by the results of the operation of other trust and risk values. The reason for doing so is that when the user's identity trust is 0, it is possible that the user is not operating the device himself. The user request should be rejected immediately. It is easy to analyze that taking this calculation method can improve the security of the model in this paper.

Once the user's request credibility is calculated, the user can be initially granted access to the device with the corresponding security level according to the request credibility. Specifically, in this paper, there are five different security levels for the device, and the correspondence between them and the request credibility size is shown in Table 2.

Table 2: Correspondence between device security level and request credibility

T_{req}	Device Security Level
[0.9,1)	1
[0.7,0.9)	2
[0.5,0.7)	3
[0.3,0.5)	4
(0,0.3)	5
(- $\propto$,0]	—

In Table 2, "—" indicates rejection of the user request.

Table 2 indicates that the security level of the devices that a user can access varies with the user request credibility, thus enabling a dynamic decision on user access requests. However, whether a user can access a specific device or not is determined by the results of fine-grained access control.

3.4.2 Fine-grained Access Control Based On Attribute Based Encryption

Fine-grained access control is achieved through access policy matching, which is also the second access control of user requests.

By matching the set of attributes corresponding to the user operation request with the device access policy stored in the cloud. In the model proposed in this paper, attribute-based encryption is used to implement this process. However, the traditional attribute-based encryption scheme has the following problem: as the attributes increase, the encryption and decryption time overhead becomes larger and larger. For this reason, in the proposed scheme, the trust and risk values calculated from each user's attributes in Section 3.3 are preprocessed.

Specifically, in this paper, each trust and risk value of the user is divided into five different risk levels, as shown in Table 3 and Table 4. Thus, the number of attributes included in the access policy of the device is controlled to be less than 6.

Table 3: Trust Level Mapping

$TrustValue$	T_{A_level}	T_{B_level}	T_{R_level}
[0.9,1]	5	5	5
[0.7,0.9)	4	4	4
[0.5,0.7)	3	3	3
[0.3,0.5)	2	2	2
[0,0.3)	1	1	1

Through the above trust and risk level classification process, the trust and risk value levels are obtained and then used as attributes for policy matching, significantly reducing the number of attributes in the access policy

Table 4: Risk Level Mapping

$RiskValue$	R_{E_level}	R_{O_level}	R_{H_level}
[0.9,1]	5	5	5
[0.7,0.9)	4	4	4
[0.5,0.7)	3	3	3
[0.3,0.5)	2	2	2
[0,0.3)	1	1	1

and reducing the time overhead of the policy matching process.

The specific process of strategy matching is divided into the following two steps:

1) Device Access Policy Formulation
First is the Formulation of device access policies. This paper formulates different access policies for different devices to achieve fine-grained permission control and improve the model's security.

Specifically, this paper uses an access structure tree as the access policy implementation scheme. The basic structure of an access structure tree is shown in Figure 2.

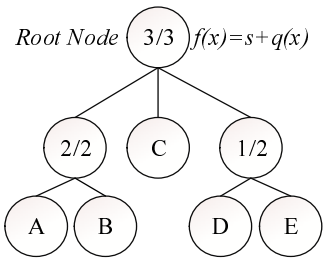


Figure 2: Access Structure Tree

In Figure 2, each non-leaf node represents a threshold gate. Each threshold gate has some child nodes and a threshold value, denoted by (t, n), where t is the threshold value and n is the number of child nodes. When $t = 2$,$n = 2$, it means "and" gate, and when t=1,n=2, it means "or" gate. The children of each node have a number i ,i=1,2,3..., in the order from left to right. Let $parent(z)$ denote the father node of node z; if z is a leaf node, let $att(z)$ denote the attribute linked to node z; let $index(z)$ denote the number of the node. The access structure tree as shown in Figure 2 represents the access policy as *(A and B) and C and (D OR E)*.

The root node corresponds to a polynomial $f(x)$ and the constant term is the secret value s , $q(x)$ is a random polynomial of order $t-1$ and t is the threshold value. Then according to Shamir's secret sharing method, a random polynomial is generated for each non-leaf node, and its constant term is the value after substituting the node number i into the polynomial $f(x)$ of the parent node. And so on until the secret value is shared to the leaf node.

When a user requests to operate the device, if the user's various trust and risk level attributes can decrypt the secret value s of the root node, the plaintext M ,can be recovered, and the subsequent access decision process can be continued.

Specifically, the access policy (AP , Access Policy) can be defined by Equation (9), where $\mathcal{T}$ represents an access control tree, and s is the secret value bound to the root node.

$$\text{result} \leftarrow \mathcal{T}\left(\begin{array}{c} s, T_{\text{A\_level}}, T_{\text{B\_level}}, T_{R\_\text{level}}, \\ R_{E\_\text{levle}}, R_{O\_\text{leve}}, R_{H\_\text{level}} \end{array}\right) \quad (9)$$

In Equation (9), result denotes the decision result, and its value has two cases of 1 and 0, which denote permission and rejection, respectively.

2) Access Policy Matching
The process of access policy matching is the process of performing attribute encryption and decryption.The main steps can be divided into the following steps:

Step 1: Initialization, $Setup(1^{\lambda}) \rightarrow (PP, MK)$.
Firstly, $\alpha, \beta \in Z_p$ is randomly selected by the attribute authorization center to generate the system common parameter PP=$(G_0, g, h = g^{\beta}, e(g,g)^{\alpha})$, where g is the generating element of G_0 and e is the bilinear pair mapping of the group G_0 to G_1. The system master private key MK=(β, g^{α}) is also generated.

Step 2: Encryption, $Encrypt(PP, M, \mathcal{T}) \rightarrow CT$.
The system randomly selects the secret value $s \in Z_p$ and uses the public key PP to encrypt the plaintext M, and embedding the access policy $\mathcal{T}$ in the ciphertext. Then, the ciphertext CT is generated as follows: $CT = (CT_1 = \mathcal{T}, CT_2 = Me(g,g)^{\alpha s}, CT_3 = h^s, \forall y \in Y : CT_y^1 = g^{q_y(0)}, CT_y^2 = H(att(y)^{q_y(0)}))$. Y is the set of leaf nodes in the access structure tree. It is worth noting that in this paper, the plaintext M here is an element randomly selected on the group G_0. When a user's operation request arrives , if the attributes corresponding with the request can successfully decrypt the element. The user can obtain the device operation privilege.

Step 3: Private key generation, $KeyGen(MK, S) \rightarrow SK$.
The private key generation algorithm will accept the attribute set S as input and output with the key identified with that set. The algorithm firstly selects a random $r \in Z_p$ .Then, for each attribute $j \in S$, a random $r_j \in Z_p$ is selected. The secret key is then computed

as SK=$(SK_1 = g^{\alpha+\beta}, \forall j \in S : SK_j^1 = g^r \bullet H(j)^{r_j}, SK_j^2 = g^{r_j})$, where the hash function H maps the attribute j to an element of group G_0.

Step 4: Decrypt, $Decrypt(PP, CT, SK) \rightarrow M$. The decryption algorithm is a recursive process, when the decryption node x is a leaf node, let $i = att(x)$, and suppose $i \in S$, execute the following algorithm:

$$\begin{aligned}
&\text{DecryptNode}(CT, SK, x) \\
&= \frac{e\left(SK_i^1, CT_x^1\right)}{e\left(SK_i^2, CT_x^2\right)} \\
&= \frac{e\left(g^r \cdot H(i)^{r_i}, h^{q_x(0)}\right)}{e\left(g^{r_i}, H(i)^{q_x(0)}\right)} \\
&= e(g,g)^{r \cdot q_x(0)}
\end{aligned}$$

If $i \notin S$, then the algorithm outputs $\perp$.

When x is a non-leaf node, the algorithm is computed as follows:

For all child nodes z of node x, invoke the function $DecryptNode(CT, SK, z)$, and store the result as F_z . Let S_x be an arbitrary set of child nodes z of size k_x and satisfy $F_z \neq \perp$. If no such set exists, the function returns $\perp$. Otherwise, the following algorithm is executed:

$$\begin{aligned}
F_x &= \prod_{z \in S_x} F_z^{A_{i,S_x'}(0)} \\
&= \prod_{z \in S_x} \left(e(g,g)^{r \cdot q_{parent(z)}(\text{ index }(z))}\right)^{A_{i,S_x'}(0)} \\
&= \prod_{z \in S_x} e(g,g)^{r \cdot q_x(i) \cdot A_{i,S_x'}(0)} \\
&= e(g,g)^{r \cdot q_x(0)}
\end{aligned}$$

Among them, $i = \text{index}(z), S_x' = \{\text{index}(z) : z \in S_x\}$ $A_{i,S}(x) = \prod_{j \in S, j \neq i} \frac{x-j}{i-j}$ is the Lagrange coefficient.

Finally, If the tree is satisfied by S, we set $A = DecryptNode(CT, SK, r) = e(g,g)^{rq_R(0)} = e(g,g)^{rs}$, the plaintext M can be calculated by performing the following algorithm:

$$\begin{aligned}
&\frac{CT_2}{(e(CT_3, SK_1)/A)} \\
= \quad &\frac{CT_2}{\left(e\left(h^s, g^{(\alpha+r)/\beta}\right)/e(g,g)^{rs}\right)} \\
= \quad &M.
\end{aligned}$$

3.4.3 Definition of Multi-access Policy

Considering that a device may correspond to more than one access policy. At this point, the final access policy can be expressed as shown in Equation (10):

$$AP\left(AP_j\right) = AP_1 \cap AP_2 \cap AP_3 \cdots \tag{10}$$

For example, Table 5 gives a set of access policy examples, and the final access policy can be represented by Equation (10).

Table 5: Example of Access Policy

	AP_1	AP_2	AP_3	...
T_{A_level}	5	4	4	...
T_{B_level}	2	3	2	...
T_{R_level}	1	3	2	...
R_{E_level}	3	3	1	...
R_{O_level}	2	2	1	...
R_{H_level}	5	4	4	...

3.5 Access Decision Process

Finally, this paper gives a figure of the whole process from the user initiating the operation request to the model executing the access decision. This is shown in Figure 3.

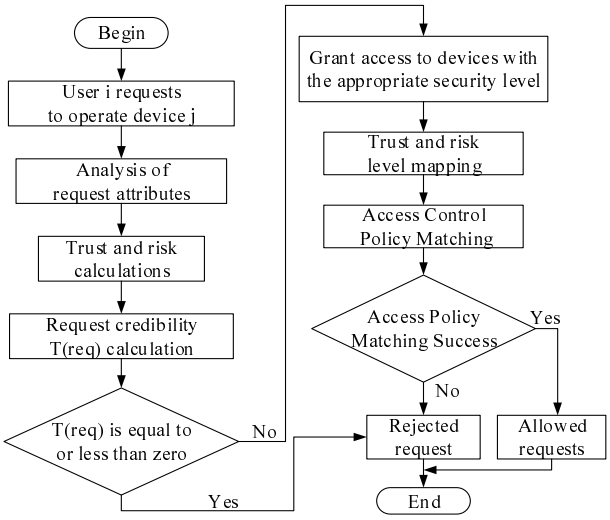


Figure 3: The Access Decision Process

In Figure 3, when the access subject, i.e., the user, initiates an operation request, the model first performs the trust and risk value calculation process based on the set of attributes corresponding to the user's request, and then implements access control based on the calculation results. Among them, the request trust calculation process is the first access control of the user request, and the access policy matching is the second access control. By implementing dual access control on user requests, the overall security of the smart home system is improved.

4 Experiments and Analysis

The experiments in this paper were accomplished under Windows 10 environment, using JetBrains IDEA as the

programming tool, Java as the programming language, JDK version 1.8, JPBC library to implement attribute encryption, and Cisco Packet Tracer network simulator to simulate smart home devices. The specific configuration of the computer is Intel CPU i5-8250U with 8GB RAM.

4.1 The Design of Experiments

The model proposed in this paper is for smart home devices in the cloud computing environment. The experiments simulate relevant smart home devices through Cisco Packet Tracer network simulator and then realize the one-to-one mapping of devices to device images in the OneNet cloud platform through programming.

In the experiments, three users are simulated in this paper: the smart home system administrator: mother, denoted as user 1; the 6-year-old son, denoted as user 2; and the home guest, denoted as user 3. Then, 20 smart home devices were simulated with different security levels, as shown in Table 6. They were mapped to the cloud platform via Transmission Control Protocol (TCP) passthrough, and a series of application programming interfaces (API) for operating the devices were generated for these devices.

Table 6: Number of devices with different security levels in the experiment

Device Security Level	1	2	3	4	5
Number of devices	5	5	5	3	2

Finally, the access control model proposed in this paper is programmatically implemented, including trust and risk value calculation algorithms, policy matching algorithms, and the definition of relevant device access policies. The experiments simulate the process of user operation of the device by controlling the API invocation, and whether the user can successfully invoke the API depends on the decision result of the model on its operation request.

4.2 Analysis of Results

1) Dynamicity Verification

In the experiment, to verify the dynamics of the model, 30 access operations were performed by simulated users 1, 2, and 3, respectively.

The first ten times simulated users perform abnormal operations, i.e., the users try to access devices they cannot operate. The middle ten times simulate users performing normal operations. Finally, after changing the user request attributes, the user is simulated to perform ten operations. The change in the request credibility of each user is recorded, which leads to the change in the number of devices that the user can operate, and the experiment results are shown in Figure 4 and Figure 5.

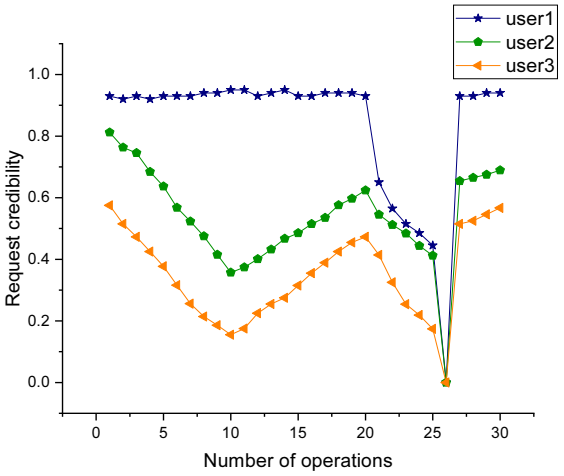


Figure 4: Changes in request credibility

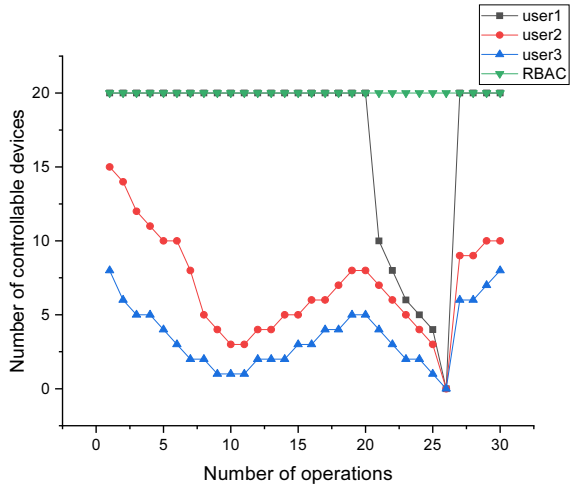


Figure 5: Changes in the number of controllable devices

Figure 5 shows that the initial number of manipulable devices for users 1, 2, and 3 differs because the model proposed in this paper considers the operational risk of users when operating different devices as well as their identity trust value and recommended trust value. User 1, as a system administrator, can access all devices. User 2, as a minor, cannot access some devices with higher risk factors, so the initial number of controllable devices is less than that of the system administrator. User 3, as a home guest, has a lower identity trust value and recommended trust value so that it can access the least number of devices.

In the first ten access operations, the number of accessible devices does not change because the behavioral trust value of user 1 does not decrease. After all, it can access all devices. User 2 generates abnormal behavior records when accessing devices it cannot manipulate. As the number of abnormal behaviors increases, its behavior trust value decreases rapidly. Thus, the request credibility

decreases, and thus, the number of manipulable devices decreases. In addition, it also cannot pass the access policy for devices with high requirements of behavior trust value level, so it also makes the number of accessible devices gradually decrease; the same for user 3.

The experiment simulates users performing normal operations in the middle ten access operations. User 1 can control all devices, so the number of controllable devices does not change. User 2 and user 3 gradually increase the number of normal behaviors due to their normal operations, so the behavioral trust value gradually increases, and the number of controllable devices gradually increases. It is worth noting that the number of devices accessible to the user during normal operation increases relatively slowly compared to the user during abnormal operation due to the different rates of decrease and increase in the behavioral trust calculation due to the "fast increase and slow decrease" strategy. The advantage of this strategy is that it can motivate users to perform normal operations and reduce the possibility of illegal access to devices.

In the last ten access operations, the number of manipulable devices for users 1, 2, and 3 is firstly reduced to zero and then increased, which is because, in the first five access operations in this paper, the environmental attributes of the user's request are changed so that their environmental risk value increases, which makes their request credibility decrease, which makes the number of manipulable devices decrease. Then, the user's identity attribute is changed so that the user's identity trust value is zero, which makes the request credibility zero. Thus, the number of controllable devices for users 1, 2, and 3 is zero. In the last five operations, the environment and identity attributes of the user's request are restored, which increases the request credibility so that the number of controllable devices increases again.

In Figure 5, the number of devices that can be controlled by the user under the role-based access control (RBAC) model remains constant, compared with the number of devices that can be controlled by the user in the proposed model in this paper, which is constantly changing, i.e., the user's control authority over the devices is constantly changing. The devices that can be controlled originally may not be controlled now, thus verifying the dynamic nature of the proposed model in this paper.

In summary, in this paper, because the device is divided into different security level devices, the proposed scheme can better protect the user's privacy and home security and prevent illegal users from accessing devices with higher security levels to protect the user's life and property security.

2) Fine-grained verification

To test whether the proposed access control model in this paper can provide fine-grained access control for different users. In the experiment, users 1, 2, and 3 are simulated to access devices with different risk levels at different times and locations to verify the fine-grained nature of the proposed model.

Specifically, the access policies of each device in this experiment are shown in Table 8, where AP_1 is represented as the access policy of device 1, and AP_2 and AP_3 are similar. The model proposed in this paper is verified by simulating the user requests shown in Table 7. The experimental results are shown in Table 9.

Table 7 indicates that nine relevant requests were simulated experimentally to verify the fine-grained nature of the model. These requests were initiated by users with different roles and ages under different environmental attributes. They attempt to control three smart home devices, namely, door1, door2, and water heater.

In Table 9, "✓" means the request is allowed, "×" means the request is rejected, and "—" means not applicable.

Tables 7 and 9 show that when the location of user 1 initiates an access request changes, its permission to operate device 1 also changes. User 1 and User 2 also have different control rights to device 2 due to their different ages, which correspond to different control risks. The location where User 3 initiates a request, even if it is in the user's house, does not allow him to control the opening of device 3 because his role is that of a guest. Therefore, it cannot enter a more private space, such as a bedroom, and when its location is changed, it cannot successfully operate device 1.

The results of this experiment indicate that by using attribute encryption and setting different access policies for different devices, the access control model proposed in this paper successfully achieves fine-grained access control for users in terms of role, age, location, and operation risk.

3) Efficiency verification

In order to verify the efficiency of the access control model proposed in this paper, the experiments simulate users initiating relevant access requests and assume that the relevant attributes contained in the user requests can be verified by the access policies of the devices. By gradually increasing the number of attributes in the user access requests, the encryption and decryption time overhead of the relevant policy matching process is recorded and compared with the time overhead of the traditional ciphertext policy attribute-based encryption(CP-ABE) scheme. The experimental results are shown in Figure 6 and Figure 7.

Figures 6 and 7 show that the encryption and decryption time overheads of the proposed scheme and the original CP-ABE scheme are the same when the number of attributes included in the user request is six or less. In contrast, the encryption and decryption time of the proposed scheme remains the same after the number of attributes increases to 6 or more. At the same time, the time overhead of the original attribute encryption and decryption scheme gradually increases. This is because in the proposed scheme, instead of using the attributes included in the user access request, we first perform trust and risk assessment on them to obtain the trust and risk

Table 7: The User Request

Requests	user_id	role	age	req_location	device_id	device_name	device_type	req_operation
Req_1	1	families	35	home	1	door1	Anti-theft door	open
Req_2	1	families	35	others	1	door1	Anti-theft door	open
Req_3	1	families	35	home	2	water heater	water heater	open
Req_4	2	families	8	home	1	door1	Anti-theft door	open
Req_5	2	families	8	home	2	water heater	water heater	open
Req_6	2	families	8	others	1	door1	Anti-theft door	open
Req_7	3	resident	25	home	1	door1	Anti-theft door	open
Req_8	3	resident	25	home	3	door2	bedroom door	open
Req_9	3	resident	25	others	1	door1	Anti-theft door	open

Table 8: The Device Access Policy

	AP_1	AP_2	AP_3
T_{A_level}	5	2	5
T_{B_level}	4	3	3
T_{R_level}	4	2	4
R_{E_level}	1	3	2
R_{O_level}	4	1	4
R_{H_level}	2	2	2

Table 9: The Access Policy Matching Results

Requests	AP1	AP2	AP3
Req_1	✓	—	—
Req_2	×	—	—
Req_3	—	✓	—
Req_4	✓	—	—
Req_5	—	×	—
Req_6	×	—	—
Req_7	✓	—	—
Req_8	—	—	×
Req_9	×	—	—

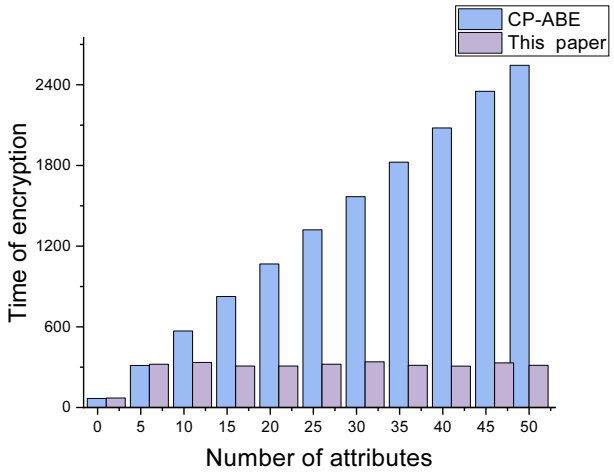


Figure 6: Comparison of encryption time

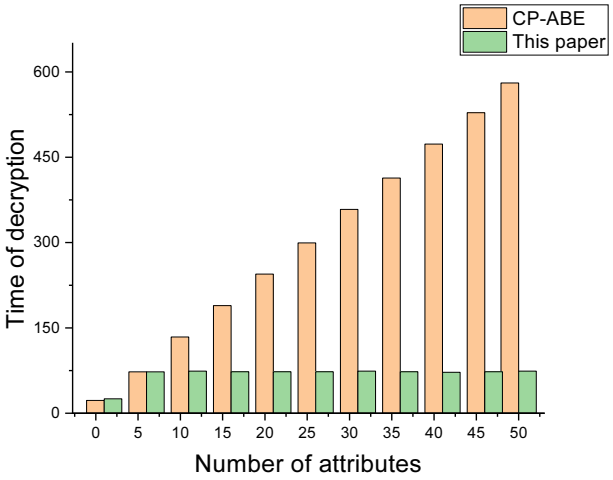


Figure 7: Comparison of decryption time

assessment values corresponding to the user request and then map them to the corresponding trust and risk levels according to the interval in which these trust and risk values are located. The policy matching is performed based on these trust and risk level attributes. Specifically, in the scheme proposed herein, the access policy for a device includes up to six attributes. Therefore, after the number of attributes in user requests increases to 6, the proposed scheme's variation of encryption and decryption time overhead tends to be flat.

It is worth noting that there is a request credibility calculation step in the decision process of the proposed model. When the user's requested attribute changes, the corresponding request confidence value will also change, and when the request credibility is lower than or equal to 0, the user's request will be rejected directly and the subsequent policy matching process based on attribute encryption will not be executed. In this paper, the cal-

culation process of request credibility is based on addition, subtraction, and multiplication, and its computation overhead is much lower than that of attribute encryption based on bilinear pairs. This approach can significantly reduce the overall overhead of the model if a large number of access requests are initiated by illegal users in a short period. This demonstrates the robustness of the proposed model in this paper. Based on the above analysis, it can be seen that the computation of request credibility can further reduce the computational overhead of the model, which shows the efficiency of the proposed access control model from another perspective.

The above analysis and experimental results indicate that by reducing the number of policy attributes, the proposed scheme reduces the time overhead of encryption and decryption compared with the traditional CP-ABE scheme, so the proposed model is efficient.

4) Comparison with existing schemes

Table 10 shows the advantages of the proposed scheme compared to existing smart home access control schemes.

In Table 10, DM is dynamic, and FG is fine-grained. The abbreviation DDAC refers to Dangerous Device Access Control. CA represents context awareness, MUM represents multi-user management, PM stands for Policy Management, and TA stands for Temporary Access.

As demonstrated in Table 10, the access control scheme proposed in this paper meets all the requirements of a smart home access control system as described in the literature [14]. Compared to the access control schemes presented in Ref. [8], Ref. [9], and Ref. [18], the proposed method introduces dynamic access control by calculating trust and risk values and separating the classification of devices into different risk levels. It accomplishes fine-grained access control for attribute levels through attribute-based encryption while reducing time overhead by mapping risk and trust levels.

Table 10: Comparison with existing schemes

	Ref. [8]	Ref. [9]	Ref. [18]	Proposed
DM	×	✓	✓	✓
FG	✓	✓	✓	✓
DDAC	×	×	×	✓
CA	✓	✓	✓	✓
MUM	×	×	✓	✓
PM	×	✓	×	✓
TA	×	×	×	✓

Further, the proposed scheme enables access control to dangerous devices by using a pre-set operation risk matrix. Temporary access and multi-user management through the computation of identity trust and environmental risk values. The classification of device risk levels allows users to participate in formulating access policies. Thus, our scheme offers superior flexibility and better suits for the intricate and ever-changing smart home environment.

5 Conclusions

In this paper, a fine-grained dynamic access control model is proposed, which applies to smart home devices in a cloud environment. The architecture of the proposed model, the definition of relevant attributes, the calculation methods of trust and risk values, and the specific process of access decision implementation by the model are elaborated in this paper. By simulating users to initiate relevant access requests, the proposed scheme is validated in terms of dynamism, fine-grained, and efficiency. The experimental results show that the proposed scheme achieves real-time awareness of user access requests by calculating the relevant trust and risk values and dynamic changes of user control authority by calculating the request credibility. Fine-grained access control of users in terms of role, age, location, etc., is achieved by using attribute-based encryption. The number of attributes in the device access policy is reduced by trust and risk level mapping, thus reducing the computational overhead of the scheme. In particular, by defining the risk factor matrix, the proposed model solves the risk problem when different users operate different devices. By performing request trust calculation and policy matching on user operation requests, it achieves dual access control on user requests, thus improving the security of the model. The comparison with existing schemes shows the advantages of our proposed model. In future research, we will focus on the dynamic update method of device access policy to further improve the dynamics of the model.

Acknowledgments

This research is supported by the National Natural Science Foundations of China under Grants No.61862040 and No.62162039. The authors gratefully acknowledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] A. I. Abdi, F. E. Eassa, K. Jambi, K. Almarhabi, M. Khemakhem, A. Basuhail, and M. Yamin, "Hierarchical Blockchain-Based Multi-Chaincode Access Control for Securing IoT Systems," *Electronics*, vol. 11, p. 711, February 2022.

[2] N. Almolhis, A. M. Alashjaee, S. Duraibi, F. Alqahtani, and A. N. Moussa, "The security issues in iot - cloud: A review," in *2020 16th IEEE International Colloquium on Signal Processing and Its Applications (CSPA)*, pp. 191–196, 2020.

[3] S. Ameer, J. Benson, and R. Sandhu, "The EGR-BAC Model for Smart Home IoT," in *2020 IEEE*

21st International Conference on Information Reuse and Integration for Data Science (IRI), pp. 457–462, Las Vegas, NV, USA, August 2020. IEEE.

[4] S. Ameer, J. Benson, and R. Sandhu, "An Attribute-Based Approach toward a Secured Smart-Home IoT Access Control and a Comparison with a Role-Based Approach," *Information*, vol. 13, p. 60, January 2022.

[5] S. Ameer, J. Benson, and R. Sandhu, "Hybrid Approaches (ABAC and RBAC) Toward Secure Access Control in Smart Home IoT," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–18, 2022.

[6] S. Bhatt, T. K. Pham, M. Gupta, J. Benson, J. Park, and R. Sandhu, "Attribute-Based Access Control for AWS Internet of Things and Secure Industries of the Future," *IEEE Access*, vol. 9, pp. 107200–107223, 2021.

[7] R. H. Dong, T. T. Xu, and Q. Y. Zhang, "Access control model of industrial control system based on multi-attribute decision making," *International Journal of Network Security*, vol. 23, no. 6, pp. 1037–1048, 2021.

[8] S. Dutta, S. Chukkapalli, M. Sulgekar, S. Krithivasan, and A. Joshi, "Context Sensitive Access Control in Smart Home Environments," in *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, pp. 35–41, Baltimore, MD, USA, May 2020. IEEE.

[9] G. Goyal, P. Liu, and S. Sural, "Securing smart home iot systems with attribute-based access control," in *Proceedings of the 2022 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems*, Sat-CPS '22, p. 37–46, New York, NY, USA, 2022. Association for Computing Machinery.

[10] Y. B. Hamdan, "Smart home environment future challenges and issues-a survey," *Journal of Electronics and Informatics*, vol. 3, no. 01, pp. 239–246, 2021.

[11] A. Haque, B. Bhushan, and G. Dhiman, "Conceptualizing smart city applications: Requirements, architecture, security issues, and emerging trends," *Expert Systems*, vol. 39, June 2022.

[12] K. Liao, "Design of the Secure Smart Home System Based on the Blockchain and Cloud Service," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–12, January 2022.

[13] H. Liu, D. Z. Han, and D. Li, "Fabric-IoT: A blockchain-based access control system in IoT," *IEEE Access*, vol. 8, 2020. Publisher: IEEE.

[14] Z. N. Mohammad, F. Farha, Aom Abuassba, S. Yang, and F. Zhou, "Access control and authorization in smart homes: A survey," *Tsinghua Science and Technology*, vol. 26, no. 6, pp. 906–917, 2021.

[15] Y. Nakamura, Y. Zhang, M. Sasabe, and S. Kasahara, "Exploiting Smart Contracts for Capability-Based Access Control in the Internet of Things," *Sensors*, vol. 20, p. 1793, March 2020.

[16] A. Qashlan, P. Nanda, and X. He, "Security and privacy implementation in smart home: Attributes based access control and smart contracts," in *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 951–958. IEEE, 2020.

[17] Y. Ren, Y. Leng, J. Qi, P. K. Sharma, J. Wang, Z. Almakhadmeh, and A. Tolba, "Multiple cloud storage mechanism based on blockchain in smart homes," *Future Generation Computer Systems*, vol. 115, pp. 304–313, February 2021.

[18] A. K. Sikder, L. Babun, Z. B. Celik, A. Acar, H. Aksu, P. McDaniel, E. Kirda, and A. S. Uluagac, "Kratos: Multi-user multi-device-aware access control system for the smart home," in *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '20, p. 1–12, New York, NY, USA, 2020. Association for Computing Machinery.

[19] A. K. Sikder, L. Babun, Z. B. Celik, H. Aksu, P. McDaniel, E. Kirda, and A. S. Uluagac, "Who's Controlling My Device? Multi-User Multi-Device-Aware Access Control System for Shared Smart Home Environment," *ACM Transactions on Internet of Things*, vol. 3, pp. 1–39, November 2022.

[20] H. Touqeer, S. Zaman, R. Amin, M. Hussain, and M. Bilal, "Smart home security: challenges, issues and solutions at different iot layers," *The Journal of Supercomputing*, vol. 77, no. 12, pp. 14053–14089, 2021.

[21] Z. Wang, M. H. Gao, L. Chen, and S. Sun, "An access control scheme based on access tree structure pruning for cloud computing," *International Journal of Network Security*, vol. 23, no. 1, pp. 143–156, 2021.

[22] P. S. Xie, H. J. Fan, T. Feng, Y. Yan, G. Q. Ma, and X. M. Han, "Adaptive access control model of vehicular network big data based on xacml and security risk.," *International Journal of Network Security*, vol. 22, no. 2, pp. 347–357, 2020.

[23] H. B. Zhang, P. C. Ma, and B. Liu, "Adaptive fine-grained access control method in social internet of things," *International Journal of Network Security*, vol. 23, no. 1, pp. 42–48, 2021.

Biography

Pengshou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Privacy Protection, Security on Internet of Vehicles, Security on Industrial Internet. E-mail: xiepsh_lut@163.com.

Pengyun Zhang was born in Dec. 1999. He is a master student at Lanzhou University of Technology. His major research field is network and information security and access control. E-mail: 2324327226@qq.com

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University

of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.edu.cn.

Minghu Zhang was born in Dec. 1986. He is an associate professor and a supervisor of Master student at Lanzhou University of Technology. His major research field is computer vision, Internet of things, network and information security technology. E-mail: zhangmh@lut.edu.cn.

Xiaoye Li was born in Oct. 1995. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 976339400@qq.com

Linge Qi was born in Oct. 1998. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: 15690755751@163.com

Improving the Transferability of Adversarial Samples through Automatically Learning Augmentation Strategies from Data

Ru-Zhi Xu and Chang-Ran Lyu
(Corresponding author: Chang-Ran Lyu)

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China
Email: 120212227100@ncepu.edu.cn

(Received Jan. 10, 2023; Revised and Accepted Aug. 5, 2023; First Online Oct. 8, 2023)

Abstract

The attackers cause the deep neural network (DNN) to misclassify the original image by adding perceptible perturbations, which brings security risks to deploying deep neural networks. Since the existing transfer-based attack algorithms overfit to the source model resulting in poor transferability of black-box attacks and data augmentation is one of the main methods to avoid overfitting the source model. We propose an auto-augmented transferable black-box attack method. Firstly, we set up a search space for data augmentation strategies, and then we use reinforcement learning to search for the best augmentation strategy automatically. We use the strategies to augment images that are used to compute gradients. Finally, we employ the fast gradient sign algorithm to generate adversarial examples. Extensive experiments on ImageNet show the superiority of our method to state-of-the-art baselines in attacking different undefended and defended models.

Keywords: Adversarial Samples; Black Box Attack; Data Augmentation; Deep Neural Network; Transferability

1 Introduction

In recent years, with the continuous development of deep learning theory and the improvement of computing power caused by graphics processing unit (GPU) technology, deep learning technology has achieved outstanding performance in various visual tasks such as image classification, speech recognition, and target detection. Deep neural networks play a vital role in industrial production and life such as face recognition, anomaly detection, and voice recognition. At the same time, more and more experts pay attention to the security of deep neural networks. Szegedy *et al.* [21] found that deep neural networks are vulnerable to adversarial sample attack, which is that the attacker adds adversarial perturbation information on the original image that is imperceptible to the human eyes and makes the neural network misjudge. Therefore, the security problem of deep neural network hinders its development in specific application fields, such as automatic driving [22], face recognition system [18], etc. A large number of studies have shown that research on adversarial samples can further improve the robustness of deep neural networks and reduce security risks for specific application areas.

According to the attacker's understanding of the model, adversarial attacks can be divided into two types: white-box attack and black-box attack. The former attacker owns specific information such as the parameters and structure of the target model, while the latter attacker doesn't have access to specific information about the model. Compared with white-box attacks, black-box attacks are more in line with real attack situations. There are usually two kinds of black-box adversarial attacks: query-based and transfer-based attacks. The query-based attack method needs to query the model multiple times to generate adversarial samples, which results in high query cost and poor concealment. In contrast, the transfer-based attack method uses the source model as the victim to launch an attack to generate adversarial samples, and directly uses the adversarial samples to attack the target model. This method has lower cost and higher concealment. Due to the difference between the source model and the target model,the existing transfer-based attack methods overfit the source model, resulting in limited attack transferability. Data augmentation is one of the main ways to avoid overfitting the source model. To improve the transferability of adversarial samples, Xie *et al.* [24] proposed to apply image transformations (random image resizing and random padding) to images with a fixed probability and a fixed magnitude. But its disadvantage is that the transformation used is relatively simple, which limits the diversity of the input samples. It is not significant on alleviating the overfitting problem of the source model. Russakovsky *et al.* [17] proposed to train an adversarial transformation network, and con-

structed a more aggressive adversarial sample that can resist adversarial transformation. Because the adversarial transformation network is composed of two layers of CNN, it can only simulate simple image transformations such as blurring and sharpening. The above studies have confirmed that data augmentation methods can alleviate the overfitting problem of the source model and improve the transferability of attacks. Our method optimizes the generation process of adversarial samples from the perspective of data augmentation.

In this paper, We use data augmentation and gradient attack algorithm to generate adversarial samples in order to improve the transferability of our attack method. The main contributions are as follows:

1) We use reinforcement learning algorithm to search for the most effective image augmentation strategy, which can effectively alleviate the overfitting phenomenon of adversarial sample to the source model and improve the transferability of the attack.

2) Unlike the classical Gradient attack algorithms, our method(Automatic data augmentation Fast Gradient Sign Method, AutoAugment-FGSM) does not maximize the loss function from the original input image. Instead, the image is augmented according to the augmentation strategy, then we feed the image into the source model to generate adversarial samples by maximizing the loss function.

3) Extensive experiments on the ImageNet dataset show that compared with the existing attack algorithms that used data augmentation to mitigate overfitting to source models, our method had a higher black-box attack success rate.

2 Related Work

Szegedy *et al.* [21] found that a clean sample which was added small perturbations will be misclassified by a CNN with high classification accuracy. The images on which the attacker adds carefully calculated perturbations are called adversarial samples, and the attack is called adversarial attack. In an image classification task, consider a deep learning model $f(x) = y^{true}, x \in R^m$ as input to the model, $y^{true} \in Y$ is the correct output of the model for the current input. The adversarial attack is that the attacker adds a carefully calculated perturbation r to the original input image x. In the untargeted attack, the attacker intends to make$f(x+r) \neq y^{true}$. In the targeted attack, the attacker intends to make $f(x+r)$=y^t. y^t is the target class for the attacker, $(x+r)$ represents an adversarial sample. The added perturbation is a small value, which is imperceptible to the human eyes. It is usually limited by l_0, l_2 and l_∞. In targeted attacks, the optimization problem for generating adversarial samples is given by the Equation (1).

$$\begin{array}{ll} \min & \|r\|_2 \\ \text{s.t.} & 1.f(x+r) = y^t \\ & 2.x + r \in R^m \end{array} \tag{1}$$

2.1 Adversarial Attack

Szegedy *et al.* [21] first proposed the L-BFGS method to generate adversarial samples, but it has the problem of large amount of calculation. Goodfellow *et al.* [7] proposed the Fast Gradient Sign Method (FGSM), which makes the deep learning model misclassify by adding a certain size of perturbation to the input sample in the direction of the gradient of the model loss function. Later, a series of attack methods based on gradient optimization are derived, which are all variants of FGSM algorithm. We first introduce the following three gradient-based adversarial attack methods.

The FGSM algorithm generates an adversarial perturbation in the gradient direction of the loss function, and adds the perturbation to the image to achieve the purpose of making the model misclassify. θ represents the parameters of the model, $L(x, y^{true}; \theta)$ represents the loss function of the model. x, y^{true} represent the original input image and the ground truth label. The process of FGSM generating adversarial samples is shown in Equation (2), where $\nabla x L(x, y^{true}; \theta)$ represents the direction of the gradient of the model loss function, and ϵ represents the maximum value of the adversarial perturbation. Since FGSM is a single-step attack, the attack success rate is low.

$$x^{adv} = x + \epsilon \cdot sign(\nabla x L(x, y^{true}; \theta)) \tag{2}$$

Kurakin *et al.* [11] proposed the iterative FGSM algorithm (Iterative Fast Gradient Sign Method, I-FGSM) , which reduces the optimization range and divides the perturbation into multi-step calculations. The calculation method of I-FGSM is shown by Equation (3) and Equation (4), where the $Clip(\cdot)$ operation clips the image within the legal range. The disadvantage of I-FGSM is that it is easy to fall into local extrema.

$$x_0^{adv} = x \tag{3}$$

$$x_{t+1}^{adv} = Clip_x^{\epsilon}\{x_t^{adv} + \alpha \cdot sign(\nabla x L(x, y^{true}; \theta))\} \tag{4}$$

Dong *et al.* [4] proposed the (Momentum Iterative Fast Gradient Sign Method, MI-FGSM) algorithm , which introduced momentum technology into the process of generating adversarial samples. This method stabilizes the direction of the updated gradient, avoids local extreme value in the direction of loss function gradient. MI-FGSM has great transferability. The MI-FGSM algorithm is shown in Equation (5) and Equation(6), where μ represents the decay factor of the momentum item, and g_{t+1} is the accumulated gradient at iteration $t+1$.

$$g_{t+1} = \mu \cdot g_t + \frac{\nabla x L(x_t^{adv}, y^{true}; \theta)}{\left|\left|\nabla x L(x_t^{adv}, y^{true}; \theta)\right|\right|_1} \tag{5}$$

$$x_{t+1}^{adv} = Clip_x^{\epsilon}\{x_t^{adv} + \alpha \cdot sign(g_{t+1})\} \quad (6)$$

The existing transfer-based attack methods over-fit the source model, resulting in limited attack transferability. Data augmentation is one of the main methods to avoid over-fitting the source model.From the perspective of data augmentation, our method uses reinforcement learning to search for the most effective augmentation strategy. The image is transformed using this strategy, and then input to the source model to generate adversarial examples according to the gradient of the loss function.

2.2 Defending against Adversarial Samples

Adversarial samples seriously threaten the security of deep learning models, causing people to question the reliability of deep learning models in some specific application fields. With the development of adversarial attack research, some effective adversarial attack defense methods are continuously proposed. Effective defense methods can be roughly divided into two types: adversarial training and data preprocessing.

In order to make the adversarial attack fail, the basic idea of data preprocessing is to reduce the impact of adversarial perturbation on the image, resulting in the failure of adversarial perturbation. Dziugaite *et al.* [6] found that JPEG compressed input images can effectively reduce the threat of adversarial perturbation. Since CNN is a high-dimensional network structure, even very small noises will be amplified after passing through CNN, which will have a greater impact on the output of CNN. Therefore, even the residual perturbation after denoising processing will have an impact on the classification accuracy of the deep learning model. Liao *et al.* [13] proposed a High Level Representation Guided Denoiser (HGD)defense method, which is effective in reducing the impact of residual noise. Jia *et al.* [10] proposed a defense model based on image compression (ComDefend), which successfully "purifies" the adversarial samples through compression and reconstruction. ComDefend processes the image block by block instead of the entire image, which makes processing the input image less time-consuming. Data preprocessing methods can effectively reduce the impact of adversarial perturbation, but cannot completely remove the impact of adversarial perturbation. Therefore, the model may still misclassify the preprocessed image.

The basic idea of adversarial training is to use the adversarial samples as part of the training set to train the model, so that the model acquires defended capabilities. Adversarial training can be regarded as an optimization problem of internal maximization and external minimization, as shown in Equation (7), where θ represents the neural network parameters, r represents the added adversarial perturbation, x represents the input image, and y^{true} represents the true label of the input image x, L represents the loss function.

$$\min_{\theta} \rho(\theta)$$
$$\rho(\theta) = E_{(x,y)\sim D}[\max_{\theta} L(\theta, x + r, y^{true})] \quad (7)$$

The internal max process is to maximize the loss function of the current model and find the most suitable perturbation to generate adversarial samples. The process of external min process is to train the deep learning model to find the appropriate network parameters θ. In this way, the model has a certain ability to resist interference and defend against attacks. Adversarial training is effective in defending against known attacks, but the defense against unknown attacks is poor or even impossible.

Cihang Xie *et al.* [23] proposed a randomization method (R&P) for adversarial training. This method includes two random operations: random transformation size and random filling. This method can adapt to different network structure models, and it is easy to combine with other defense methods. Jeremy Cohen *et al.* [2] proposed a new random smoothing method (RS) to perform adversarial training to improve the performance of deep learning models. This method enables the model to classify data correctly within a certain random smoothing radius, which is locally robust. In this paper, we exploit these state-of-the-art defenses to evaluate the effectiveness of our attack against defended models.

2.3 Data augmentation

Data augmentation is a technology that generates new samples from existing samples under a constraint. The data is processed by using preset data transformation rules. Data augmentation generates more data by adding prior knowledge to the original data without drastically increasing the amount of data. It improves the diversity of the data and reduces the overfitting of the model.

The transfer-based attack method uses the source model as the victim to launch an attack to generate an adversarial sample, and uses the generated adversarial sample to attack the target model. The existing transfer-based attack methods overfit the source model , resulting in limited attack transferability. Data augmentation is one of the main ways to avoid overfitting the source model. Our method optimizes the generation process of adversarial samples from the perspective of data augmentation. We use reinforcement learning to search for a set of effective data augmentation strategies, and combine gradient attack algorithms to generate adversarial samples.

3 Method

The problem of existing transfer-based attack methods is overfitting the source model, resulting in limited attack transferability. Data augmentation is one of the main ways to avoid overfitting the source model. Firstly, we propose an automatic augmentation method combined

with the gradient algorithm to find the most effective data augmentation strategy. Then, we apply the augmentation strategies to transform the image and finally use the FGSM gradient algorithm to generate adversarial sample. Our method can effectively alleviate the overfitting of adversarial samples to the source model.

3.1 Search Augmentation Strategies

Our search method consists of two parts: the search algorithm and the search space. We use reinforcement learning as the search algorithm to find the most effective data augmentation strategy. The controller of reinforcement learning is RNN, which samples the augmentation strategies S. The source model uses the strategies to transform the image. Then we use the gradient algorithm to generate adversarial samples. The white-box attack success rate of the adversarial sample attack source model as the reward, which is fed back to the controller. The controller is updated through back propagation.

A strategy in the search space includes two sub-strategy. Each sub-strategy consists of two sequential image transformations. Each image transformation is associated with two hyperparameters: 1) the probability of the transformation, and 2) the magnitude of the transformation. In each mini-batch, each image randomly selects a sub-strategy for image transformation. The 14 image augmentation operations [3] we used are Shear X/Y, Translate X/Y, Rotate, AutoContrast, Equalize, Solarize, Posterize, Contrast, Color, Brightness, Sharpness, Cutout. Each operation has a default magnitude range, as shown in Table 1.

The search algorithm is reinforcement learning. The data augmentation strategy search framework is shown in Figure 1. The training algorithm is PPO. The controller consists of a layer of LSTM and a fully connected layer, and the data augmentation strategy is generated through the softmax layer. The controller can generate a total of 30 decisions, predicting 5 sub-strategies, each sub-strategy includes two image operations, and the probability and magnitude of the operation. The controller updates the parameters according to the reward. Each sample randomly selects a sub-strategy from the five sub-strategies for image augmentation processing. Our method takes the augmented samples as the input data of the source model, and then applies the gradient algorithm to generate adversarial samples. The white-box attack success rate of the adversarial sample attack source model as the reward. Finally, we use the strategy with the highest attack success rate as the data augmentation strategy for generating adversarial samples.

3.2 Adversarial Sample Generation

In image classification tasks, data augmentation is often used to improve the generalization ability of the model. Similarly, data augmentation can be applied to generate adversarial samples, which reduces overfitting to the source model and improves the transferability of adversarial samples.

Our method uses the data augmentation strategy S learned by reinforcement learning, which includes five sub-strategies F_i, as shown in Equation (8). Each sub-strategy includes two image augmentation operations as shown in Equation (9).

$$S = \{F_1, F_2, F_3, F_4, F_5\} \tag{8}$$

$$F_i = \{oration_1^i, p_1^i, m_1^i, oration_2^i, p_2^i, m_2^i\} \tag{9}$$

Table 1: List of all image transformations

Operation Name	Range of magnitudes
Shear X/Y	[0,1]
Translate X/Y	[0.0,0.3]
Rotate	[0,30]
AutoContrast	[0,1]
Equalize	[0,1]
Posterize	[0,4]
Contrast	[0.1,1.9]
Color	[0.1,1.9]
Brightness	[0.1,1.9]
Sharpness	[0.1,1.9]
Cutout	[0,40]
Solarize	[0,110]

Firstly, We randomly select the top-2 most effective sub-strategies from the 5 sub-strategies to perform image transformation processing on each image. The transformation function of the image is shown in Equation (10). Based on this augmentation function, the optimization problem of the objective function as shown in Equation (11). Finally, we use the transformed image as input to generate adversarial samples according to the gradient algorithm. We combine the gradient algorithm (FGSM, MI-FGSM) to generate adversarial samples, and the generation process is shown in Equation (12), Equation (13) and Equation (14).

$$x_0^{adv} = A(x, F_i, F_j) \tag{10}$$

$$arg\ max\ L(A(x, F_i, F_j), y^{true}; \theta) \tag{11}$$

$$x^{adv} = x + \epsilon \cdot sign(\nabla x L(A(x, F_i, F_j), y^{true}; \theta)) \tag{12}$$

$$g_{t+1} = \mu \cdot g_t + \frac{\nabla x L(A(x, F_i, F_j), y^{true}; \theta)}{||\nabla x L(A(x, F_i, F_j), y^{true}; \theta)||_1} \tag{13}$$

$$x_{t+1}^{adv} = Clip_x^{\epsilon}\{x_t^{adv} + \alpha \cdot sign(g_{t+1})\} \tag{14}$$

Our adversarial sample generation algorithm details are described in Algorithm 1. For further research, we combine the learned augmentation strategy with the MI-FGSM algorithm for multiple iterative optimizations, and the details of the process in Algorithm 2.

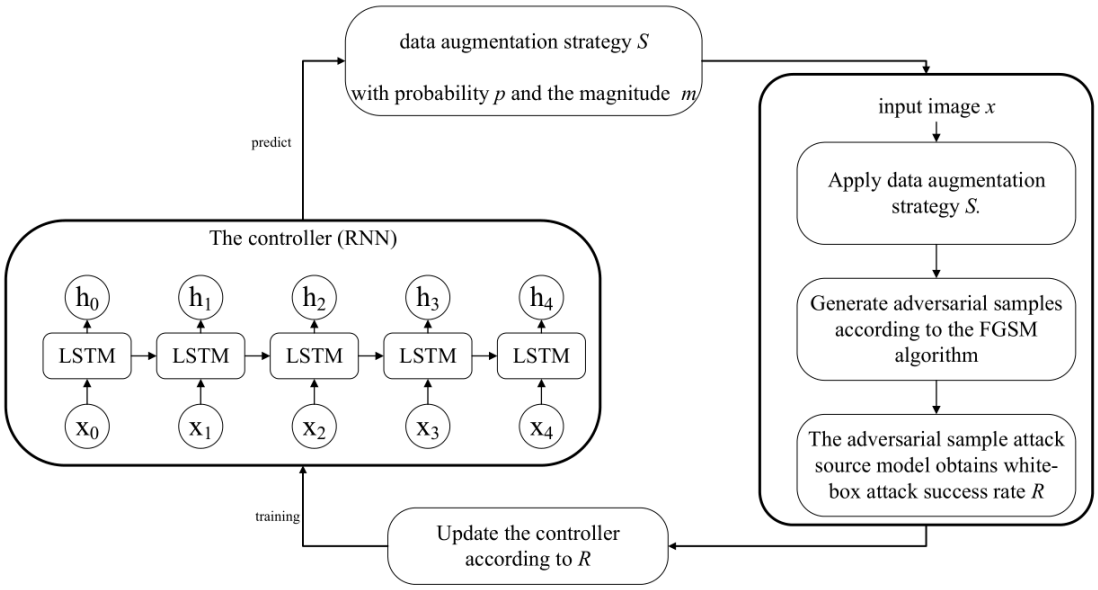


Figure 1: Overview of Neural Architecture Search

Algorithm 1 AutoAugment-FGSM

Input: A clean image x and its ground-truth label y^{true}; A classifier f; The loss function L and the perturbation budget ϵ

Output: An adversarial sample x^{adv}

Begin
Select the top-two sub-strategies F_1, F_2 from strategy S
$x^{adv} = A(x, F_1, F_2)$
Calculate the gradient $\nabla x L(A(x, F_1, F_2), y^{true}; \theta)$
$x^{adv} = x + \epsilon \cdot sign(\nabla x L(A(x, F_1, F_2), y^{true}; \theta))$
return x^{adv}

Algorithm 2 AutoAugment-MI-FGSM

Input: A clean image x and its ground-truth label y^{true}; A classifier f; The loss function L and the perturbation budget ϵ; Iterations T

Output: An adversarial sample x^{adv}

Begin
Select the top-two sub-strategies F_1, F_2 from strategy S
$x_0^{adv} = x, g_0 = 0$
For $t = 0$ to $T - 1$ **do**
    $x_{t+1}^{adv} = A(x_t^{adv}, F_1, F_2)$.
    Calculate the gradient $\widetilde{g}_{t+1}$ $\widetilde{g}_{t+1} = \nabla x L(A(x_t^{adv}, F_1, F_2), y^{true}; \theta)$.
    $g_{t+1} = \mu \cdot g_t + \frac{\widetilde{g}_{t+1}}{||\widetilde{g}_{t+1}||_1}$
    $x_{t+1}^{adv} = Clip_x^{\epsilon}\{x_t^{adv} + \alpha \cdot sign(g_{t+1})$
return $x^{adv} = x_T^{adv}$

4 Experiment

4.1 Experiment Setup

Dataset. We randomly sampled 1000 images of different categories from the ILSVRC 2012 validation set [23] that are classified correctly by all the networks which we test on. All these images are resized to 299×299×3 beforehand.

Target model. We attack both undefended and defended models. Undefended models include ResNet-v2(Res-v2) [8, 9], Inception-v3(Inc-v3) [20], Inception-v4 (Inc-v4) [19] and Inception-ResNet-v2(IncRes-v2) [19]. For defended models, we focus on several adversarially trained [16] models, including adversarially trained Inc-v3(Adv-Inc-v3) [12] and ensemble-adversarially trained IR-v2(Ens-adv-IR-v2) [12].

To further verify the attack capability of our method, we investigate some state-of-the-art defenses aimed at correcting adversarial sample. These defenses include high-level representation guided denoiser (HGD) [13], random resizing and padding (R&P) [23], feature distillation (FD) [15], compression defense (ComDefend) [10], and randomized smoothing (RS) [2]. To evaluate our attack method, we attack the above advanced defense models.

Baseline. We compare the AutoAugment-FGSM method with two classes of baseline methods. The first category is gradient sign series algorithms with extremely high attack migration ability, including FGSM [7], I-FGSM [11] and MI-FGSM [4]. The second type of attack algorithms based on image transformation, which includes DIM [24], TIM [5], SIM [14], ATTA [26], ODI-MI-TI [1],

Figure 2: Original image and corresponding adversarial samples

VMI-FGSM [25].

Parameter. we set the perturbation budget $\epsilon = 15$. The iteration numbers T was set to 18. The step size α was set to 1.0 .The decay factor μ for MI-FGSM was set to 1.0.

Table 2: Augmentation strategy of Inception-v3

Operation 1	p1	m1	Operation 2	p2	m2
AutoContrast	0.72	0.78	Rotate	0.96	0.97
ShearX	0.94	0.89	Rotate	0.96	0.76
ShearY	0.33	0.85	Rotate	0.92	0.99
Rotate	0.93	0.22	Rotate	0.71	0.07
Rotate	0.97	0.16	Rotate	0.72	0.50

4.2 Attacking a Single Network

We synthesis adversarial samples only on normally trained net-works, and test them on all six networks. The success rates are shown in Table 3. The attack success rate is the misclassification rate of an adversarial sample on the target model. Its calculation method is shown in Equation (15). The $sumNUM(\cdot)$ represents the number of samples. The higher the value of ASR, the higher the success rate of the sample attack, and the stronger the attack of the algorithm. Besides, we visualize 4 randomly selected pairs of such generated adversarial images and their clean samples in Figure 2. These visualization results show that these generated adversarial perturbations are human imperceptible. The data augmentation strategy learned by using the source model Inception-v3 is shown in Table 2.

$$ASR = \frac{sumNUM(lable(x^{adv}) \neq y^{true})}{sumNUM(lable(x) = y^{true})} \tag{15}$$

Experiments evaluate various attack methods against undefended and defended models. Specifically, our attack method utilizes the source model as the victim to generate adversarial samples. Our method directly applies this adversarial sample to attack other models, which is equivalent to a black-box attack. As shown in Table 3, * stands for white-box attack. In white-box attack, our attack method AutoAugment-FGSM can achieve 99.99% attack success rate. In the black-box setting, compared with the FGSM attack algorithm, our method improves the average attack success rate based on 4 source models by 60.26%, 30.76%, 33.22%, 46.02%. Besides, Our attack consistently outperforms all state-of-the-art base-lines under the black-box settings, which further corroborates the superiority of our strategy on generating transferable adversarial samples.

4.3 Attacking Advanced Defense Models

Experiment with the current five defense models, including HGD, R&P, FD, ComDefend, and RS. The attack success rates of the adversarial samples generated by AutoAugment-FGSM under the above five defense models are shown in Table 4. The average success rate of AutoAugment-FGSM attacking the above five defense models is 77.04. DIM improves the transferability of adversarial samples by increasing the diversity of images, which utilizes fixed image transformation probabilities and magnitudes to process images. TIM leverages translation invariance to optimize perturbations on an ensemble of transformed images to generate adversarial samples. SIM avoids overfitting by optimizing adversarial perturbations at different scales. ATTA proposes an adversarial transform network for simulated data augmentation. The gradient series algorithm can combine the above four attack methods to generate adversarial samples. Our method AutoAugment-FGSM attack success rate is higher than the above four attack methods 35.94, 22.26, 12.96, 10.68. Experiments further demonstrate the effectiveness of AutoAugment-FGSM attack undefended and defended models. This research further raises new security concerns in researching stronger defense methods.

4.4 Further Analysis

We further investigate the influence of the size of the maximum perturbation ϵ on the success rate of the

Table 3: The success rates(%) on six networks where we attack a single network

Model	Attack	Res-v2	Inc-v3	Inc-v4	IncRes-v2	Adv-Inc-v3	Ens-adv-IR-v2
Res-v2	FGSM	57.7*	24.6	16.5	11.4	22.0	6.0
	BIM	91.0*	30.4	18.7	15.2	23.9	7.6
	MI-FGSM	98.2*	41.9	35.6	28.5	23.5	7.2
	DIM	97.8*	78.0	70.7	71.7	39.8	19.4
	TIM	98.8*	65.2	59.8	57.4	35.5	32.8
	SIM	98.8*	67.3	57.4	57.4	38.3	26.7
	ATTA	99.8*	64.3	61.8	59.2	35.5	18.5
	ODI-MI-TI	**99.9***	70.1	45.6	67.3	30.7	16.9
	VMI-FGSM	**99.9***	74.9	70.1	66.1	44.2	34.2
	AutoAugment-FGSM	**99.9***	**80.2**	**73.4**	**88.6**	**74.2**	**65.4**
Inc-v3	FGSM	31.6	84.3*	43.9	40.1	32.1	11.9
	BIM	21.7	93.8*	30.4	26.8	28.8	8.5
	MI-FGSM	37.0	**99.9***	69.0	60.4	26.9	10.2
	DIM	34.4	95.9*	61.9	53.9	39.8	19.4
	TIM	39.2	**99.9***	44.3	45.8	23.2	25.8
	SIM	40.1	**99.9***	42.9	46.4	22.8	33.5
	ATTA	44.8	**99.9***	52.9	53.2	25.1	14.4
	ODI-MI-TI	50.8	**99.9***	42.5	57.6	26.3	18.8
	VMI-FGSM	59.2	**99.9***	66.5	59.6	32.8	17.5
	AutoAugment-FGSM	**62.1**	**99.9***	**69.8**	**62.4**	**65.8**	**53.3**
Inc-v4	FGSM	27.7	44.9	65.9*	31.2	28.1	9.0
	BIM	24.0	45.5	91.6*	28.9	29.1	7.6
	MI-FGSM	36.9	65.3	**99.9***	55.5	27.9	8.6
	DIM	43.8	68.5	97.3*	58.9	28.8	12.4
	TIM	41.4	64.3	99.6*	48.2	29.6	28.7
	SIM	41.4	61.9	99.6*	49.7	31.3	42.9
	ATTA	43.8	66.8	99.6*	59.2	35.5	15.6
	ODI-MI-TI	52.4	48.8	**99.9***	54.3	22.7	19.6
	VMI-FGSM	58.6	67.9	**99.9***	59.2	38.2	23.2
	AutoAugment-FGSM	**61.1**	**68.8**	**99.9***	**59.2**	**63.3**	**54.6**
IncRes-v2	FGSM	24.4	40.4	29.6	50.6*	26.6	10.0
	BIM	23.2	44.5	33.3	91.0*	31.2	7.9
	MI-FGSM	35.2	67.3	64.5	**99.9***	30.3	9.8
	DIM	46.6	75.2	73.3	98.3*	36.5	18.0
	TIM	43.1	62.9	55.4	98.9*	37.9	49.3
	SIM	42.1	60.9	52.7	98.9*	35.4	57.1
	ATTA	44.8	68.9	65.2	98.9*	39.5	22.4
	ODI-MI-TI	60.0	70.1	45.6	**99.9***	32.7	17.9
	VMI-FGSM	63.3	77.9	72.1	**99.9***	40.8	34.4
	AutoAugment-FGSM	**67.9**	**80.2**	**73.4**	**99.9***	**74.2**	**65.4**

Table 4: Success rates (%) of different attacks against advanced defense methods.

Attack	HGD	R&P	FD	ComDefend	RS	Average
FGSM	21.2	18.1	52.2	66.6	47.5	41.1
BIM	16.8	17.9	50.0	63.4	48.9	39.4
MI-FGSM	17.9	18.7	56.7	74.6	49.2	39.4
DIM	16.5	23.1	70.5	66.6	28.8	41.1
TIM	45.1	45.2	**78.2**	69.2	36.2	54.8
SIM	71.5	57.6	76.6	75.4	39.3	64.1
ATTA	70.0	58.8	75.3	79.8	47.9	66.4
ODI-MI-TI	68.5	54.2	66.3	68.2	38.2	59.1
VMI-FGSM	70.1	62.3	72.4	80.3	51.9	67.4
AutoAugment-FGSM	**72.0**	**64.6**	78.0	**81.7**	**88.9**	**77.0**

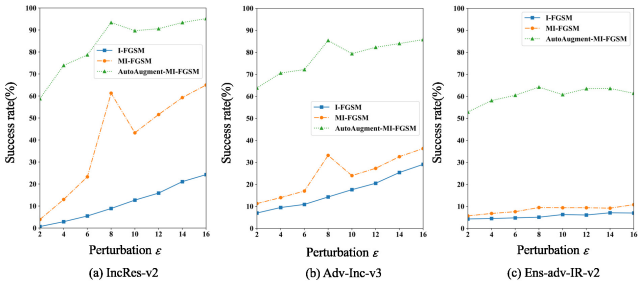


Figure 3: The success rates of I-FGSM, MI-FGSM and AUTO-MI-FGSM when T=4

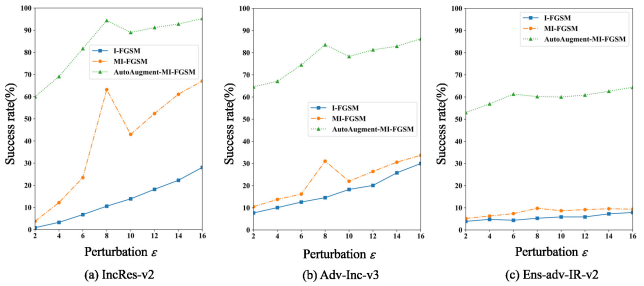


Figure 4: The success rates of I-FGSM, MI-FGSM and AUTO-MI-FGSM when T=8

attack. We combined the image augmentation strategy learned by Inc-v3 with the MI-FGSM gradient algorithm to generate adversarial samples. The method is called AutoAugment-MI-FGSM. AutoAugment-MI-FGSM, I-FGSM, MI-FGSM attack IncRes-v2, Adv-Inc-v3, Ens-adv-IR-v2 models to calculate the attack success rate. In this experiment, the size of the perturbation ϵ increases from 2 to 16 in steps of 2. Figure 3 and Figure 4 respectively show the attack success rates of iteration T=4 and T=8 on IncRes-v2, Adv-Inc-v3, and Ens-adv-IR-v2. As ϵ increases, the trend of attack success rate of I-FGSM is an upward trend, while the trend of MI-FGSM and AutoAugment-MI-FGSM is first up and then down, and the inflection point is around $\epsilon = 8$. Compared with other methods, our method AutoAugment-MI-FGSM has the highest attack success rate. The results show that AutoAugment-MI-FGSM is highly aggressive.

5 Conclusion

In this paper, we propose an adversarial example generation method(AutoAugment-FGSM) based on image enhancement. Firstly, our method regards the source model as the victim and uses reinforcement learning to search for the most effective image augmentation strategy. We select the top-2 sub-strategies for image transformation on clean samples. Finally, we use the FGSM gradient algorithm to generate adversarial sample. Our method effectively alleviates the problem of overfitting the source model. Moreover, our strategy can be conveniently combined with other transfer-based attacks to further promote their performance. Experimental results show that compared with the traditional FGSM series gradient algorithms, our method greatly improves the success rate of black box attacks. Demonstrate the superiority of our method synthesize adversarial samples attacking both state-of-the-art undefended and defended models. Our study further provokes the investigation of new security issues for more robust defense methods.

Acknowledgments

This work was supported by the Natural Science Foundation of China (No.61972148). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] J. Byun, S. Cho, M.-J. Kwon, H.-S. Kim, and C. Kim, "Improving the transferability of targeted adversarial examples through object-based diverse input," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 15 244–15 253.

[2] J. Cohen, E. Rosenfeld, and Z. Kolter, "Certified adversarial robustness via randomized smoothing," in *International Conference on Machine Learning*. PMLR, 2019, pp. 1310–1320.

[3] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "Autoaugment: Learning augmentation strategies from data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 113–123.

[4] Y. Dong, F. Liao, T. Pang, H. Su, J. Zhu, X. Hu, and J. Li, "Boosting adversarial attacks with momentum," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9185–9193.

[5] Y. Dong, T. Pang, H. Su, and J. Zhu, "Evading defenses to transferable adversarial examples by translation-invariant attacks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4312–4321.

[6] G. K. Dziugaite, Z. Ghahramani, and D. M. Roy, "A study of the effect of jpg compression on adversarial images," *arXiv preprint arXiv:1608.00853*, 2016.

[7] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," *arXiv preprint arXiv:1412.6572*, 2014.

[8] 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*, 2016, pp. 770–778.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*. Springer, 2016, pp. 630–645.

[10] X. Jia, X. Wei, X. Cao, and H. Foroosh, "Comdefend: An efficient image compression model to defend adversarial examples," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 6084–6092.

[11] A. Kurakin, I. J. Goodfellow, and S. Bengio, "Adversarial examples in the physical world," in *Artificial intelligence safety and security*. Chapman and Hall/CRC, 2018, pp. 99–112.

[12] A. Kurakin, I. Goodfellow, S. Bengio, Y. Dong, F. Liao, M. Liang, T. Pang, J. Zhu, X. Hu, C. Xie *et al.*, "Adversarial attacks and defences competition," in *The NIPS'17 Competition: Building Intelligent Systems*. Springer, 2018, pp. 195–231.

[13] F. Liao, M. Liang, Y. Dong, T. Pang, X. Hu, and J. Zhu, "Defense against adversarial attacks using high-level representation guided denoiser," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1778–1787.

[14] J. Lin, C. Song, K. He, L. Wang, and J. E. Hopcroft, "Nesterov accelerated gradient and scale invariance for adversarial attacks," *arXiv preprint arXiv:1908.06281*, 2019.

[15] Z. Liu, Q. Liu, T. Liu, N. Xu, X. Lin, Y. Wang, and W. Wen, "Feature distillation: Dnn-oriented jpeg compression against adversarial examples," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2019, pp. 860–868.

[16] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," *stat*, vol. 1050, p. 9, 2017.

[17] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[18] Y. Sun, Q. Xu, Y. Li, C. Zhang, Y. Li, S. Wang, and J. Sun, "Perceive where to focus: Learning visibility-aware part-level features for partial person re-identification," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 393–402.

[19] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.

[20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[21] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," in *2nd International Conference on Learning Representations, ICLR 2014*, 2014.

[22] L. Tabelini, R. Berriel, T. M. Paixao, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "Polylanenet: Lane estimation via deep polynomial regression," in *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021, pp. 6150–6156.

[23] C. Xie, J. Wang, Z. Zhang, Z. Ren, and A. Yuille, "Mitigating adversarial effects through randomization," in *International Conference on Learning Representations*, 2018.

[24] C. Xie, Z. Zhang, Y. Zhou, S. Bai, J. Wang, Z. Ren, and A. L. Yuille, "Improving transferability of adversarial examples with input diversity," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2730–2739.

[25] X. Wang and K. He, "Enhancing the transferability of adversarial attacks through variance tuning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 1924–1933.

[26] W. Wu, Y. Su, M. R. Lyu, and I. King, "Improving the transferability of adversarial samples with adversarial transformations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 9024–9033.

Biography

Ru-Zhi Xu is a professor at the School of Control and Computer Engineering, North China Electric Power University. Her research directions are smart grid and AI security.

Chang-Ran Lyu is a master student at the School of Control and Computer Engineering, North China Electric Power University. Her research directions are smart grid and AI security.

Effective Blockchain Assisted Certificateless Proxy Re-encryption Scheme for Medical Internet of Things

Xin Qi[1], Shu Wu[1], and Shu-Hao Yu[1,2]
(Corresponding author: Shu Wu)
School of Electronic and Information Engineering, West Anhui University[1]
Intelligent Connected Vehicle Key Technology Laboratory[2]
Lu'an, China
Email: wuyshu@126.com
(Received Mar. 15, 2023; Revised and Accepted Sept. 22, 2023; First Online Oct. 8, 2023)

Abstract

With the rapid development of the medical Internet of Things, there is a growing trend of sharing medical data via the cloud. However, in the current complex medical IoT environment, users' medical data is at risk of attack and leak during transmission. Traditional encryption schemes involve complex key management protocols, making them unsuitable for data sharing. To protect medical data privacy, this paper proposes a certificateless proxy re-encryption scheme based on blockchain. This scheme enhances the mutual trust of medical parties by making blockchain tamper-proof while protecting sensitive medical privacy data from the prying eyes of honest and interested servers. Our scheme effectively guarantees data security and privacy between medical users and service providers. Furthermore, experiments on the JPBC library demonstrate that our proposed scheme has advantages in computational and communication complexity compared to the comparison scheme.

Keywords: Blockchain; Certificateless Public Key Encryption; Privacy Security; Proxy Re-encryption

1 Introduction

1.1 Background

In the past, when people needed medical care, they often had to go to a nearby hospital. For better medical care, they may have to travel farther, which is inconvenient for most people. With the development of internet and medical service, people use proxy re-encryption technology to combine with medical service system [12,33,42]. The earliest proxy re-encryption concept was proposed by Blaze [4]. People don't need to go to the hospital, they can directly send their data to the doctor over the network. In this way, users can not only get answers faster and more convenient, but also choose the doctor they want to serve them [25, 28].

Nowadays, more and more people choose to conduct medical services on the medical internet. While providing convenience, online medical service also faces many problems. For example, when users transmit their health information, they do not want others to see their private information [23]. What's more, data may also be stolen or tampered with in the process of transmission. What's worse, in order to save energy, the malicious cloud server may return part of the search results [31, 32], or even the wrong data, as proposed by Hewa [20]. Therefore, the privacy and security of information become particularly important in the transmission process. In addition, the time is also urgent when the patient consults the treatment results. If the transmission time is too long, the medical process will be delayed, thus delaying the patient's treatment. Therefore, in the process of data transmission, it requires efficient transmission and the protection of users' privacy information. For example, solutions [35,43] are often complex and time-consuming, and are not suitable for use in the medical IoT [27].

Ensuring the safety and confidentiality of online medical consultations is crucial to avoid tampering of consultation information or leakage of patient conditions, which can have serious and irreversible consequences if left unaddressed. To tackle this issue, we have developed a design scheme and conducted research to enable patients to securely and confidentially consult their medical conditions with their preferred doctors online, and receive accurate feedback in a timely manner. Our solution also addresses the privacy and anonymity of patients during the inquiry process, and takes into account the importance of speed in emergency situations where time is of the essence. By implementing our scheme, patients can have peace of mind knowing that their medical information is protected, and they can receive the care and attention they need from

their trusted medical professionals. In this paper, a certificateless proxy re-encryption scheme is proposed to ensure the security and privacy of the data between the medical user and the medical service provider. At the same time, the communication cost and computing cost in the process of data transmission should be as small as possible.

1.2 Blockchain

Blockchain is an evolving, untampered, shared ledger of data [37], a chain of blocks. Each block holds a certain amount of information, which is linked in the chronological order of their respective generation [7–9]. This information is kept on all the servers. These servers, called nodes in the blockchain system, provide storage and computing power for the entire blockchain system. To modify information in a blockchain, it is extremely difficult to tamper with information in a blockchain because more than half of the nodes must agree and modify information in all of the nodes, which are often in the hands of different actors [10, 30, 38].

Compared with traditional networks, blockchain has two core features: one is hard to tamper with data, and the other is decentralized. Based on these two characteristics, the information recorded by blockchain is more authentic and reliable, which can help solve the problem of mutual distrust between people. Blockchain will increase trust and the integrity of transaction information flow between participating nodes [46]. Simultaneously, blockchain data sharing technology has also been widely applied [21, 40, 45].

1.3 Contribution

In the medical Internet of Things environment, this paper focuses on how to ensure the security of users' sensitive medical data, so as to effectively protect the privacy of patients' medical data in the medical IoT environment. Specifically, our major contributions are as follows:

- We propose a healthcare data privacy protection scheme based on blockchain and proxy re-encryption. This scheme uses certificateless proxy to re-encrypt, hash the user's Identity, and then sends the hash data together with the encrypted ciphertext. In our scheme, the privacy and security of the user's information and identity are effectively protected. What's more, the authorized user can also know the sender's user information accurately by comparing the hash value.

- We propose a solution to store data based on blockchain.Users can transfer their re-encryption keys to the blockchain database, which use blockchain to ensure that the user agent's re-encryption keys can't be tampered with by attackers and semi-honest cloud agents, which solves the semi-honest problem of cloud agents and the possibility that the agent's re-encryption keys can be replaced.

- Compared with the comparison scheme, our scheme achieves more security objectives while protecting user privacy. Meanwhile, experiments on JPBC library show that our scheme has more advantages in computational complexity and communication complexity.

2 Related Work

The medical system usually deal with critical information, so health care is one of the most important sensitive areas that requires a lot of interest in privacy, security, access control. The security of such a sector must be guaranteed in all aspects of the stage:transmission, storage and operation. In this section, we describe some work records that deal with medical security using blockchain technology. At present, many solutions are proposed for the network medical system. Jiang *et al.* [26] designed a lightweight privacy protection proxy data transmission scheme, which realized data and identity privacy protection, but it had high computational complexity and secret key hosting problems. Liang *et al.* [29] proposed an efficient remote health monitoring system with low delay and low cost, but it did not consider the problem of data security transmission. Dwivedi *et al.* [15] proposed a novel blockchain framework to protect private medical IoT devices. They used lightweight encryption to encrypt data and verify transactions. However, they have yet to implement the system. Boonyarattaphan *et al.* [5] proposed a framework containing effective protocols based on e-health services. Their scheme automatically adapted authentication technology and corresponding different encryption algorithms to process data when two risks occur in the system, but it did not protect patient identity privacy. Literature [14, 34] intended to obtain and share medical data through Ethereum platform and intelligent contract to solve such serious problems as fragmentation of medical data, low sharing efficiency, insecure transmission process, insufficient data integrity verification, and insufficient privacy information protection. But it was not realized at that time.Hwang *et al.* [24], any t1 original signers can delegate the signing capability to the proxy group. Alam [3] applying blockchain to the Internet of Things. Hamian [19] applying blockchain to biometrics. And we're going to apply blockchain to the medical Internet of things.

Esposito *et al.* [16] comprehensively analyzed the use of blockchain to protect medical data in the cloud. The actual challenges and the work ahead were also indicated. Abu-Elezz *et al.* [1] elaborated on the benefits and threats of blockchain technology in healthcare. Schemes [2, 17, 22] either rely on agents to fully manage their data or require agents to always delegate data users' decryption keys online, which means that data transparency and timeliness are impossible to achieve and need to always be online.

Chakraborty *et al.* [6] proposed a framework for healthcare system design based on blockchain and IoT. Rifi *et al.* [39] suggested using Ethereum to manage contracts for access systems to medical data. It was used to process medical sensors based on IPFS and store data in an off-chain database. However, there were no implementation details for these last two works. In a previous study, the authors proposed an attribute-based encryption system based on a combination of blockchain and IPFS to store and protect EMRs [41]. However, in their work, they did not integrate IoT devices and realize that the system has not yet completed debugging. Another work proposes a secure intrusion detection method that used blockchain to detect a physical system classification model of information in healthcare [36]. The Chong *et al.* [11] demonstrate that how the replay attack and known key attack can be defeated.

In this paper, a new privacy protection scheme is proposed for sensitive medical data of patients under the environment of medical Internet of Things. In our scheme, only authorized doctors can view the medical data of patients. The cloud server, as the agent, cannot obtain the encrypted plaintext, which avoids the disclosure of patient privacy by the cloud server and effectively protects the privacy of users. In addition, the use of blockchain ensures that sensitive medical data cannot be tampered with and improves the security and availability of data.

3 System Model

In our proposal, we put forward a complete set of system architecture model, including: patient, cloud service system, blockchain, doctor. In addition, a series of threat hypothesis models are proposed for our system structure, and different attack modes of different attackers are used to verify our model:

3.1 System Structure

The system model consists of four entities, which are cloud service system, blockchain database, doctor. Specifically, our system model is designed as shown in Figure 1.

Patient: The patient encrypts private information, such as their health data and medical conditions, and sends it to the cloud server. Additionally, the patient generates and sends their own proxy key to the blockchain database.

Cloud service system: This is a large semi-honest intermediate shared database that collects most of the patient's various encrypted privacy, such as a variety of physical health indicators. It downloads the corresponding proxy key from the blockchain database and can carry out the semi-honest proxy encryption process.

Blockchain: This is a public database of data records that is difficult to tamper with. Messages from this database are completely trusted.

Doctor: The doctor can download the corresponding patient's private information from the cloud service, decrypt it, and then diagnose the patient's condition.

3.2 Threat Model

In our system model, we assume that the cloud service database provides proxy re-encryption service for patients in accordance with relevant privacy regulations. However, we also consider that the cloud service agent may not be entirely trustworthy and may not provide honest services to patients. Under these assumptions, an attacker could launch various attacks on the proxy re-encryption key, ciphertext, and re-encryption ciphertext. These attacks include tampering, identity forging, eavesdropping, and information injection. Based on the above assumptions, it is consider that an attacker who might launch the following attacks:

- **Man-in-the-middle attack**: When the patient encrypts the data to the cloud service system, there is a risk that the attacker will eavesdrop and tamper with the data in the transmission process, and the information downloaded by the doctor from the cloud database is also at risk.

- **Impersonation attack**: During the transmission of ciphertext and re-encrypted ciphertext, the forger may intercept intelligence and then forge false identities to transmit wrong identity information to the recipient, resulting in serious consequences if the transmitter does not get the corresponding response.

- **Tampering attack**: When the cloud server performs proxy re-encryption of ciphertext, the obtained proxy key may be tampered with. In the process of re-encryption, the wrong re-encryption key is used, resulting in the failure of the subsequent receiver to decrypt the re-encrypted ciphertext.

- **Server spoofing attack**: Cloud agents are not completely trustworthy, and cloud agents may tamper with the agent key, thus transmitting wrong information.

3.3 Design Objective

For the medical Internet of Things model, we mainly set the following goals, so as to better realize our scheme.

- In the process of consultation with doctors, we need to solve relevant identity verification problems, so that patients can better choose the doctors they want to consult, and doctors can better identify patients when they download relevant patient information from the cloud database.

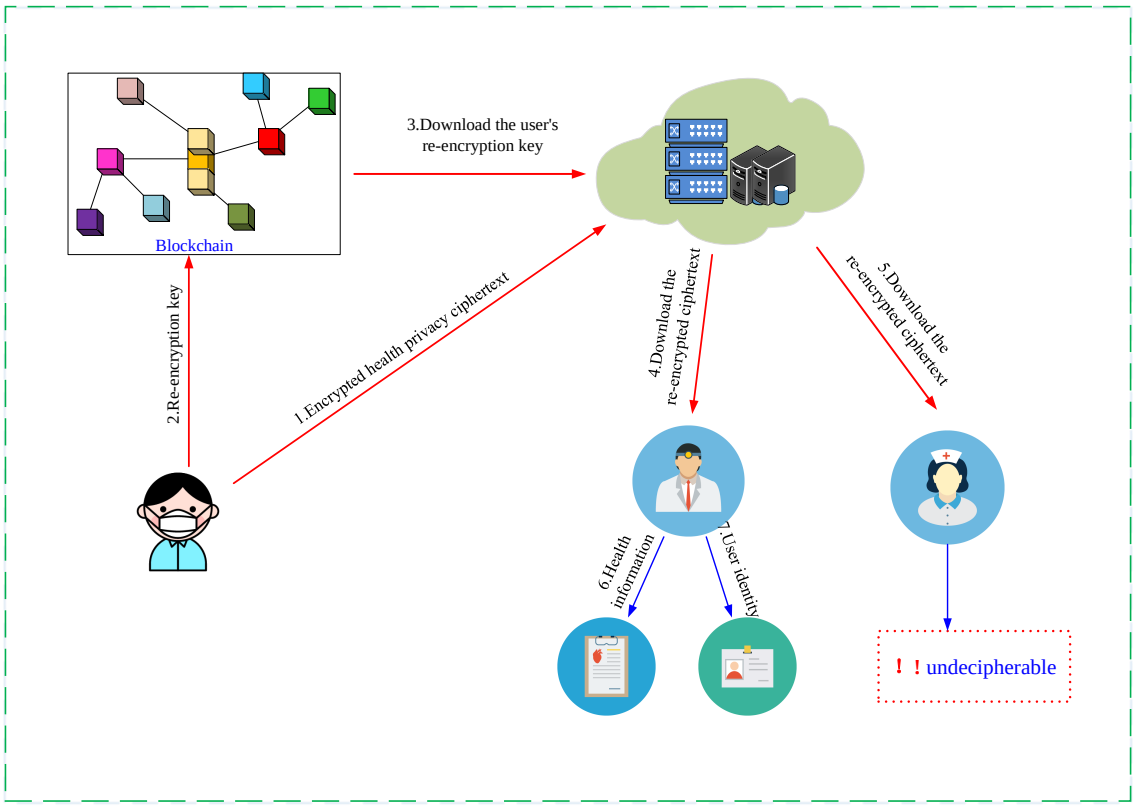


Figure 1: Systems model

- Ensure that the data transferred by the patient to the cloud service database and downloaded by the doctor are complete, so that the patient can send the correct information and the doctor can give the correct judgment to the patient.

- Our scheme uses bilinear mapping to ensure absolute confidentiality in the process of data transmission, so that different attacks cannot decipher the patient's health data, so as to avoid serious consequences.

- Some patients do not want to reveal their true identity when they send their data to the doctor. Our solution also addresses this problem, ensuring patient privacy and anonymity, addressing patient concerns, and enabling patients to better consult their doctors.

- In the previous proxy re-encryption schemes, the security of the proxy re-encryption key cannot be well guaranteed, and the cloud agent cannot be trusted completely. However, in the scheme we designed, we need to solve the problems encountered by the agent in the transmission process of the re-encryption key, and solve the problem that the cloud agent cannot be completely trusted. In this way, the security of agent re-encryption key and decryption of final ciphertext and re-encryption ciphertext are guaranteed.

4 Protocol Process

Table 1 is our specific interpretation of some elements in the protocol, and also some elements generated during the initialization of the entire system.

In order to effectively improve data privacy under the medical Internet of Things, we proposes a kind of classification proxy re-encryption algorithm which can resist the selective ciphertext attack.Our protocol uses bilinear mapping in order to better guarantee the security of the protocol. What's more, our protocol addresses the issue of distrust between patients, doctors, and cloud servers, thus well protecting privacy in the medical Internet of Things. Specifically, our agreement consists of six phases, namely setup phase, key generation phase, encryption phase, re-encrypt key generation phase, re-encryption phase and decryption phase. Which are described as follows.

4.1 $setup(1^\lambda)$

The patient needs to encrypt the ciphertext, and the algorithm outputs a pair of linear maps after entering security parameters during data transmission$(g, \mathbb{G}_1, \mathbb{Z}_r, \hat{e}, H)$, $H : \{0,1\}^* \rightarrow \{0,1\}^\lambda$ is a cryptographically unidirectional collusion-resistant hash function, instantiated using SHA-256, etc. Let g_1 be the generator different from $\mathbb{G}_1$ in g, and get the public parameter $param = (g, \mathbb{G}_1, \mathbb{Z}_r, \hat{e}, g_1, H)$.

Table 1: System parameters

Parameters	Description
$\hat{e}$	A bilinear mapping
$\mathbb{G}_1$	A cyclic additive group
$\mathbb{G}_T$	A cyclic multiplicative group
$\mathbb{Z}_r$	A cyclic group of integers
g	The generator on $\mathbb{G}_1$
g_1	Another generator on $\mathbb{G}_1$
i	The patient i
j	The doctor j
pk	The public key
sk	The private key
$rk_{i\to j}$	The re-encryption key between i and j
C_i	Encrypted ciphertext
C_j	Re-encrypt ciphertext
m	Plaintext
$\mathbb{G}_1 * \mathbb{G}_1 \to \mathbb{G}_T$	The pairing function
$H : \{0,1\}^* \to \mathbb{G}_T$	A map to point Hash function

4.2 $KegGen(param) \to (pk, sk)$

Input system exposure $param$, The algorithm selects two different random numbers x_{id} in Z_r,The public keys of patient i and doctor j are set as $pkid_1 = g^{x_{id}}$ and $pkid_2 = g_1^{x_{id}}$ respectively and disclosed to the public. The private keys are set to $skid = x_{id}$ and used for a long time.

4.3 $Enc(param, pki_1, pkj_2,) \to c_i$

Enter system exposure $param$, Entity public key and plaintext m, The first encrypted ciphertext is obtained. The encryption process is as follows:

The patient i randomly selects a random number $u \in Z_r$ and uses his public key and doctor j's public key to calculate the ciphertext as follows:

$$C_i = (a, b, c) = (g^u, me(pkj_2, pki_1)^u, H(H(a)||H(b))) \quad (1)$$

After encrypting the plaintext, the patient i obtains ciphertext C_i and C_i to the cloud service system. After receiving ciphertext C_i, the cloud server stores it and exposes it in its own database.

4.4 Re$KeyGen(param, ski, pk) \to rk_{i\to j}$

Enter system exposure $param$, Entity public key, the process of generating the re-encryption key is as follows:

The patient i then randomly selects a random number $v \in Z_r$ and calculates the re-encryption key through the following formula:

$$rk_{i\to j} = (X, Y) = (g^v, pkj_2^{-ski} pkj_1^v) \quad (2)$$

After calculating the re-encryption key $rk_{i\to j}$, patient i sends the re-encryption key to the blockchain database for public disclosure and storage.

4.5 Re$Enc(param, c_i, rk_{i\to j}) \to C_j$

Enter system exposure $param$, Ciphertext C_i and $rk_{i\to j}$, the re-encryption ciphertext is generated as follows:

After receiving the encrypted ciphertext C_i stored by patient i in its own database, the cloud server downloads the re-encryption key placed by the patient i from the blockchain database. Then it obtains the re-encrypted ciphertext by the following calculation (3).

After calculating the re-encrypted ciphertext C_j, the cloud server saves C_j and shares it in its own service system.

$$\begin{aligned} c_j &= (a', b', c', d') \\ a' &= a \\ b' &= b \cdot \hat{e}(a, Y) \\ c' &= X \\ d' &= H(H(a')||H(b')||H(c')) \end{aligned} \quad (3)$$

4.6 Decrypted Ciphertext

For patient i, ciphertext C_i can be obtained from the cloud shared database and decrypted. The decryption process is as follows:

The patient i first verifies that equation $c = H(H(a)||H(b))$ is true, and if it is not, outputs $\perp$. Otherwise, user i decrypts and outputs the message plaintext through the following calculation:

$$m \leftarrow \frac{b}{\hat{e}(a, pkj_2)^{ski}} \quad (4)$$

After decryption, patient i can get the health data uploaded by herself and the privacy of her medical condition.

For the doctor j, the re-encrypted ciphertext can be obtained from the cloud shared database, and the re-

encrypted ciphertext C_j can be decrypted. The decryption process is as follows:

For ciphertext C_j after re-encryption, the doctor j first verifies whether equation $d' = H(H(a')||H(b')||H(c'))$ is true. If not, output $\perp$. Otherwise, the doctor j decrypts and outputs the message plaintext through the following calculation:

$$m \leftarrow \frac{b'}{\hat{e}(a', c')^{skj}} \tag{5}$$

After successful decryption, the doctor j can obtain relevant health data and disease privacy of patient i, and make corresponding diagnosis for patient i according to the data.

4.7 Ciphertext Decryption

We deduce the decryption formulas of formula (4) and (5) respectively, and verify that the decryption method of this algorithm can successfully decrypt the plaintext m. The derivation process is (6) and (7) respectively. The analysis is as follows:

Formula (6) decrypts the ciphertext sent by patient i. The decryption process is as follows:

$$\begin{aligned}
\frac{b}{\hat{e}(a, pkj_2)^{ski}} &= \frac{m\hat{e}(pkj_2, pki_1)^u}{\hat{e}(g^u, pkj_2)^{ski}} \\
&= \frac{m\hat{e}(pkj_2, pki_1)^u}{\hat{e}(g^u, pkj_2)^u} \\
&= \frac{m\hat{e}(pkj_2, g^u)^u}{\hat{e}(g^u, pkj_2)^u} \\
&= \frac{m\hat{e}(pkj_2, g)^{uu}}{\hat{e}(g, pkj_2)^{uu}} \\
&= m
\end{aligned} \tag{6}$$

Formula (7) is for doctors j to download the corresponding agent re-encrypted ciphertext from the cloud server and decrypt it accordingly. The decryption process is as follows:

$$\begin{aligned}
\frac{b'}{\hat{e}(a', c')^{skj}} &= \frac{b \cdot \hat{e}(a, Y)}{\hat{e}(a, X)^{skj}} \\
&= \frac{m\hat{e}(pkj_2, pki_1)^u \hat{e}(g^u, pkj_2^{-ski} pkj_1^v)}{\hat{e}(g^u, g^v)^{skj}} \\
&= \frac{m\hat{e}(pkj_2, pki_1)^u \hat{e}(g^u, pkj_1^v)\hat{e}(g^u, pkj_2^{-ski})}{\hat{e}(g^u, g^v)^{skj}} \\
&= \frac{m\hat{e}(pkj_2, pki_1)^u \hat{e}(g^u, pkj_1^v)}{\hat{e}(g^u, g^v)^{skj}\hat{e}(g^u, pkj_2^{ski})} \\
&= \frac{m\hat{e}(pkj_2, g^u)^u \hat{e}(g^u, pkj_1^v)}{\hat{e}(g^u, g^v)^v \hat{e}(g^u, pkj_2^u)} \\
&= \frac{m\hat{e}(pkj_2, g^u)^u \hat{e}(g^u, g^{v*v})}{\hat{e}(g^u, g^v)^v \hat{e}(g^u, pkj_2^u)} \\
&= \frac{m\hat{e}(pkj_2, g)^{uu} \hat{e}(g^u, g^{v*v})}{\hat{e}(g^u, g^v)^v \hat{e}(g, pkj_2)^{uu}} \\
&= m
\end{aligned} \tag{7}$$

Through the above two verification, we can know that the correctness and feasibility of the scheme can be passed.

5 Safety Analysis

The proposed protocol can mitigate the attacks in the threat model. The detailed analysis is as follows:

- **Defense against man-in-the-middle attacks**: Suppose that g and g_1 are generator on $\mathbb{G}_1$ respectively. x is a prime number on $\mathbb{Z}_r$. It's almost impossible to figure out x if an attacker know g^x and g_1^x according to discrete logarithm. Therefore, it can achieve polynomial security for ciphertext and re-encrypted ciphertext in the transmission process. Even if an attacker intercepts information in the middle, he still cannot steal or tamper with it, which can ensure that the message can be accurately transmitted to the receiver.

- **Defense against impersonation attack**: In the event of an attacker attempting to forge an identity, the receiver can verify the authenticity of the message by comparing the hash identity data in the received message. By checking whether the hash value has been transformed, the receiver can determine whether the message was sent by the original author or if it has been tampered with. This process ensures that the sender's identity is accurately determined, and that the receiver can trust the authenticity of the message.

- **Defense against tampering attack**: It is virtually impossible for an attacker to tamper with the stored proxy key in the large blockchain database or the ciphertext encrypted with elliptic curves. Any attempt to modify the ciphertext or proxy key would require the attacker to have control over a majority of the nodes in the blockchain network, which is highly unlikely due to the decentralized nature of the network. Moreover, elliptic curve cryptography provides a high level of security that makes it difficult to tamper with the ciphertext. Therefore, our proposed system provides robust protection against tampering attempts by attackers.

- **Defense against server spoofing attack**: Since the data in the blockchain database in this scheme is completely reliable and cannot be tampered by anyone, which can guarantee the security of the re-encryption key. We propose to transfer the re-encryption key to the blockchain database for preservation. In this way, the proxy key obtained by the cloud is completely correct, and the recipient can determine whether the cloud has tampered with the re-encryption key based on the data in the blockchain.

6 Performance Analysis

We have carried out experiments on the four schemes respectively on the JPBC library, and also summarize the two schemes in terms of security performance, computing cost and communication cost. We analyze and compare the experimental results of the four schemes in different degrees.

We use the Type A curves defined within the JPBC library because they are commonly used in primitive encryption. In the JPBC library, the Type A curve is chosen as $E(F_q): y^2 = x^3 + x$.The group order of $\mathbb{G}_1$ is 160bits and the order of the base field is 512 bits. So g is a 512 bits prime number and g_1 is also a 512 bits prime number. The length of the element in $\mathbb{G}_1$ is 1024 bits. The output length of hash map is 256 bits. The specific experimental is analyzed as follows.

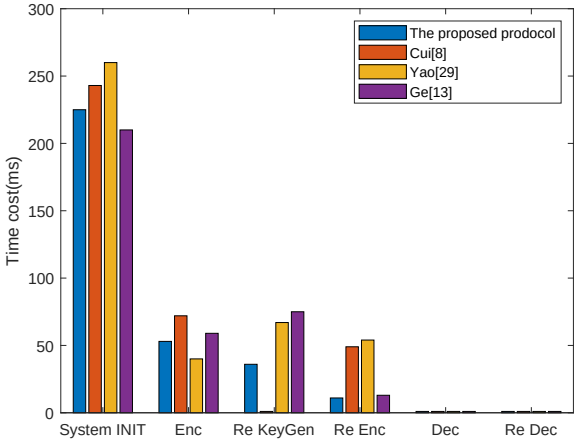


Figure 2: Computational overhead for different stages

6.1 Safety Performance Comparison

In this subsection, we compare the proposed scheme with three different proxy re-encryption schemes in terms of security performance. The comparison results are shown in Table 2.

All the schemes can achieve identity authentication, but in Cui [13], because the corresponding proxy re-encryption key is generated when both parties know each other's private key, data integrity can not be fully guaranteed, and therefore decryption of the recipient can not be guaranteed. For Yao [44], data confidentiality is not possible. In Ge [18], the privacy of the sender is not well protected, so it can not be guaranteed anonymity . In addition, they have not completely solved the problem of cloud agents' complete trust in their schemes, so they cannot guarantee the accuracy of the proxy key well. However, in our scheme, we can well guarantee the above situation, and solve the problem of semi-integrity of cloud agents. By comparing the safety performance data, our scheme has better safety performance.

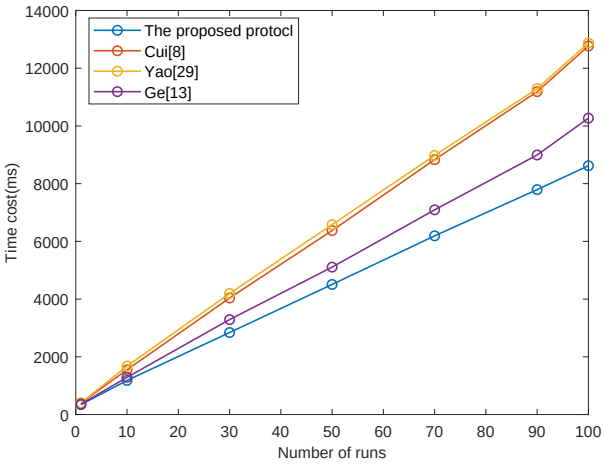


Figure 3: Total computation overhead of different runs

6.2 Computational Overhead

Our scheme is also compared with Cui [13],Yao [44],Ge [18] scheme on computational overhead. Compared with other operations, basic operations such as addition and hash are extremely fast, so the time consumed is negligible. In this paper, we mainly consider the calculation cost of dot multiplication and bilinear pair operations on elliptic curves.

As can be seen from the Figure 2: In comparison, the time required to initialize the system in our scheme is 90.7% of Cui [13], 86.5% of Yao [44], and Ge [18] is 93.5% of our scheme. The time required for encrypted ciphertext is 72.1% of Cui [13], 88.3% of Ge [18],Yao [44] is 81.1% of our scheme and respectively. Although the generation of proxy re-encryption key is slower than Cui [13], it is 22.4%, 20.3%, and 84.6% of the other three schemes in re-encrypting ciphertext, respectively. Although we may be slightly ahead of one in some parts, in the overall algorithmic data, our solution is preferred over the other three.

Furthermore, we also conducted different times of operation for the whole scheme. As the number of runs increases, the running speed of our scheme is significantly improved compared with the other three schemes (See Figure 3). The advantage in time of our scheme is more significant.

To be specific, the time required by our scheme to run the whole system is 343ms. The time taken by Cui [13] is 399ms, Yao [44] is 420ms,Ge [18] is 360ms. With the gradual increase of the overall run times, the advantage of our scheme is more obvious in the calculation cost. When the whole system runs 100 times, the calculation cost of our scheme is 67.5% of Cui [13], 67.3% of Yao [44], 83.9% of Ge [18]. So our solution is more in line with the application of the medical Internet of Things.

Table 2: Safety performance comparison

Safety performance	Our protocol	Cui [13]	Yao [44]	Ge [18]
Identity authentication	✓	✓	✓	✓
Data integrity	✓	×	✓	✓
Data confidentiality	✓	✓	×	✓
Anonymity	✓	✓	✓	×
Security of proxy keys	✓	×	×	×
Decryption security	✓	×	✓	✓

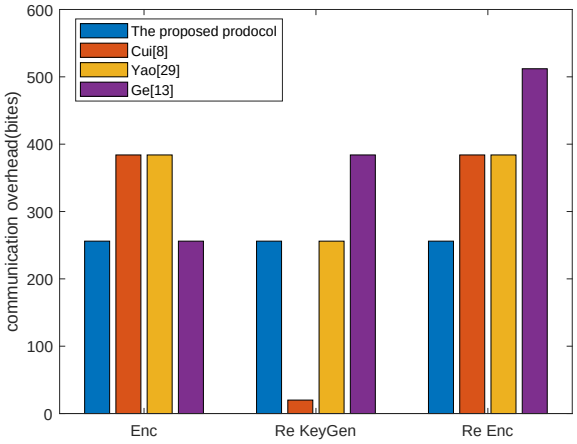


Figure 4: Communication overhead comparison

6.3 Communication Overhead

Our protocol evaluates the size of the message during delivery, where the size of the element on group $|\mathbb{G}_1|$ is 128 bytes. The size of the last element in group $|\mathbb{G}_T|$ is 128 bytes. The size of an element in group $|Z_r|$ is 20 bytes.

Through the analysis of communication cost, compared with Cui [13] and Yao [44], the ciphertext communication cost of this scheme is smaller than the size of the group element $|\mathbb{G}_T|$, which is the same as Ge [18]. Although the communication cost of reencryption key is slightly higher than Cui [13], this scheme has higher security on reencryption key agent. In addition, the re-encryption ciphertext requires the same communication cost is lower than the other three schemes (See Figure 4). By overall comparison, our scheme is close to the communication cost of Cui [13], and lower than Yao [44] and Ge [18]. But our plan has a great advantage in calculating the cost.

7 Conclusion

In this paper, we proposed a certificateless proxy re-encryption scheme for storing data based on blockchain in the medical Internet of Things. We transmitted the generated proxy re-encryption key to the blockchain, which can solve the problems such as possible tampering and cloud agents not being completely trustworthy. We have evaluated the performance and results of the protocol to show that our protocol can reduce the cost of computing overhead. It also spend less space on communication overhead. For the most part, it allows patients to get to doctors more quickly. In addition, for the third-party cloud services that are not completely trusted, this paper also proposes a cloud service system based on blockchain, so as to enhance the credibility of the third party and ensure the security of communication authentication between them.

Our future work includes coming up with new authentication protocols and authentication with higher computational efficiency and less cost.

Acknowledgments

This work was supported in part by Anhui Province Scientific Research Planning Project under Grant No.2022AH051670, in part by the Natural Science Foundation of West Anhui University under Grant No.WXZR202210. The authors are very appreciative of the anonymous reviewers for their detailed remarks and suggestions on this paper.

References

[1] I. Abu-Elezz, A. Hassan, A. Nazeemudeen, M. Househ, and A. Abd-Alrazaq, "The benefits and threats of blockchain technology in healthcare: A scoping review," *International Journal of Medical Informatics*, vol. 142, p. 104246, 2020.

[2] K. O.-B. O. Agyekum, Q. Xia, E. B. Sifah, C. N. A. Cobblah, H. Xia, and J. Gao, "A proxy re-encryption approach to secure data sharing in the internet of things based on blockchain," *IEEE Systems Journal*, vol. 16, no. 1, pp. 1685–1696, 2021.

[3] T. Alam, "A survey on the use of blockchain for the internet of things," *International Journal of Electronics and Information Engineering*, vol. 13, no. 3, pp. 119–130, 2021.

[4] M. Blaze, G. Bleumer, and M. Strauss, "Divertible protocols and atomic proxy cryptography," in *Advances in Cryptology (EUROCRYPT'98): International Conference on the Theory and Application of*

Cryptographic Techniques Espoo, Finland, May 31–June 4, 1998 Proceedings 17. Springer, pp. 127–144, 1998.

[5] A. Boonyarattaphan, Y. Bai, and S. Chung, "A security framework for e-health service authentication and e-health data transmission," in *2009 9th International Symposium on Communications and Information Technology*. IEEE, 2009, pp. 1213–1218.

[6] S. Chakraborty, S. Aich, and H.-C. Kim, "A secure healthcare system design framework using blockchain technology," in *21st International Conference on Advanced Communication Technology*, IEEE, pp. 260–264, 2019.

[7] P. Y. Chang, M.S. Hwang, C.C. Yang, "A blockchain-based traceable certification system", in *Security with Intelligent Computing and Big-data Services*, pp. 363-369, 2018.

[8] L. Chen, Q. Fu, Y. Mu, L. Zeng, F. Rezaeibagha, M. S. Hwang, "Blockchain-based random auditor committee for integrity verification", *Future Generation Computer Systems*, vol. 131, pp. 183-193, 2022.

[9] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on the secure financial surveillance blockchain systems", *International Journal of Network Security*, vol. 22, no. 4, pp. 708-716, 2020.

[10] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on blockchain technologies in bidding systems", *International Journal of Network Security*, vol. 22, no. 6, pp. 897-904, 2020.

[11] S.-F. Chiou, M.-S. Hwang, and S.-K. Chong, "A simple and secure key agreement protocol to integrate a key distribution procedure into the dss 1," *International Journal of Advancements in Computing Technology*, vol. 4, no. 19, 2012.

[12] P. S. Chung, C. W. Liu, M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments", *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[13] N. Cui and Y. Li, "A certificateless proxy re-encryption scheme based on bilinear pairs," *Information technology*, vol. 7, pp. 34–36, 2013.

[14] G. G. Dagher, J. Mohler, M. Milojkovic, and P. B. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable cities and society*, vol. 39, pp. 283–297, 2018.

[15] A. D. Dwivedi, G. Srivastava, S. Dhar, and R. Singh, "A decentralized privacy-preserving healthcare blockchain for iot," *Sensors*, vol. 19, no. 2, p. 326, 2019.

[16] C. Esposito, A. De Santis, G. Tortora, H. Chang, and K.-K. R. Choo, "Blockchain: A panacea for healthcare cloud-based data security and privacy?" *IEEE Cloud Computing*, vol. 5, no. 1, pp. 31–37, 2018.

[17] Y. J. Galteland and S. Wu, "Blockchain-based privacy-preserving fair data trading protocol," *Cryptology ePrint Archive*, 2021.

[18] C. Ge, Z. Liu, J. Xia, and L. Fang, "Revocable identity-based broadcast proxy re-encryption for data sharing in clouds," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 3, pp. 1214–1226, 2019.

[19] N. Hamian, M. Bayat, M. R. Alaghband, Z. Hatefi, and S. M. Pournaghi, "Blockchain-based user re-enrollment for biometric authentication systems," *IJ of Electronics and Information Engineering*, vol. 14, no. 1, pp. 18–38, 2022.

[20] T. Hewa, M. Ylianttila, and M. Liyanage, "Survey on blockchain based smart contracts: Applications, opportunities and challenges," *Journal of Network and Computer Applications*, vol. 177, p. 102857, 2021.

[21] B. Huang, Z. Liu, J. Chen, A. Liu, Q. Liu, and Q. He, "Behavior pattern clustering in blockchain networks," *Multimedia Tools and Applications*, vol. 76, pp. 20 099–20 110, 2017.

[22] H. Huang, P. Zhu, F. Xiao, X. Sun, and Q. Huang, "A blockchain-based scheme for privacy-preserving and secure sharing of medical data," *Computers & Security*, vol. 99, p. 102010, 2020.

[23] L. C. Huang, L. Y. Tseng, M. S. Hwang, "The study on data hiding in medical images", *International Journal of Network Security*, vol. 14, no. 6, pp. 301–309, 2012.

[24] M. S. Hwang, C. C. Lee, and S.-F. Tzeng, "A new proxy signature scheme for a specified group of verifiers," *Information Sciences*, vol. 227, pp. 102–115, 2013.

[25] M. S. Hwang, S. F. Tzeng, C. S. Tsai, "Generalization of proxy signature based on elliptic curves," *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 73-84, 2004.

[26] S. Jiang, X. Zhu, R. Hao, H. Chi, H. Li, and L. Wang, "Lightweight and privacy-preserving agent data transmission for mobile healthcare," in *IEEE International Conference on Communications*, pp. 7322–7327, 2015.

[27] C. T. Li, M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks",*Information Sciences*, vol. 181, no. 23, pp. 5333-5347, 2011.

[28] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245–255, 2003.

[29] X. Liang, M. Barua, L. Chen, R. Lu, X. Shen, X. Li, and H. Y. Luo, "Enabling pervasive healthcare through continuous remote health monitoring," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 10–18, 2012.

[30] C. Y. Lin, L. C. Huang, Y. H. Chen, M. S. Hwang, "Research on security and performance of blockchain with innovation architecture technology", *International Journal of Network Security*, vol. 23, no. 1, pp. 1-8, 2021.

[31] C. W. Liu, W. F. Hsien, C. C. Yang, M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900-916, 2016.

[32] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing", *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[33] E. J. L. Lu, M. S. Hwang, and C. J. Huang, "A new proxy signature scheme with revocation", *Applied Mathematics and Computation*, vol. 161, no. 3, PP. 799-806, 2005.

[34] C. McFarlane, M. Beer, J. Brown, and N. Prendergast, "Patientory: a healthcare peer-to-peer emr storage network v1," *Entrust Inc.: Addison, TX, USA*, 2017.

[35] Y. Miao, Q. Tong, R. H. Deng, K.-K. R. Choo, X. Liu, and H. Li, "Verifiable searchable encryption framework against insider keyword-guessing attack in cloud storage," *IEEE Transactions on Cloud Computing*, vol. 10, no. 2, pp. 835–848, 2020.

[36] G. N. Nguyen, N. H. Le Viet, M. Elhoseny, K. Shankar, B. Gupta, and A. A. Abd El-Latif, "Secure blockchain enabled cyber–physical systems in healthcare using deep belief network with resnet model," *Journal of parallel and distributed computing*, vol. 153, pp. 150–160, 2021.

[37] A. Panarello, N. Tapas, G. Merlino, F. Longo, and A. Puliafito, "Blockchain and iot integration: A systematic survey," *Sensors*, vol. 18, no. 8, p. 2575, 2018.

[38] Z. P. Peng, M. L. Chiang, I. C. Lin, C. C. Yang, M. S. Hwang, "CBP2P: Cooperative electronic bank payment systems based on blockchain technology", *Journal of Internet Technology*, vol. 23, no. 4, pp. 683-692, 2022.

[39] N. Rifi, E. Rachkidi, N. Agoulmine, and N. C. Taher, "Towards using blockchain technology for iot data access protection," in *IEEE 17th International Conference on Ubiquitous Wireless Broadband*, pp. 1–5, 2017.

[40] H. Shafagh, L. Burkhalter, A. Hithnawi, and S. Duquennoy, "Towards blockchain-based auditable storage and sharing of iot data," in *Proceedings of the 2017 on Cloud Computing Security Workshop*, pp. 45–50, 2017.

[41] J. Sun, X. Yao, S. Wang, and Y. Wu, "Blockchain-based secure storage and access scheme for electronic medical records in ipfs," *IEEE Access*, vol. 8, pp. 59 389–59 401, 2020.

[42] S. F. Tzeng, M. S. Hwang, C. Y. Yang, "An improvement of nonrepudiable threshold proxy signature scheme with known signers", *Computers & Security*, vol. 23, no. 2, pp. 174-178, 2004.

[43] W. Yang and Y. Zhu, "A verifiable semantic searching scheme by optimal matching over encrypted data in public cloud," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 100–115, 2020.

[44] S. Yao, R. Sankar, and I.-H. Ra, "A collusion-resistant identity-based proxy reencryption scheme with ciphertext evolution for secure cloud sharing," *Security and Communication Networks*, vol. 2020, pp. 1–16, 2020.

[45] M. Zhang, C. Chen, T. Wo, T. Xie, M. Z. A. Bhuiyan, and X. Lin, "Safedrive: Online driving anomaly detection from large-scale vehicle data," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2087–2096, 2017.

[46] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *IEEE International Congress on Big Data*, pp. 557–564, 2017.

Biography

Xin Qi is an undergraduate student, majoring in communication engineering, specializing in information security. His research interests include applied cryptography, privacy and security of medical data, etc. Email: 1789424660@qq.com

Shu Wu received the M.S. degree in pattern recognition and intelligent system from Nanjing University of Posts and Telecommunications, China, in 2011. He is currently a lecturer with West Anhui University, China. He has more than 7 years of work experience in the relevant industries, and has authored nearly 10 articles. He is currently pursuing the Ph.D. degree in Physical information and intelligence systems with the School of Physical and Electronic Information Engineering, Anhui Normal University, Wuhu, China. His research interests include healthcare blockchain, applied cryptography, and security of cyberspace. Email: wuyshu@126.com

Shu-hao Yu is a professor and a master's supervisor. He received his Bachelor's degree in Computer Education from Huaibei Normal University in 1999, his master's degree in computer software and theory from Hohai University in 2006, and his Doctor's degree in Management Science and Engineering from Hefei University of Technology in 2015. He has published more than 30 academic papers as the first author (corresponding author), including 6 SCI papers and 8 EI papers. He has obtained one national invention patent and 10 computer software Copyrights.His research interests include information security, data privacy protection, etc. Email: 1956345025@qq.com

Research on Image Copy-paste Tamper Detection Based on Gray Scale Co-occurrence Matrix and Graph Neural Network

Jiwei Sun

(Corresponding author: Jiwei Sun)

School of Smart Manufacturing, Zhengzhou University of Economics and Trade
No. 2 Shuanghu Avenue, Longhu Town, Xinzheng City, Zhengzhou 451191 China
Email: zxcvfdsa5024@foxmail.com

(Received July 15, 2023; Revised and Accepted Oct. 1, 2023; First Online Oct. 10, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

With the rapid development of image processing technology, the function of digital image editing software is more and more powerful; even non-professionals can easily tamper with the image content by using this software. The copy-and-paste of digital images is the most common and covert method of tampering. To solve the problem of uneven feature extraction and single-scale superpixel division, this paper proposes a grayscale co-occurrence matrix and graph neural network for image copy-paste tamper detection. Firstly, the image to be detected is divided into multiple overlapping blocks of the same size. The gray co-occurrence matrix statistics represent each image's texture features, and the feature vector of the image is obtained. Then, a new graph neural network feature matcher based on the attention mechanism is introduced to perform iterative matching between superpixels, and the false matching is eliminated by the random sample consensus (RANSAC) algorithm. Finally, the multi-scale matching results are fused to locate the tamper region accurately. Many experiments show that this new algorithm has higher detection quality and efficiency than the traditional algorithms.

Keywords: Graph Neural Network; Gray Scale Co-occurrence Matrix; Image Copy-paste Tamper Detection; Random Sample Consensus

1 Introduction

At present, digital images have been widely used in People's Daily life and work. At the same time, the rapid development of powerful image processing software makes it easier to tamper with digital images [8, 9, 12]. If false images are misused in the fields of law, medicine, military, etc., it will bring immeasurable negative impact on society and personal life. Therefore, the forensics research on digital image tampering is of great significance [11,13,21]. At present, digital image forensics mainly include active forensics and passive forensics. Active forensics requires digital signature or watermark technology to be added to the image in advance, so its application scope is limited [7, 16, 34, 35, 37]. Passive forensics does not need to add any information to the image in advance, and only needs to use the characteristics of the image itself to achieve authentication, so it has a wider application [4–6]. Passive forensics has become a new hotspot in the field of digital image security [26, 40].

There are many ways to tamper with digital images, among which regional copy-paste tampering is the simplest and most effective method. It copies a certain part of the image, and then pastes it to another area in the same image that does not intersect [38], so as to cover up or remove some information in the image. Since the copy area and paste area in the image are basically similar, the tampering traces can be detected by looking for similar areas in the image.

In recent years, many researchers have introduced deep learning methods into CMFD [20, 36]. Lin *et al.* [41] carried out adaptive segmentation of the image, detected feature points through segmentation based key point distribution strategy, and extracted image depth features by convolutional kernel network (CKN), and finally conducted K-nearest neighbor search to find matching pairs. Abhishek *et al.* [10] proposed the use of deep convolutional neural networks by training VGG-16 to classify real images and forged images, and then using semantic segmentation method to train images with color pixel labels to locate the forged region. However, the above methods had implemented a complete tamper detection step in the same neural network model. Therefore, Zhao *et*

al. [27] proposed a tamper detection step based on end-to-end DNN to realize a complete tamper detection step, which not only avoided setting various parameters and thresholds, but also jointly trained all modules on the reconstruction loss of forged masks. Relying on neural network model can accurately classify whether an image has been tampered with, but due to the complexity of the tampered region, it is often difficult for deep learn-based methods to obtain a high recall rate.

Using graph neural network (GNN) [17,28] to learn image feature matching has outstanding application effect in the fields of image matching, image segmentation, image retrieval, etc., but it has not been widely used in the field of image tampering detection. In this paper, a new GNN model based on attention mechanism is introduced as a feature matcher [14]. After super pixel partitioning and depth feature extraction, super pixel is used as the input of the feature matcher to carry out feature matching between super pixels. Since the effect of super pixel segmentation has a great influence on the positioning results, this paper adopts the multi-scale segmentation and fusion method to obtain accurate positioning results.

2 Proposed Method

In the field of image tampering detection, aiming at the problem that uneven feature extraction and single-scale super-pixel division have great influence on the tampering location results, this paper proposes an image CMFD algorithm based on gray co-occurrence matrix and GNN matching, and introduces a new GNN model based on attention mechanism for feature matching, so as to improve the efficiency and accuracy of feature matching. Figure 1 is the flow chart of the algorithm in this paper.

2.1 Feature Extraction Based on Gray Co-occurrence Matrix

A sliding window of size $b \times b$ is used to scan the detected image by sliding one pixel point at a time from the upper left corner of the image to the lower right corner of the image. Thus, an $M \times N$ image to be detected is divided into $(M-b+1) \times (N-b+1)$ image sub-blocks.

Gray co-occurrence matrix is a statistics-based texture analysis method, which describes texture by the spatial correlation of pixel gray [25, 42]. The gray co-occurrence matrix is defined as a pixel (x, y) with gray level i as the starting point, and the frequency $P(i, j, d, \theta)$ of the pixel $(x+\Delta x, y+\Delta y)$ with distance d, direction θ and gray level j is counted.

The expression of gray co-occurrence moment is as follows:

$$P(i, j, d, \theta) = (x+\Delta x, y+\Delta y) | f(x, y) = i \qquad (1)$$

Where (x, y) is pixel coordinate. i and j stand for gray level. Δx and Δy represent the offset of coordinates. d is the displacement. θ is the matrix generation direction. When the position relation of two pixels (d and θ) is determined, the gray co-occurrence matrix under the relation can be generated. For example, Figure 2(a) is a 4×5 image, and Figure 2(b) is its corresponding co-occurrence matrix with gray level $n = 8$, $d = 1$,$\theta = 0°$.

Since the gray co-occurrence matrix can not be used to describe the texture features directly, some statistics are defined to extract the texture features reflected by it. In this paper, inertia, energy, entropy, contrast, deficit moment, cluster shadow, cluster prominence and correlation are used. (i, j) represents the pixels of the image, $P(i, j)$ represents the gray co-occurrence matrix, and n represents the number of gray levels.

1) Inertance.

$$F1 = \sum_{i,j=0}^{n} (i-j)^2 P(i, j). \qquad (2)$$

2) Energy.

$$F2 = \sum_{i,j=0}^{n} P^2(i, j). \qquad (3)$$

3) Entropy.

$$F3 = -\sum_{i,j=0}^{n} P(i, j) ln P(i, j). \qquad (4)$$

4) Contrast.

$$F4 = -\sum_{i,j=0}^{n} P(i, j)^5 |i-j|^5. \qquad (5)$$

5) Deficit moment.

$$F5 = \sum_{i,j=0}^{n} \frac{P(i, j)}{1+(i-j)^2}. \qquad (6)$$

6) Cluster shadow.

$$F6 = -\sum_{i,j=0}^{n} (i-ux+j-uy)^3 P(i, j). \qquad (7)$$

7) Cluster prominence.

$$F7 = -\sum_{i,j=0}^{n} (i-ux+j-uy)^4 P(i, j). \qquad (8)$$

8) Correlation.

$$F8 = \frac{-\sum_{i,j=0}^{n} P(i, j) ln P(i, j) - hxy}{max(hx, hy)}. \qquad (9)$$

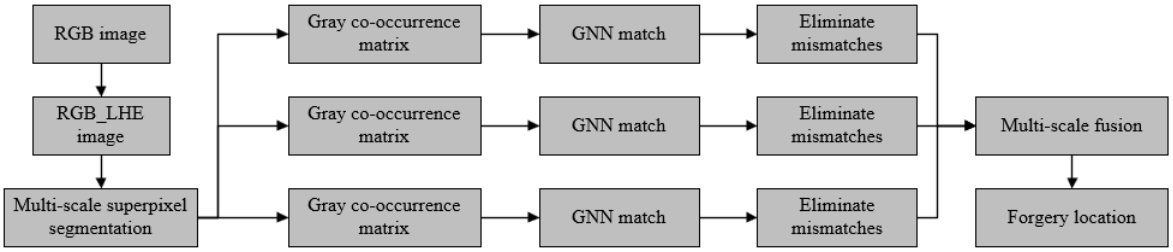


Figure 1: The proposed scheme

1	1	5	6	8
2	3	5	7	1
4	5	7	1	2
8	5	1	2	5

(a) 4×5 image

1	2	0	0	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
1	0	0	0	0	1	2	0
0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0

(b) Gray-level co-occurrence matrix

Figure 2: Gray-level co-occurrence matrix

Where,

$$ux = \sum_{i,j=0}^{n} iP(i,j), uy = \sum_{i,j=0}^{n} jP(i,j). \tag{10}$$

$$hx = -\sum_{i=0}^{n}\sum_{j=0}^{n} P(i,j) ln \sum_{j=0}^{n} P(i,j). \tag{11}$$

$$hy = -\sum_{j=0}^{n}\sum_{i=0}^{n} P(i,j) ln \sum_{i=0}^{n} P(i,j). \tag{12}$$

$$hxy = -\sum_{i,j=0}^{n} P(i,j) ln(\sum_{j=0}^{n} P(i,j) \sum_{i=0}^{n} P(i,j)). \tag{13}$$

Since the features of gray co-occurrence matrix are closely related to the direction, if the extracted texture features are to remain unchanged during image rotation, the results of gray co-occurrence matrix should be properly processed. The specific processing method is that: take the migration parameters in different directions, find their gray co-occurrence matrix, respectively find their 8 eigenvalues, and then calculate the mean value of these eigenvalues. In order to avoid the huge amount of computation caused by a large gray quantization level, the algorithm reduces the gray level by taking the gray level of 16, $d = 1$, and θ to take the gray co-occurrence matrix of 0°, 45°, 90° and 135° respectively. The 8 eigenvalues of the gray co-occurrence matrix of each image after segmentation are calculated, and their mean values are calculated and stored in a certain line of matrix A respectively. In this case, A is a matrix with 8 columns and $(M - b + 1) \times (N - b + 1)$ rows.

The matrix A after extracting the eigenvalues is lexicographically sorted, and the displacement vector of the coordinate values of the corresponding image blocks is calculated for the two adjacent rows in A, and stored in a counter C. If $(i1, i2)$ and $(j1, j2)$ represent the positions of two matching blocks, the displacement vector between the two matching blocks is calculated according to the following formula:

$$s = (s1, s2) = (i1 - j1, i2 - j2). \tag{14}$$

Since the displacement vector s and $-s$ correspond to the same displacement, taking the absolute value of s normalizes it. Incrementing the standard displacement vector counter C by 1 each time, a matching block pair is found:

$$C(s1, s2) = C(s1, s2) + 1. \tag{15}$$

For adjacent matching rows in the sorted matrix A, calculate their displacement vector and the value of counter C. At the beginning of the algorithm, the counter C is initialized to zero. For the displacement vector s, if it exceeds the specified threshold $T : C(s) > T$, then it is considered that the matching block pairs that generate the displacement vector s correspond to the copy region and the paste region, and the positions of these block pairs are identified. If the representation block is isolated, it may be misjudged [42]. The obtained preliminary detection image can be corroded by mathematical morphological filtering, and then expanded to eliminate those small areas that are misjudged as tampering areas.

3 Feature Matching Based on GNN Model

This paper introduces a new GNN feature matcher based on attention mechanism. The feature matcher mainly consists of two modules: attention GNN and optimal matching, and its structure is shown in Figure 3.

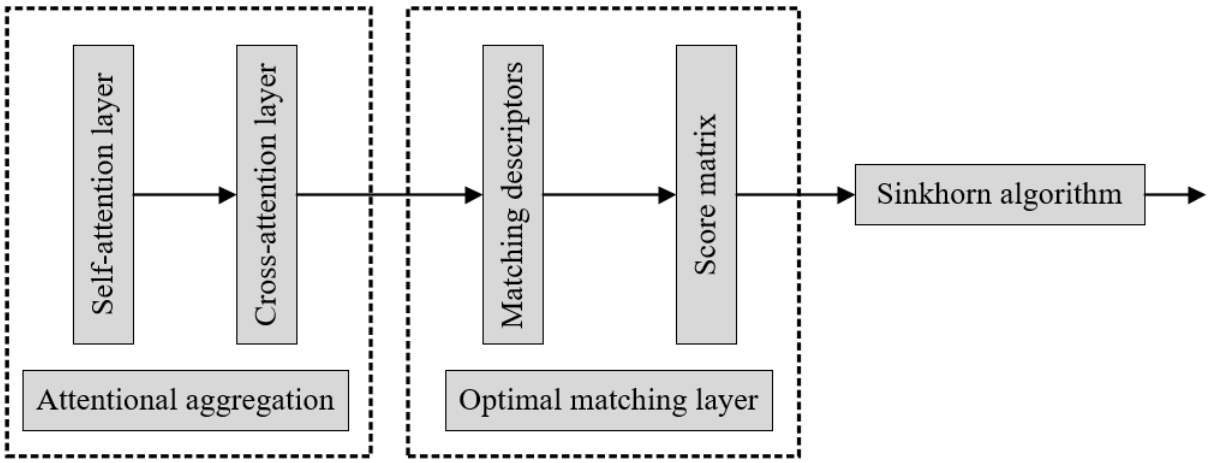


Figure 3: Structure diagram of GNN feature matcher

In the attention GNN module, this module uses nine alternating layers of inner attention layer and cross attention layer. The inner attention layer is used to transfer the feature information within the super-pixels, and the cross attention layer is used to transfer the feature information between the super-pixels. Starting from the high-dimensional state of each super-point feature point, all edge information of all nodes is aggregated, and a robust feature descriptor is generated through 9 layers of attentional GNN aggregation [19, 29, 30, 39]. The optimal matching module generates an allocation matrix to get matching pairs and filter out false matching pairs. First calculate the similarity score $S_{i,j}$ between the key points. Secondly, in order to make the network suppress some unmatched feature points, a Dustbin score is added to the score matrix and the unmatched feature points are assigned to it. Finally, the Sinkhorn algorithm is used to eliminate the Dustbin score, solve the optimal matching matrix $\bar{P}$, and minimize the loss function (16) to maximize the accuracy of the feature matcher.

$$Loss = -D - E - F. \tag{16}$$

Here, $D = \sum_{(i,j)\in\varpi} log\bar{P}_{i,j}$, $E = \sum_{i\in I} log\bar{P}_{i,N+1}$, $F = \sum_{j\in J} log\bar{P}_{M+1,j}$. In the formula, M and N represent the number of feature points in super-pixels. The location information and feature vector of matching pairs are indexed according to the optimal matching matrix to obtain the set of matching pairs. Under the single-scale super-pixel division, GNN features are used for matching.

4 Experimental Results and Analysis

In order to evaluate the ability of this algorithm to detect copy-paste tampering in image areas, images in Waterloo BragZone standard image library are selected in the experiment and compared with the classical GNN algorithm [23]. For RGB images, it is first converted into grayscale images and then detected. The experimental environment is Matlab7b, and the threshold of similarity is set to $\theta = 0.978$.

A 512×512 peppers gray scale image is randomly selected from the standard image library, and regional copy-paste tampering operation is performed on the image, and then the detection effect of the proposed algorithm and GNN algorithm on the tampered image is compared. The original image and the altered image are shown in Figure 4(a) and Figure 4(b) respectively. The detection results of the proposed algorithm and GNN algorithm on the tampered image are shown in Figure 4(c) and Figure 4(d) respectively. It can be seen from Figure 4(c) and Figure 4(d) that both the proposed algorithm and the GNN algorithm can accurately detect the copy-paste tampering area of the image, and the detection effect is basically the same.

In the above experiment, when pasting the copy area of the image, the copy area was rotated at different angles, and then the detection effect of the proposed algorithm and GNN algorithm on the tampered image was compared. The original image is rotated 90° and 180° as shown in Figure 5(a) and Figure 5(b), respectively. The detection effects of the proposed algorithm and GNN algorithm on Figure 5(a) and Figure 5(b) respectively are shown in figures 5(c-f). By comparing Figure 5(c) with Figure 5(d), Figure 5(e) and Figure 5(f), it can be seen that GNN algorithm does not have the ability to detect the image after rotation, while the algorithm in this paper can accurately detect the image region after rotation operation, and has strong robustness for image post-processing.

In order to further evaluate the image CMFD algorithm based on gray co-occurrence matrix feature extraction and GNN matching, this paper calculates three index values of Precision, Recall and F on GRIP data set for

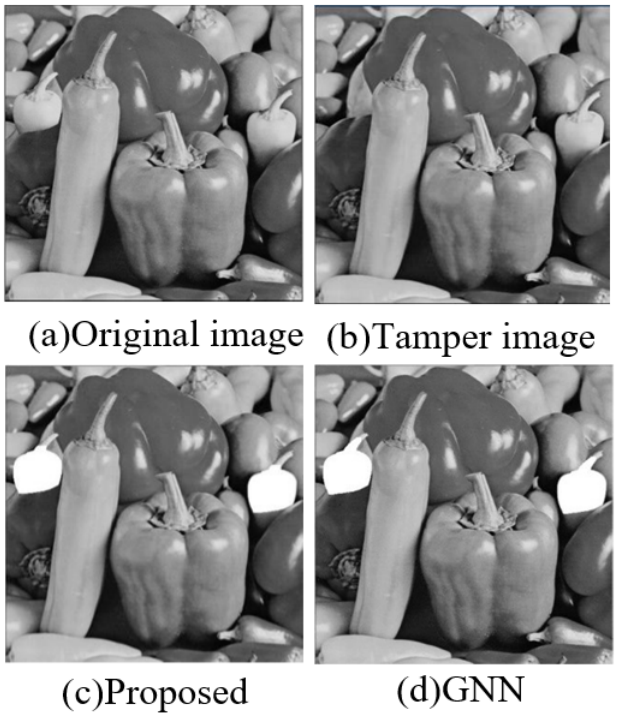


Figure 4: Area copy and paste test results

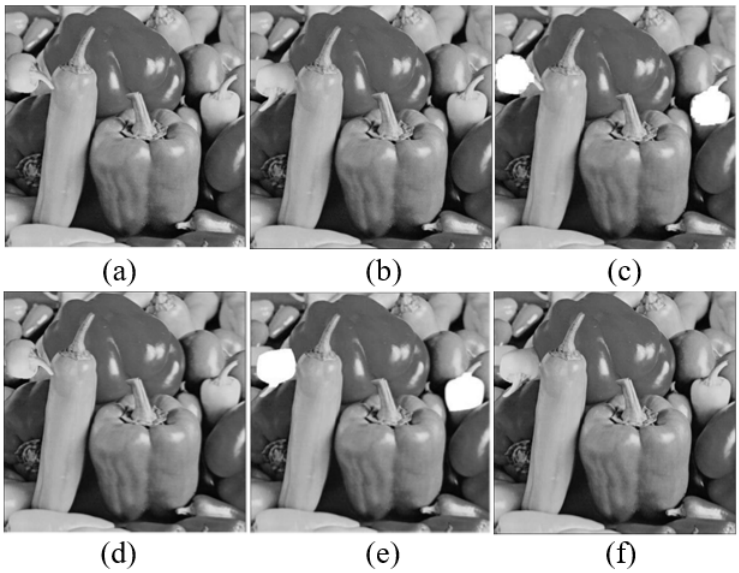


Figure 5: The detection result of the tampered area after rotation. (a) The replication area is rotated 90°; (b) The replication area is rotated 180°; (c) Proposed detection method with rotated 90°; (d) GNN detection method with rotated 90°; (e) Proposed detection method with rotated 180°; (f) GNN detection method with rotated 180°;

evaluation [2, 24, 32, 33]. Table 1 shows the comparison of experimental results of different algorithms on GRIP data set.

Table 1: Typical states of SEIR model

Method	P	R	F
ASP[25]	0.918	0.934	0.926
AKF[26]	0.813	0.888	0.848
GNN[27]	0.922	0.834	0.875
Proposed	0.935	0.930	0.932

In the table, not only the evaluation index values of the above comparison algorithm are calculated, but also the index values of SIFT feature and SURF feature combined with GNN feature matcher respectively are calculated. It can be concluded from the table that compared with GNN, the P, R and F values of superpoint combined with SIFT and GNN, SURF and GNN are the maximum values, that is, the best detection effect is achieved, because SIFT and SURF features have fewer feature points in the smooth region. Many undetected areas exist.

5 Conclusions

In this paper, the image is divided into super-pixels in the pre-processing step, and the super-point feature points of the image are extracted by using gray scale co-existence matrix. In order to ensure sufficient number of feature points, uniform feature points are obtained by adaptive adjustment of threshold values within the range of super-pixels. In the subsequent matching process, GNN feature matcher based on attention mechanism is introduced, and hierarchical clustering and RANSAC algorithm are used to eliminate the influence of wrong matching, and the matched super-pixel is obtained. In order to locate the tamper region accurately, a method of fusion of multi-scale matching results is proposed to obtain accurate detection results. The P, R and F of the proposed algorithm in GRIP data set reach 93.5%, 93.0% and 93.2% respectively, which can effectively locate the forged area, which fully proves that the proposed algorithm has strong robustness.

Acknowledgments

The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] Abhishek and N. Jindal, "Copy move and splicing forgery detection using deep convolution neural network, and semantic segmentation," *Multimedia Tools and Applications*, vol. 80, pp. 3571-3599, 2021.

[2] A. Bera, Z. Wharton, Y. Liu, N. Bessis and A. Behera, "SR-GNN: Spatial Relation-Aware Graph Neural Network for Fine-Grained Image Categorization," *IEEE Transactions on Image Processing*, vol. 31, pp. 6017-6031, 2022.

[3] S. Bhalerao, I. A. Ansari, A. Kumar, "A secure image watermarking for tamper detection and localization," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1057-1068, 2021.

[4] C. C. Chang, K. F. Hwang, M. S. Hwang, "A digital watermarking scheme using human visual effects", *Informatics*, vol. 24, no. 4, pp. 505–511, Dec. 2000.

[5] C. C. Chang, K. F. Hwang, M. S. Hwang, "A block based digital watermarks for copy protection of images," in *Asia-Pacific Conference on Communications*, vol. 2, pp. 977-980, 1999.

[6] C. C. Chang, K. F. Hwang, M. S. Hwang, "Robust authentication scheme for protecting copyrights of images and graphics", *IEE Proceedings-Vision, Image and Signal Processing*, vol. 149, no. 1, pp. 43-50, 2002.

[7] C. C. Chang, K. F. Hwang, M. S. Hwang, "A feature-oriented copyright owner proving technique for still images," *International Journal of Software Engineering and Knowledge Engineering*, vol. 12, no. 3, pp. 317-330, 2002.

[8] C. H. Chen, Y. L. Tang, W. S. Hsieh, M. S. Hwang, "Image tamper detection and recovery by intersecting signatures", *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 12, no. 4, pp. 1123-1131, 2014.

[9] T. Y. Chen, M. S. Hwang, J. K. Jan, "A secure image authentication scheme for tamper detection and recovery", *The Imaging Science Journal*, vol. 60, no. 4, pp. 219-233, 2012.

[10] Y. Fan, J. Li, Y. Guo, L. Xie, G. Zhang, "Digital image colorimetry on smartphone for chemical analysis: A review," *Measurement*, vol. 171, pp. 108829, 2021.

[11] Y. Hao, L. Zhang, S. Qiao, Y. Bai, R. Cheng, H. Xue, Y. Hou, W. Zhang, G. Zhang, "Breast cancer histopathological images classification based on deep semantic features and gray level co-occurrence matrix," *Plos one*, vol. 17, no. 5, pp. e0267955, 2022.

[12] L. C. Huang, L. Y. Tseng, M. S. Hwang, "The study on data hiding in medical images", *International Journal of Network Security*, vol. 14, no. 6, pp. 301–309, 2012.

[13] M. Hussan, S. Gull, S. A. Parah, G. J. Qureshi, "An efficient encoding based watermarking technique for tamper detection and localization," *Multimedia Tools and Applications*, pp. 1-23, 2023.

[14] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548–555, Jan. 2000.

[15] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks",

Journal of Electronic Imaging, vol. 9, no. 4, pp. 301–555, Jan. 2000.

[16] M. S. Hwang, K. F. Hwang, C. C. Chang, "A time-stamping protocol for digital watermarking", *Applied Mathematics and Computation*, vol. 169, pp. 1276–1284, 2005.

[17] P. Li, A. A. Laghari, M. Rashid, J. Gao, T. R. Gadekallu, A. R. Javed, S. Yin, "A Deep Multimodal Adversarial Cycle-Consistent Network for Smart Enterprise System," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 693-702, 2023.

[18] X. Liang, S. He, "Deep Learning Based Image Forgery Detection Methods," *Journal of Cybersecurity*, vol. 4, no. 2, pp. 119, 2022.

[19] C. C. Lin, T. L. Lee, Y. F. Chang, P. Shiu, B. Zhang, "Fragile Watermarking for Tamper Localization and Self-Recovery Based on AMBTC and VQ," *Electronics*, vol. 12, no. 2, pp. 415, 2023.

[20] I. C. Lin, H. H. Ou, M. S. Hwang, "A user authentication system using back-propagation network", *Neural Computing & Applications*, vol. 14, pp. 243-249, 2005.

[21] J. Liu, J. Feng, "Design of embedded digital image processing system based on ZYNQ," *Microprocessors and Microsystems*, vol. 83, pp. 104005, 2021.

[22] C. Min, J. Xu, L. Xiao, D. Zhao, Y. Nie and B. Dai, "Attentional Graph Neural Network for Parking-Slot Detection," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3445-3450, 2021.

[23] A. O. Mulani, G. N. Shinde, "An approach for robust digital image watermarking using DWT-PCA," *Journal of Science and Technology*, vol. 6, no. 1, 2021.

[24] N. R. NR, R. Shreelekshmi, "Fragile watermarking scheme for tamper localization in images using logistic map and singular value decomposition," *Journal of Visual Communication and Image Representation*, vol. 85, pp. 103500, 2022.

[25] E. E. Osagiede, M. Rosenau, A. Rotevatn, R. Gawthorpe, A. Jackson, M. Rudolf, "Influence of zones of pre-existing crustal weakness on strain localization and partitioning during rifting: Insights from analog modeling using high-resolution 3D digital image correlation," *Tectonics*, vol. 40, no. 10, pp. e2021TC006970, 2021.

[26] I. V. Pantic, A. Shakeel, G. A. Petroianu, P. R. Corridon, "Analysis of vascular architecture and parenchymal damage generated by reduced blood perfusion in decellularized porcine kidneys using a gray level co-occurrence matrix," *Frontiers in Cardiovascular Medicine*, vol. 9, pp. 797283, 2022.

[27] C. Qu, C. Liu, Y. Liu, X. Chen, D. Peng, F. Guo, L. Jin, "Towards Robust Tampered Text Detection in Document Image: New Dataset and New Solution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5937-5946, 2023.

[28] D. A. K. Sahu, "A logistic map based blind and fragile watermarking for tamper detection and localization in images," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 8, pp. 3869-3881, 2022.

[29] E. A. K. Sahu, M. Hassaballah, R. S. Rao, G. Suresh, "Logistic-map based fragile image watermarking scheme for tamper detection and localization," *Multimedia Tools and Applications*, pp. 1-32, 2022.

[30] F. S. Schaefer, D. Gehrig, D. Scaramuzza, "Aegnn: Asynchronous event-based graph neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12371-12381, 2022.

[31] X. Shi, P. Li, H. Wu, Q. Chen, H. Zhu, "A lightweight image splicing tampering localization method based on MobileNetV2 and SRM," *IET Image Processing*, 2023.

[32] L. Teng, Y. Qiao, M. Shafiq, G. Srivastava, A. R. Javed, T. R. Gadekallu, "FLPK-BiSeNet: Federated Learning Based on Priori Knowledge and Bilateral Segmentation Network for Image Edge Extraction," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 1529-1542, 2023.

[33] M. Z. Wang, Z. Wang, X. Li, Z. Yu, B. Guo, L. Chen, X. Zhou, "Exploring Multi-Dimension User-Item Interactions With Attentional Knowledge Graph Neural Networks for Recommendation," *IEEE Transactions on Big Data*, vol. 9, no. 1, pp. 212-226, 2023.

[34] C. C. Wu, S. J. Kao, W. C. Kuo, M. S. Hwang, "A robust-fragile watermarking scheme for image authentication," in *3rd International Conference on Innovative Computing Information and Control*, pp. 176, 2008.

[35] C. C. Wu, S. J. Kao, W. C. Kuo, M. S. Hwang, "A digital watermarking scheme using human visual effects," *Informatics*, vol. 24, no. 4, 2000.

[36] H. J. Wu, Y. H. Chang, M. S. Hwang, I. C. Lin, "Flexible RFID location system based on artificial neural networks for medical care facilities", *ACM SIGBED Review*, vol. 6, no. 2, pp. 1-8, 2009.

[37] N. I. Wu, C. M. Wang, C. S. Tsai, M. S. Hwang, "A certificate-based wtermarking scheme for coloured images", The Image Science Journal, vol. 56, no. 6, pp. 326-332, 2008.

[38] Y. Wu, D. Wang, L. Wang, Z. Shang, C. Zhu, J. Wei, A. Yuan, H. Zhang, F. Zeng, "An analysis of the meso-structural damage evolution of coal using X-ray CT and a gray-scale level co-occurrence matrix method," *International Journal of Rock Mechanics and Mining Sciences*, vol. 152, pp. 105062, 2022.

[39] S. Yin, Y. Zhang, S. Karim, "Region search based on hybrid convolutional neural network in optical remote sensing images," *International Journal of Distributed Sensor Networks*, vol. 15, no. 5, 2019.

[40] D. Zeng, W. Liu, W. Chen, L. Zhou, M. Zhang, H. Qu, "Substructure aware graph neural networks," in

Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, no. 9, pp. 11129-11137, 2023.

[41] G. Zhao, C. Qin, H. Yao, Y. Han, "DNN self-embedding watermarking: Towards tampering detection and parameter recovery for deep neural network," *Pattern Recognition Letters*, vol. 164, pp. 16-22, 2022.

[42] J. Zhu, R. Hong, H. Zhang, R. Gu, H. Wang, F. Sun, "Fired bullet signature correlation using the finite ridgelet transform (FRIT) and the gray level co-occurrence matrix (GLCM) methods," *Forensic science international*, vol. 330, pp. 111089, 2022.

Biography

Jiwei Sun biography. Sun Jiwei, female, Member of the Communist Party of China, Master of Engineering, associate professor, the current director of automation Teaching and Research Department. She is mainly engaged in the scientific research of control theory and control science and the teaching of "Motion control System" and "Electrical Control and PLC". Around the above directions and courses, In recent years, She has presided over and participated in more than 10 provincial, ministerial and university level scientific research, teaching and research projects, such as the key technology research project of precise control of inverters in distributed power generation system, the research project of sudden failure and variable parameter control technology of rotating machinery, the research project of LCL grid-connected inverter control technology, and the project of government inspection system of distributed food business transaction security based on SOA architecture. She has won the second prize of Henan Province Information Technology Education Outstanding Achievement Award, outstanding Communist Party Member, advanced individual of Sanyu, excellent instructor of discipline competition and other honors.

Secure Encryption of Parallel Chaotic English Education Data Based on 4-order Cellular Neural Network

Mengya Wei

(Corresponding author: Mengya Wei)

School of Foreign Languages, Zhengzhou University of Science and Technology

Zhengzhou 450064, China

Email: ljnan127@163.com

(Received July 15, 2023; Revised and Accepted Oct. 1, 2023; First Online Oct. 10, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

In English education, the security of English data is very important. It is related to the leakage of English data in the transmission process. In order to solve the problem of uncontrollable processing behavior of parallel chaotic English data encryption and realize secure transmission of English data of network encrypted information, a parallel chaotic English education data based on a 4-order cellular neural network is designed. Firstly, the 4-order cellular neural network is synchronized by the active-passive synchronization method, generating chaotic signals. Then, the synchronous random sequence generated by the synchronous chaotic signal is used for data scrambling and diffusion. Finally, the Advanced Encryption Standard (AES) encryption algorithm is combined with the spiral matrix-based scrambling diffusion model, and the elliptic curve is used to encrypt the AES key. The key and encrypted data are transmitted through the public channel. The experimental results show that compared with the application system based on parameter optimization, the memory ratio and mixed ratio values of host components are closer to the ideal value level under the action of the 4-order cellular neural network, which not only solves the problem of uncontrollable data encryption behavior but also realizes the secure transmission and processing of network encrypted information.

Keywords: 4-order Cellular Neural Network; AES; Parallel Chaotic English Education Data; Scrambling Diffusion

1 Introduction

In the modern era of information diversification, data has become a part of People's daily life, and people use it all the time. Database management technology is widely used in data storage, reading and updating, it can provide corresponding services according to the needs of users, so as to achieve the purpose of easy data management for users [16]. However, when performing specific tasks, databases are vulnerable to a wide range of threats, such as excessive privilege abuse, legal privilege abuse, weak authentication, and backup data exposure [12, 18–20]. Therefore, it is essential to ensure the security of the database [1, 44, 47].

Domestic and foreign scholars have carried out a lot of research on how to ensure the security of database. In 2016, Malik *et al.* [27] proposed a layered reference model of database security agent to improve the ability of key applications to resist malicious attacks on database management System. It enhanced the ability of firewall reference model to defend against enemies by adding transmission unit modification function and attribute value mapping function. With the development of Data information security technology, Transparent Data Encryption (TDE) technology [26] has entered the field of vision of researchers, and since then, due to its excellent characteristics of encryption and decryption processes are completely transparent to applications and users, it has been widely used.

Today, TDE is the best choice for bulk database encryption. Based on the principle of transparent data encryption and the deployment of transparent database encryption process, Almuzaini *et al.* [2] implemented and tested the whole process of transparent database encryption in the SQL Server 2008 environment. Shmueli *et al.* [33] studied five database encryption architectures, including Always Encrypted (AE) and Transparent Data encryption (TDE), Cell Level Encryption (CLE), Dynamic Data Masking (DDM) and Vormetric Transparent Encryption (VTE). The study showed that choosing the right encryption strategy was the most important to keep-

ing your organization safe and protecting data and information. Bhattacharjya *et al.* [3] studied how the massive transparent data encryption of data security solutions on Microsoft SQL Server affected the performance of database management systems. The authors conducted stress and load tests on the performance of each system, and the results showed that using transparent data encryption on standard databases has many advantages.

Cellular Neural Network (CNN) [8, 29] is a feedback neural network proposed on the basis of artificial neural network. It is a nonlinear system formed by a finite number of cells arranged and connected according to certain rules [35]. When the connection and arrangement rules of cells meet certain conditions, CNN can produce high-dimensional hyperchaos [11,46], so it is often used to generate high-dimensional chaotic signals and applied to various encryption systems [22, 28]. At the same time, in order to more effectively resist various attacks, some digital encryption algorithms, such as Elliptic Curve Cryptography (ECC) [15, 17, 36, 38, 41], RSA encryption algorithm [6, 32, 37], etc., are also applied to encryption systems. In reference [31], the plaintext image was converted into two-dimensional code, and then ECC was used for encryption, and an anti-noise image encryption scheme was proposed based on the anti-noise property of two-dimensional code [7, 39, 40]. In reference [42], RSA algorithm was improved to control the initial values of 4-wing and Chen 4D hyperchaotic systems, and an image encryption system with larger key space and shorter encryption time was proposed [13]. Reference [43] firstly compressed the plaintext image, and then used 4D cat mapping and EC-ElGamal algorithm to globally scramble the compressed image [10, 21, 23]. RSA and ECC are the two most mainstream asymmetric cryptography technologies, which have the characteristics of simple key distribution and high key security [4, 14]. However, because of the slow encryption speed of asymmetric key system, it is difficult to be used for big data encryption. AES is a symmetric cryptographic technology that divides plaintext into independent plaintext blocks for block encryption. It has the advantages of flexible key length and fast encryption speed [34], but the key distribution of AES is more complicated. Aiming at the defects of existing image encryption schemes and the advantages of symmetric and asymmetric encryption schemes, an image encryption system based on 4-order cellular neural network and AES encryption is proposed in this paper. The key is distributed through the public channel, and the plaintext image is scrambled and diffused by synchronous random sequence and spiral matrix, which can effectively break the correlation between the plaintext image pixels The ability of encryption algorithm to resist differential attack, noise attack and clipping attack is improved.

In recent years, with the complexity of the Internet construction environment, parallel chaotic data, which combines parallel characteristics with chaotic characteristics, has become the core object of information processing. However, the encryption behavior of this type of data is always affected by various uncontrollable factors, and ultimately the security transmission ability of encrypted information is affected. Although the data encryption system based on parameter optimization can define the unit transmission capability of parallel chaotic data, it cannot encode and decode the transmitted information according to specific read and write rules, which is the main reason why the secure transmission capability of network encrypted information cannot reach the ideal standard level all the time. In order to solve the above problems, a novel parallel chaotic data encryption system is designed under the function of 4-order cellular neural network.

2 Proposed Data Encryption Scheme

The system framework of an image encryption algorithm based on 4-order cellular neural network and AES proposed in this paper is shown in Figure 1. The 4-order CNN hyperchaotic system at the sending and receiving ends realizes chaos synchronization through the active-passive synchronization method [30]. The plaintext image Hash function is used to generate AES encryption key and control the generation of random sequence. After the AES encryption key is encrypted through ECC, it is sent to the receiving end by the public channel, and the generated random sequence is used for image scrambling and diffusion at the sending end respectively.

2.1 Chaotic System

The model of the fourth-order cellular neural network at the sending end is represented as follows:

$$\begin{cases} \dot{x}_1 = -x_3 - x_4 \\ \dot{x}_2 = 2x_2 + x_3 \\ \dot{x}_3 = 14x_1 - 14x_2 \\ \dot{x}_4 = 100x_3 - 100x_4 + 100(|x_4 + 1| - |x_4 - 1|) \end{cases} \tag{1}$$

Where, x_i represents the state variable of the $i-th$ cell at the sending end.

In order to realize chaos synchronization between the sender and the receiver, active and passive synchronization methods are used in this paper, and $s(t) = x_3 + 4x_2$ is used as the driver signal of synchronization. Therefore, Formula (1) can be expressed as:

$$\begin{cases} \dot{x}_1 = -x_3 - x_4 \\ \dot{x}_2 = s(t) - 2x_2 \\ \dot{x}_3 = 14x_1 - 14x_2 \\ \dot{x}_4 = 100x_3 - 100x_4 + 100(|x_4 + 1| - |x_4 - 1|) \end{cases} \tag{2}$$

At the same time, the model of the receiving end is represented as:

$$\begin{cases} \dot{x}'_1 = -x'_3 - x'_4 \\ \dot{x}'_2 = s(t) - 2x'_2 \\ \dot{x}'_3 = 14x'_1 - 14x'_2 \\ \dot{x}'_4 = 100x'_3 - 100x'_4 + 100(|x'_4 + 1| - |x'_4 - 1|) \end{cases} \tag{3}$$

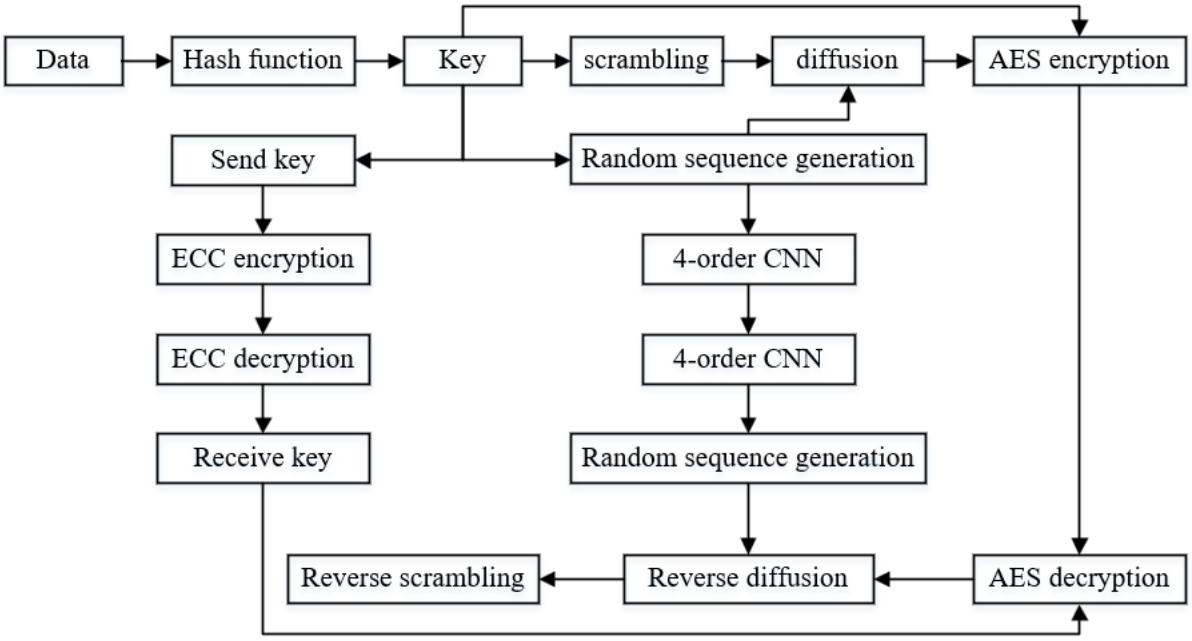


Figure 1: The proposed scheme

Its initial value is [-2,-2,1,1]. In formula (3), x'_i represents the state variable of the $i-th$ cell at the receiving end.

2.2 Key Generation

SHA-512, as a summarization algorithm, has one-way irreversibility, uniqueness and unpredictability, and can generate arbitrary data into a 128-bit hexadecimal string summarization with low algorithmic complexity [9]. This paper uses SHA-512 function to generate 128-bit hexadecimal string from plaintext image as AES key, and is used to generate random sequence, so that different plaintext image has different key and random sequence, which greatly strengthens the security of image.

2.3 Random Sequence Generation

Suppose the size of the encrypted image is $M \times N$, and the chaotic signal generated by CNN and sampled is $C_j(i)$, $i = 1, 2, \cdots, M \times N$, $j = 1, 2, \cdots, 4$. The random sequence is generated as follows:

1) Calculate the SHA-512 value of the image and get the vector H represented by the hexadecimal number;

2) Convert each digit of H to a decimal number, and sum these decimal numbers to get K;

3) Generate random sequences R_1, R_2.

$$R_1(i) = \lfloor K \times (C_1(i) + C_2(i)) \times 10^5 \rfloor \bmod (M \times N) \quad (4)$$

$$R_2(i) = \lfloor K \times (C_3(i) + C_4(i)) \times 10^5 \rfloor \bmod 256. \quad (5)$$

Where, mod(·) represents modulo operation, and $\lfloor$ and $\rfloor$ represents two different random sequences R_1 and R_2 generated by round down operation, which are respectively used for image scrambling and diffusion operations at the sending end, and the same random sequence is used for image inverse diffusion and inverse scrambling operations at the receiving end.

2.4 Scrambling and Diffusion

The scrambling and diffusion operations can greatly reduce the correlation between adjacent pixels. At the same time, in order to compensate for the small key space of AES encryption algorithm, additional nonlinear operations are provided before AES encryption [25]. Therefore, an index-based spiral matrix scrambling method is proposed, the steps are as follows:

1) Obtain the length M and width N of the plaintext image;

2) The spiral matrix S is generated according to M and N;

3) Rearrange the spiral matrix S into one-dimensional vector form s and the plaintext image into one-dimensional vector form p;

4) Arrange the random sequence R_1 from the largest to the smallest, and obtain the index sequence I after the arrangement;

5) The index sequence U is obtained by rearranging the index vectors using the spiral matrix;

6) The scrambling sequence Q is obtained by indexing the scrambling vector;

7) Rearrange the scrambled matrix G according to M and N;

8) The random sequence R_2 is rearranged into a diffusion matrix D according to the spiral matrix S;

9) The scrambling matrix G and diffusion matrix D are specified or operated to obtain the ciphertext matrix E;

2.5 Data Parallel Authentication

The data parallel authentication process must consider three physical indexes: two-factor factor factor, secure storage space of encrypted information [24] and encrypted file marking factor. Two-factor factor factor is also called PIN source protection factor, for parallel chaotic data, because the neural network environment does not limit the information transmission behavior, so the data source upper and lower limit positioning coefficient have unlimited expansion ability, which is easy to affect the encryption execution ability of the system host, but the performance ability of this influence is not unique. The secure storage space of encrypted information refers to the maximum storage capacity of the secure encryption system for parallel chaotic data. In a neural network environment, the larger the remaining space of the database host, the stronger the storage capacity of related data information. The markup coefficient of encrypted files can restrict the practical ability of the data read and write interface module, and generally satisfies the change rule that the larger the coefficient value, the stronger the module structure's ability to process data information [5, 45].

Let ξ indicate the two-factor coefficient and μ indicate the encryption file marking coefficient. $F(d)$ represents an indicator function based on the secure storage space of encrypted information. $F'(d)$ represents the inverse function of $F(d)$. d represents a given parallel chaotic data index. The data parallel authentication condition can be defined as:

$$K = \frac{Q/\xi[F(d) - F'(d)]^2}{\mu\sqrt{x_1^2 + \cdots + x_n^2}}. \tag{6}$$

Where $x_1^2 + \cdots + x_n^2$ represents n different data encryption calibration values, and n represents the maximum query result of the calibration instruction. In order to ensure the application security of encryption system, all parallel identities of chaotic data to be stored must be authenticated by neural network host.

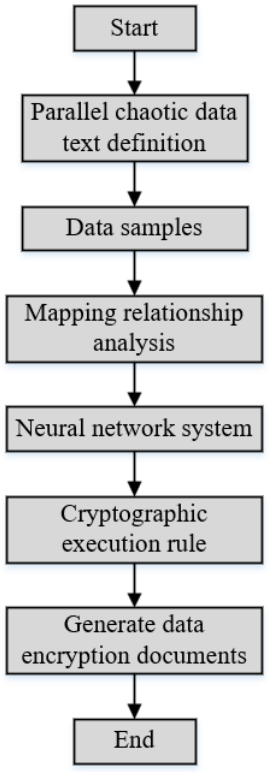


Figure 2: Encryption algorithm execution flowchart

2.6 Encryption Execution Flow

For parallel chaotic data, the encryption execution process should start with the data text definition. In the neural network environment, the storage form of these data samples is not unique, and its specific order of magnitude level should match the real-time input volume of parallel chaotic data samples. In order to satisfy the chaotic mapping relationship of data information, the neural network host should follow the principle of parallel identity authentication when processing data samples. Generally speaking, only the data samples that fully meet the authentication standards can have the ability of secondary transmission. The specific execution process is shown in Figure 2.

In the cognitive range of encryption execution law, whether the information parameter conforms to the data parallel authentication standard is the only criterion to judge whether the information parameter belongs to the parallel chaotic data. At this point, the implementation of the implementation of the hardware and software environment, the combination of the two to complete the parallel chaotic data security encryption system based on neural network design.

3 Instance Analysis

The control ability of network host for parallel chaotic data encryption processing can be considered from two

aspects: memory ratio and mixed ratio. Under normal circumstances, with the increase of parallel chaotic data transmission volume, the two indicators of memory ratio and mixed ratio will show a trend of increasing change, but when the ratio value exceeds the rated limit value (ideal maximum value), it can be judged that the network host is uncontrollable for the current parallel chaotic data encryption processing behavior. Therefore, it can be considered that when the memory ratio and mixing ratio curves are close to the ideal numerical curve, but do not exceed the rated maximum value, the network host has the strongest control ability for parallel chaotic data encryption processing.

Two Internet hosts with identical configurations were selected as experimental objects, in which the experimental host was equipped with a parallel chaotic data security encryption system based on neural network, and the control host was equipped with a data encryption system based on parameter optimization. The specific experimental execution process is as follows:

Step 1: Configure the Internet host of the experimental group and the control group at the same time, and input the neural network control program and parameter optimization control program into the established host components;

Step 2: Control the actual input of parallel chaotic data, and record the value change of correlation coefficient index after removing unreasonable information parameters;

Step 3: Record the value changes of memory ratio and mixed ratio, and compare the actual recorded results with the ideal values.

Table 1 records the change of ideal values of memory ratio index and mixed ratio index in experimental group and control group. Here, PCD: Parallel chaos data input, MER: memory ratio, MIR: mix ratio.

Table 1: Ratio indicator ideal value

PCD/Mb	MER/%	MIR/%
0	12.6	30.1
10	15.8	37.7
20	24.5	45.0
30	26.4	60.1
40	32.0	65.2
50	36.3	68.8
60	37.5	72.9
70	42.6	77.4

According to the analysis of Table 1, as the input volume of parallel chaotic data increases, both the memory ratio index and the mixed ratio index show a continuous increasing trend, but the unit increase value of the mixed ratio index is significantly greater than that of the memory ratio index.

Figure 3 reflects the actual changes of the memory ratio index in the experimental group and the control group.

The analysis of Figure 3 shows that the initial value of the memory ratio index of the experimental group is completely consistent with the initial value of the ideal curve. If this node is not considered, when the input volume of parallel chaotic data is equal to 60Mb, the difference between the memory ratio index of the experimental group and the ideal index is the smallest, only 1.2%. The initial value of the memory ratio index of the control group is slightly larger than the initial value of the ideal curve. If this node is not considered, when the input amount of parallel chaotic data is equal to 10Mb, the difference between the memory ratio index of the control group and the ideal index is at least 4.3%, which is higher than the physical value difference of the experimental group.

Figure 4 reflects the actual changes of the mixed ratio index of the experimental group and the control group.

According to figure 4, the initial value of the mixed proportion index of the experimental group is also completely consistent with the initial value of the ideal curve. If this node is not considered, when the input volume of parallel chaotic data is equal to 60Mb, the difference between the mixed proportion index of the experimental group and the ideal index is the smallest, only 2.8%. The initial value of the mixed proportion index of the control group was significantly higher than the initial value of the ideal curve. During the whole experiment, when the input of parallel chaotic data was equal to 30Mb, the difference between the mixed proportion index of the control group and the ideal index was at least 5.5%, which was higher than the difference of the physical value of the experimental group.

In summary, under the effect of neural network system, both the network host memory ratio index and the mixed ratio index can better fit the ideal value change curve. Compared with the application system based on parameter optimization, the difference between the experimental index and the ideal index is significantly reduced under the effect of the new system. It can not only solve the problem that the parallel chaotic data encryption processing behavior is uncontrollable, but also realize the secure transmission of network encrypted information, which meets the practical application requirements.

4 Conclusion

Aiming at the new parallel chaotic data security encryption system, the neural network system is taken as the entry point, and under the role of BP topology, the encryption processing module and data reading and writing interface module are combined to authenticate the parallel identity of the transmitted information by constructing a chaotic mapping relationship, so as to improve the encryption execution process of the system. From the perspective of experimental results, the trend of the memory

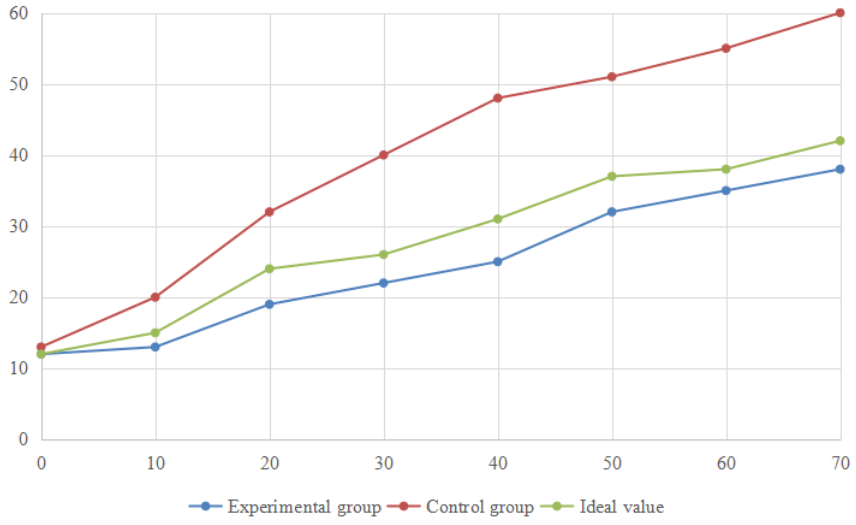


Figure 3: Memory ratio change curve

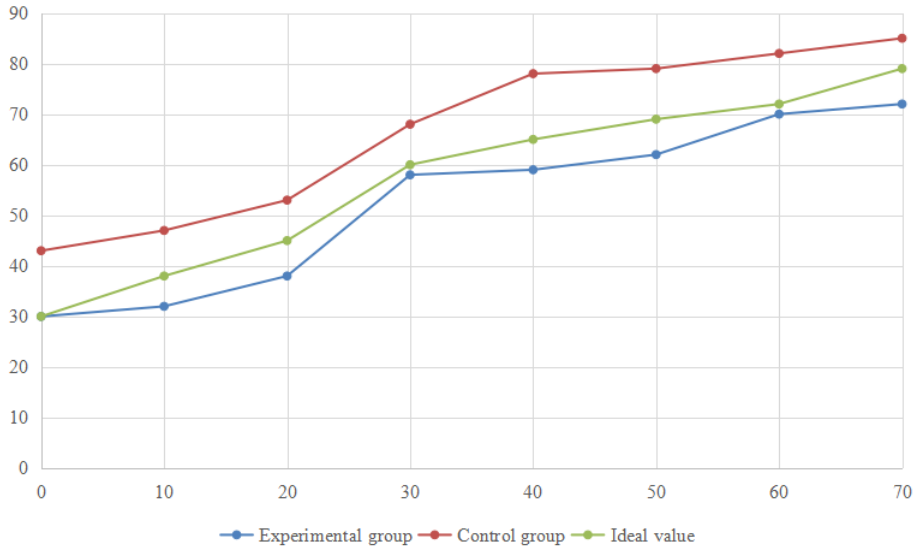


Figure 4: Mixing ratio change curve

ratio index and the mixed ratio index can fit the ideal numerical curve well, and the physical difference between them is low, which not only solves the problem of uncontrollable parallel chaotic data encryption processing, but also realizes the secure transmission of network encrypted information.

Acknowledgments

The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] K. Ahmad, M. Maabreh, M. Ghaly, K. Khan, J. Qadir, A. Al-Fuqaha, "Developing future human-centered smart cities: Critical analysis of smart city security, Data management, and Ethical challenges," *Computer Science Review*, Vol. 43, pp. 100452, 2022.

[2] K. K. Almuzaini, A. K. Sinhal, R. Ranjan, V. Goel, R. Shrivastava, and A. Halifa, "Key Aggregation Cryptosystem and Double Encryption Method for Cloud-Based Intelligent Machine Learning Techniques-Based Health Monitoring Systems," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[3] A. Bhattacharjya, "A holistic study on the use of blockchain technology in CPS and IoT architectures maintaining the CIA triad in data communication," *International Journal of Applied Mathematics and Computer Science*, vol. 32, no. 3, pp. 403-413, 2022.

[4] F. Bao, C. C. Lee, M. S. Hwang, "Cryptanalysis and improvement on batch verifying multiple RSA digital signatures", *Applied Mathematics and Computation*, vol. 172, no. 2, pp. 1195-1200, Jan. 2006.

[5] G. Cano-Quiveu, P. Ruiz-de-clavijo-Vazquez, M. J. Bellido, J. Juan-Chico, J. Viejo-Cortos, D. Guerrero-Martos and E. Ostua-Aranguena, "Embedded LUKS (E-LUKS): A Hardware Solution to IoT Security," *Electronics*, vol. 10, no. 23, pp. 3036, 2021.

[6] C. C. Chang, M. S. Hwang, "Parallel computation of the generating keys for RSA cryptosystems", *Electronics Letters*, vo1. 32, no. 15, pp. 1365–1366, 1996.

[7] C. C. Chang, M. S. Hwang, T. S. Chen, "A new encryption algorithm for image cryptosystems", *Journal of Systems and Software*, vol. 58, no. 2, pp. 83–91, 2001.

[8] Y. Guo, S. S. Ge, A. Arbi, "Stability of traveling waves solutions for nonlinear cellular neural networks with distributed delays," *Journal of Systems Science and Complexity*, vol. 35, no. 1, 18-31, 2022.

[9] A. Hosoyamada, Y. Sasaki, "Quantum collision attacks on reduced SHA-256 and SHA-512," in *Annual International Cryptology Conference. Cham: Springer International Publishing*, volume. 12825, pp. 616-646, 2021.

[10] L. C. Huang, M. S. Hwang, "Two-party authenticated multiple-key agreement based on elliptic curve discrete logarithm problem", *International Journal of Smart Home*, vol. 7, no. 1, pp. 9-18, 2013.

[11] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548–555, Jan. 2000.

[12] M. S. Hwang, Chii-Hwa Lee, "Secure access schemes in mobile database systems", *European Transactions on Telecommunications*, vol. 12, no. 4, pp. 303-310, 2001.

[13] M. S. Hwang, K. F. Hwang, I. C. Lin, "Cryptanalysis of the batch verifying multiple RSA digital signatures", *Informatica*, vol. 11, no. 1, pp. 15-18, Jan. 2000.

[14] M. S. Hwang, C. C. Lee, Y. C. Lai, "Traceability on RSA-based partially signature with low computation", *Applied Mathematics and Computation*, vol. 145, no. 2-3, pp. 465-468, Dec. 2003.

[15] M. S. Hwang, C. C. Lee, J. Z. Lee, C. C. Yang, "A secure protocol for bluetooth piconets using elliptic curve cryptography", *Telecommunication Systems*, vol. 29, no. 3, pp. 165-180, 2005.

[16] M. S. Hwang, T. H. Sun, C. C. Lee, "Achieving dynamic data guarantee and data confidentiality of public auditing in cloud storage service," *Journal of Circuits, Systems, and Computers*, vol. 26, no. 5, 2017.

[17] M. S. Hwang, S. F. Tzeng, C. S. Tsai, "Generalization of proxy signature based on elliptic curves," *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 73-84, 2004.

[18] M. S. Hwang and W. P. Yang, "A two-phase encryption scheme for enhancing database security", *Journal of Systems and Software*, vo1.31, no.12, pp. 257-265, 1995.

[19] M. S. Hwang, W. P. Yang, "Multilevel secure database encryption with subkeys", *Data & Knowledge Engineering*, vol. 22, no. 2, pp. 117-131, 1997.

[20] M. S. Hwang, W. P. Yang, "Integrating different semantics of classification levels in heterogeneous distributed database systems", *Journal of Applied Sciences*, vol. 2, no. 5, pp. 553-557, 2002.

[21] C. C. Lee, M. S. Hwang, L. H. Li, "A new key authentication scheme based on discrete logarithms", *Applied Mathematics and Computation*, vol. 139, no. 2, pp. 343-349, July 2003.

[22] B. Li, Y. Feng, Z. Xiong, W. Yang, G. Liu, "Research on AI security enhanced encryption algorithm of autonomous IoT systems," *Information sciences*, vol. 575, pp. 379-398, 2021.

[23] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245-255, 2003.

[24] C. W. Liu, W. F. Hsien, C. C. Yang, M. S. Hwang, "A survey of attribute-based access control with user re-

vocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900-916, 2016.

[25] Y. Liu, L. Wang, A. Qouneh, X. Fu, "Enabling PIM-based AES encryption for online video streaming," *Journal of Systems Architecture*, vol. 132, pp. 102734, 2022.

[26] E. D. Madyatmadja, A. N. Hakim, D. Sembiring, "Performance testing on Transparent Data Encryption for SQL Server's reliability and efficiency," *Journal of Big Data*, vol. 8, pp. 1-14, 2021.

[27] M. Malik, T. Patel, "Database security-attacks and control methods," *International Journal of Information*, vol. 6, no. 1/2, pp. 175-183, 2016.

[28] M. A. Midoun, X. Wang, M. Z. Talhaoui, "A sensitive dynamic mutual encryption system based on a new 1D chaotic map," *Optics and Lasers in Engineering*, vol. 139, pp. 106485, 2021.

[29] F. Musanna, D. Dangwal, S. Kumar, "Novel image encryption algorithm using fractional chaos and cellular neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 2205-2226, 2022.

[30] M. A. Nugroho and V. Suryani, "AADC 3: Active-Active Distributed Controller with 3-in-1 Asynchronous Heartbeat Synchronization Method in Software-Defined Networks," in *9th International Conference on Information and Communication Technology*, pp. 275-279, 2021.

[31] K. Sharma, A. Agrawal, D. Pandey, R. A. Khan, S. K. Dinkar, "RSA based encryption approach for preserving confidentiality of big data," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 5, pp. 2088-2097, 2022.

[32] M. Sharma, V. Choudhary, R. S. Bhatia, S. Malik, A. Raina, H. Khandelwal, "Leveraging the power of quantum computing for breaking RSA encryption," *Cyber-Physical Systems*, vol. 7, no. 2, pp. 73-92, 2021.

[33] E. Shmueli, R. Vaisenberg, E. Gudes, Y. Elovici, "Implementing a database encryption solution, design and implementation issues," *Computers & security*, vol. 44, pp. 33-50, 2014.

[34] L. Teng, H. Li, S. Yin, "Im-MobiShare: An improved privacy preserving scheme based on asymmetric encryption and bloom filter for users location sharing in social network," *Journal of Computers*, vol. 30, no. 3, pp. 59-71, 2019.

[35] L. Teng, Y. Qiao, M. Shafiq, G. Srivastava, A. R. Javed, T. R. Gadekallu, "FLPK-BiSeNet: Federated Learning Based on Priori Knowledge and Bilateral Segmentation Network for Image Edge Extraction," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 1529-1542, 2023.

[36] S. F. Tzeng, M. S. Hwang, "Digital signature with message recovery and its variants based on elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 61–71, Mar. 2004.

[37] S. F. Tzeng, C. Y. Yang, M. S. Hwang, "A new digital signature scheme based on factoring and discrete logarithms", *International Journal of Computer Mathematics*, vol. 81, no. 1, pp. 9-14, 2004.

[38] S. Ullah, J. Zheng, N. Din, M. T. Hussain, F. Ullah, M. Yousaf, "Elliptic Curve Cryptography; Applications, challenges, recent advances, and future trends: A comprehensive survey," *Computer Science Review*, vol. 47, pp. 100530, 2023.

[39] C. C. Wu, M. S. Hwang, S. J. Kao, "A new approach to the secret image sharing with steganography and authentication", *The Imaging Science Journal*, vol. 57, no. 3, pp. 140–151, 2009.

[40] C. C. Wu, S. J. Kao, and M. S. Hwang, "A high quality image sharing with steganography and adaptive authentication scheme", *Journal of Systems and Software*, vol. 84, no. 12, pp. 2196–2207, 2011.

[41] C. C. Yang, T. Y. Chang, M. S. Hwang, "A new anonymous conference key distribution system based on the elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 25, no. 2, pp. 141-145, 2003.

[42] M. Yang, I. Tjuawinata, K. Y. Lam, J. Zhao, L. Sun, "Secure hot path crowdsourcing with local differential privacy under fog computing architecture," *IEEE Transactions on Services Computing*, vol. 15, pp. 4, pp. 2188-2201, 2020.

[43] M. Yang, I. Tjuawinata, K. Y. Lam, T. Zhu, J. Zhao, "Differentially Private Distributed Frequency Estimation," *IEEE Transactions on Dependable and Secure Computing*, 2022.

[44] I. Yaqoob, K. Salah, R. Jayaraman, Y. Al-Hammadi, "Blockchain for healthcare data management: opportunities, challenges, and future recommendations," *Neural Computing and Applications*, vol. 34, pp. 11475-11490, 2022.

[45] B. You, X. Xiao, "Data encryption technology application in enterprise cost operation management based on cloud computing," *Soft Computing*, pp. 1-13, 2023.

[46] D. Zhang, M. Shafiq, L. Wang, G. Srivastava, S. Yin, "Privacy-preserving remote sensing images recognition based on limited visual cryptography," *CAAI Transactions on Intelligence Technology*, 2023.

[47] P. Zhang, C. Wang, C. Jiang and Z. Han, "Deep Reinforcement Learning Assisted Federated Learning Algorithm for Data Management of IIoT," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8475-8484, Dec. 2021.

Biography

Mengya Wei biography. Mengya Wei is with the School of Foreign Languages, Zhengzhou University of Science and Technology, Zhengzhou China. Her research interests include English education, English data security analysis.

Abnormal Traffic Detection Scheme Based on RBF Fuzzy Neural Network and Attention Mechanism in Robot Environment

Jintao Liu, Zhenxing Hao, Jiyue Wang, and Xi Zhang
(Corresponding author: Jintao Liu)
College of Mechanical Engineering, Zhengzhou University of Science & Technology
Zhengzhou Henan 450064, China
Email: aqiufenga@163.com

(Received July 19, 2023; Revised and Accepted Oct. 1, 2023; First Online Oct. 10, 2023)
The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection
Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

The rapid increase in terminal devices in the robot environment has brought many security risks. Detecting abnormal traffic efficiently has become an essential task in robot safety research. The existing detection methods have the problem of high computational cost. They cannot explicitly capture the relationship and structure of traffic data, so it is difficult to deal with new network attacks. This paper proposes a novel abnormal traffic detection scheme based on an RBF fuzzy neural network in a robot environment. Fuzzy rules are used to simulate the relationship between factors. The incremental fuzzy neural network training method and batch fuzzy neural network training method are combined to train the network. The risk level derived from the fuzzy rules is de-blurred to get the risk index of the information system. On this basis, the attention module is introduced to enhance the extraction of key features, enhance the interpretability of the model, and further improve the detection accuracy. Experimental results on the open data set CTU-13 show that the proposed method can detect traffic effectively and save more time than other advanced methods.

Keywords: Abnormal Traffic Detection; Attention Module; RBF Fuzzy Neural Network; Robot Environment

1 Introduction

With the development of the robot ecosystem, a large number of intelligent terminal devices are widely used in a number of IoT application fields, such as smart home [17,18], smart healthcare, smart transportation 4.0 and so on. However, the sharp increase in the number of robot terminal equipment has brought many serious security risks, and the complex network environment makes the data generated by robot equipment easy to be leaked, attacked or interrupted [11, 29, 39]. On the one hand, the robot terminal equipment is often limited by computing, memory, bandwidth and other resources, and its own limitations bring higher security challenges to the robot. On the other hand, there is a close correlation between robot devices. Once a device is invaded, it may lead to user privacy data disclosure, network infrastructure failure, network congestion or paralysis, etc., and even cause huge economic and social losses, seriously threatening enterprise and national security [40, 41].

In the past few years, the rise and development of machine learning and deep learning have promoted the research in the field of Internet of Things security [19,20,28], and various types of neural networks (such as convolutional neural networks [6], long short-term memory networks [37], and auto-encoders [31]) have been extensively applied in the intrusion detection of Internet of Things. Reference [2] proposed a hierarchical intrusion detection system based on three different classifiers (decision tree, JRP algorithm, random forest). The first two classifiers run in parallel and feed the results to the third classifier, achieving good results in the IDS2017 data set, but the system model was relatively simple and the accuracy and false positive rate are not ideal. In order to solve the problem of unbalanced samples, a CNN-FDC method based on convolutional neural network was proposed in reference [30]. After converting KDD-CUP99 data set into gray image, the original loss function was replaced with focal length loss, which weakened the influence caused by fewer attack samples. However, this model often had low accuracy in the face of high-dimensional data. Reference [10] proposed a hybrid deep learning model CNN-LSTM, which used long short-term memory network to learn the time features of high-dimensional traffic data. Compared with other advanced intrusion detection al-

gorithms [13, 14], although the accuracy of the model was 99.03%, the main disadvantage was that the method based on back-propagation random gradient was used to update the weight, which required a long time for training and updating, could not meet the low time ductility requirements of the Internet of Things system, and the operation cost was large. In reference [35], Transformer, a popular method in the field of natural language processing, was introduced to improve the model combined with the traffic data set, which improved the detection accuracy and reduced the delay. However, the trained sample was a statistic-based attack sample, and the detection effect was poor when it encountered propagating attacks, which could not meet the high dynamic requirements of the Internet of Things system [8, 9, 21, 24, 25].

Machine learning and deep learning methods are mostly applied to Euclidian Spaces with fixed neighbor nodes, but in real IoT scenarios, a large number of edge devices and sensors are connected together in a complex, non-linear manner, thus forming a non-Euclidian space with non-fixed neighboring nodes [15, 27, 38]. However, most of the traditional methods are shallow learning methods, which only analyze the anomalies of the traffic data of a single node from a statistical point of view, and do not explicitly learn the existing relationships or structures between variables [12, 32]. Therefore, the performance of conventional deep learning methods in processing non-European spatial data is still difficult to be satisfactory. Some cunning intruders will launch attacks with low-intensity and highly targeted abnormal traffic, in which the packets are very similar to legitimate traffic and do not cause significant changes at the level of statistical analysis [34], and such new attacks are often difficult to detect by traditional methods.

Based on neural network, this paper proposes a distributed abnormal traffic detection scheme for robot environment. RBF neural network is used to remove the message passing module in the network. The distributed traffic anomaly detection architecture will perform anomaly detection at the active node of each robot. The attention mechanism is introduced to calculate the attention of each adjacent node, and the detection accuracy of the model is further improved by optimizing the weight selection process of each fully connected layer.

2 Abnormal Network Traffic Detection Model

This paper considers deploying an AI distributed detection module on the side of the fog node or on the SDN edge transponder to replace the detection methods running on the virtual server or the cloud. These detection modules are implemented by a low-power AI processor on the edge transponder. Each distributed unit focuses on a subset of the data transmission business, including the detection of module information and abnormal status of neighboring nodes, and ultimately the localization of abnormal traffic detection.

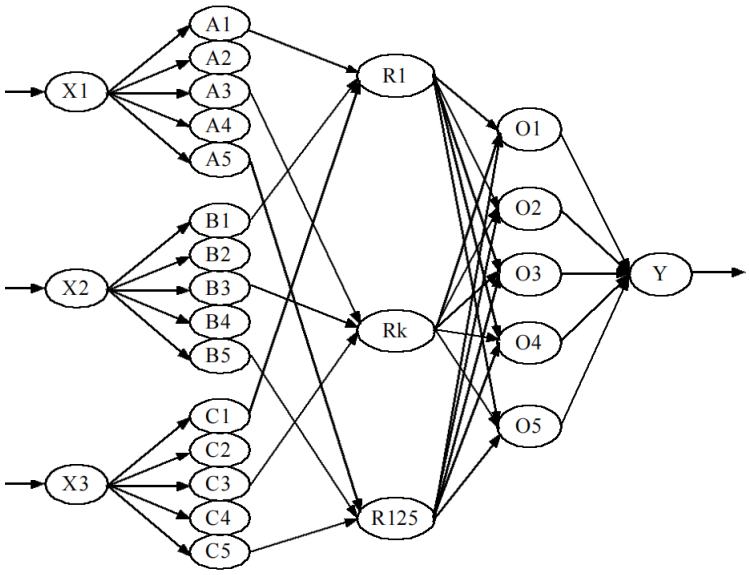


Figure 1: RBF fuzzy neural network structure

According to the basic principle of information security risk assessment and the characteristics of RBF fuzzy neural network [4, 23], the information structure diagram shown in Figure 1 is constructed, which is composed of five layers, namely, input layer, membership function layer, regularization layer, result layer and defuzzifying output layer.

First layer. Input layer, whose input is consistent with output. The three inputs are the factors that affect the information security risk level including the probability of threat occurrence, the value of assets, and the severity of vulnerability.

$$In_1(i) = out_1(i) = X = [x_1, x_2, x_3]. \tag{1}$$

Second layer. The blurring layer, for each input factor, is divided into 5 levels. For example, factor 1 is divided into $A_j(j \in [1,5])$, which respectively represents high, high, medium, low and very low, and the other two factors are similarly divided into B_j and $C_j(j \in [1,5])$. These factors belong to different levels according to different membership degrees. Here, the membership function adopts Gaussian radial basis function. c_{ij} and b_{ij} by are the center and base width of the membership function of the $j-th$ fuzzy set of the $i-th$ input variable, respectively.

$$out_2(i,j) = e^{\frac{(out_1(i)-c_{ij})^2}{(b_{ij})^2}}. \tag{2}$$

Third layer. In the rule layer, each node in layer 3 is only connected to a single level node of each factor in layer 2, then the number of nodes in layer 3 is 125, which is expressed as $R_i, i \in [1,125]$. That is, each rule is determined by three factors, and the input is the result

of the multiplication of the three.

$$Out_3(k) = out_2(1, j_1) \times out_2(2, j_2) \times out_2(3, j_3). \quad (3)$$

Fourth layer is the result layer. The nodes are represented by $O_i, i \in [1, 125]$. All rules produce different results with different combinations of weights. Each node of this layer is connected to the third layer. w_{kj} is the weight from the $k-th$ node of the third layer to the $j-th$ node of the fourth layer. The transfer function uses the $logsig()$ function to compress the output value between 0 and 1.

$$in_4(j) = \sum_{k=1}^{125} in_3(k) w_{kj}. \quad (4)$$

$$out_4(j) = logsig(in_4(j)). \quad (5)$$

The fifth layer is the deblurring layer. This layer has only one node, which is represented by Y. The 5 level values are combined according to certain weights to produce the final output result. w_j represents the weight of the fifth node of the fourth layer to the output layer, and the output function uses $logsig()$ function that compresses the output value between 0 and 1.

$$in_5 = \sum_{j=1}^{5} w_j^*(j). \quad (6)$$

$$out_5 = logsig(in_5). \quad (7)$$

As shown in Figure 2, this paper presents two independent models for updating the attributes of nodes and their corresponding edges. Based on this, a distributed detection unit deployed on the edge transponder is constructed. The core module of the architecture consists of edge RBF and node RBF, which are used to classify the state of nodes and edges respectively, and update the attributes of nodes and their corresponding edges. Edge detection units are used to classify features and predict the probability of anomalies on adjacent nodes, while node detection units are used to update features of nodes and calculate the probability of causing their own abnormal state.

Different from the conventional RBF neural network model, a communication channel is implemented in this paper, and the information exchange neighborhood is established in the channel, which is used to combine the information of edge RBF and node RBF. The inputs to the model represent three properties of the edge feature and five properties of the node feature, and each neuron is connected by a one-way link. In detail, the input and output are defined: suppose there is a node j and its adjacent nodes $i = 1, 2, \cdots, N$. The input of the edge is composed of the edge feature vector corresponding to the neighbor, the information of the node itself and the edge feature vector corresponding to the neighbor. Update the edge feature vector through the output of the fully connected layer. At the same time, the node RBF module also updates the node's own feature representation according to the collected information, and then concatenates the updated edge feature vector with the features of the $i-th$ node as the input of Softmax classifier [16, 42], and finally gets the anomaly probability of node j through classification. Compared with other centralized intrusion detection systems, this way of information exchange does not require explicit message passing and effectively reduces the resource occupation.

In the forward propagation of RBF, node information that plays an important role should be paid attention to, while node information that plays a secondary role should be ignored. In order to further improve the detection accuracy, this paper adds the attention mechanism module before the last layer classification. When each node updates the output of the hidden layer, different weights are assigned to each adjacent node by calculating the attention of adjacent nodes, and nodes with higher weights are taken as the focus of the neural network. The introduction of the attention mechanism reduces the computational burden of processing high-dimensional data, and makes the detection system more focused on finding significant relevant and useful information in the data, thus improving the output quality.

3 Experimental Results and Analysis

3.1 Experimental Environment

In order to evaluate the detection performance of the scheme in this paper, Python, NumPy, Pandas, Pytorch and other tools are used. Simulation experiments were performed on a 64-bit computer using Intel i9-9700K 16GB RAM, Nvidia GeForce RTX2080Ti 32GB and version 10.2 of CUDA.

3.2 Experimental Data

The data set used in this paper is CTU-13 [5, 33], which is a botnet traffic dataset captured at CTU University in 2011. The dataset contains 13 different attacks, with each packet containing information about various clients and servers. The network consists of 30 transponders and 170 iot devices that exchange data based on distributed devices in the CTU.

3.3 Evaluation Index

The performance index of abnormal traffic detection depends on the confusion matrix. In the confusion matrix, the true class (TP) is the correctly classified abnormal traffic instances; False positive class (FP) is a misclassified normal traffic instance; True and inverse class (TN)

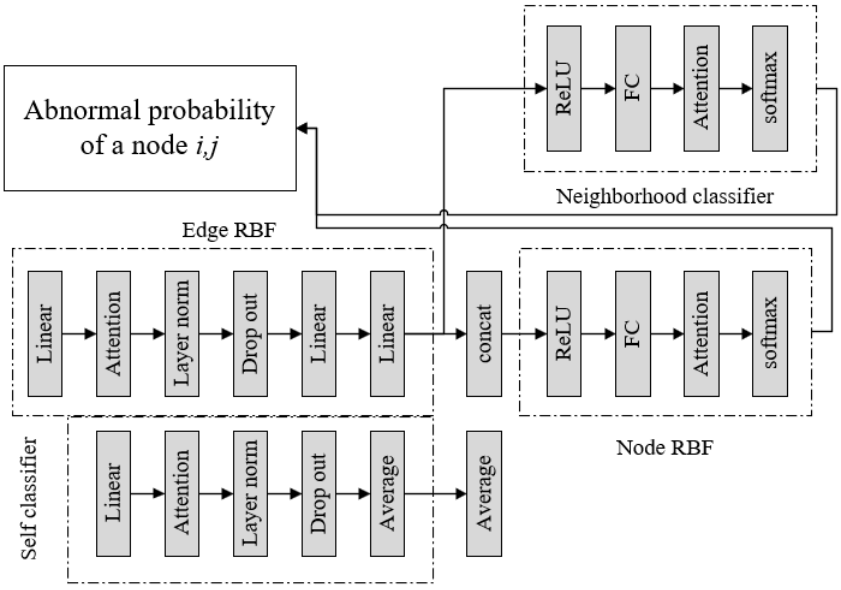


Figure 2: Structure of traffic anomaly detection

is a correctly classified normal traffic instance; False anti-class (FN) is an abnormal traffic instance that is misclassified. These 4 items are used to generate the following performance evaluation indicators [7].

Accuracy, that is, the ratio of the number of samples correctly classified by the model to the total number of samples, is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}. \tag{8}$$

Precision, that is, the ratio of the number of normal samples correctly classified by the model to the total number of normal samples, is calculated as follows:

$$Precision = \frac{TP}{TP + FP}. \tag{9}$$

Recall, that is, the ratio of the number of intrusion samples correctly classified by the model to the total number of correctly classified samples, is calculated as follows:

$$Recall = \frac{TP}{TP + FN}. \tag{10}$$

False Positive Rate (FPR) is the ratio of the number of normal samples wrongly reported as intrusions to the total number of normal samples. The formula is as follows:

$$FPR = \frac{FP}{FP + TN}. \tag{11}$$

3.4 Analysis of Experimental Results

In order to prove the advantages of the proposed scheme, classical machine learning algorithms and deep learning methods were used for experimental comparison on dataset CTU-13, including three machine learning methods and three deep learning methods. The Adam optimizer is used to train the model, uniformly setting the batch size to 1024, the learning rate to 0.0001, the batch to 30, and the dropout to 0.5. Pytorch is used to build a detection model, and the performance indicators of the experiment are shown in Table 1. The results show that the accuracy rate, accuracy rate, recall rate and false positive rate in the CTU-13 dataset are up to 0.9995, 0.9831, 0.9964 and 0.0041. This is due to the fact that the improved RBF neural network model in this paper can better learn complex features in large data sets, because the larger the data set, the more complex the communication mode, the more IP nodes and interaction edges. Compared with other methods, the improved model in this paper can more easily play an advantage and detect abnormal traffic more accurately. In addition, the introduction of attention mechanism also further improves the detection effect, making the detection performance of the proposed scheme better than other schemes.

Table 1: Detection comparison with different schemes/%

Scheme	Accuracy	Precision	Recall	FPR
Decision tree [1]	0.7884	0.7492	0.7789	0.1845
Naive Bayes [3]	0.8089	0.8175	0.8764	0.1553
SVM	0.8575	0.9028	0.9226	0.0774
PCA-SSH [36]	0.9587	0.9044	0.9360	0.0293
BGA [22]	0.9194	0.9217	0.9517	0.0171
DAE-GAN [26]	0.9726	0.9815	0.9705	0.0097
Proposed	0.9995	0.9831	0.9964	0.0041

Table 2: Computational complexity comparison of different schemes

Method	Number of trainable parameters/10^3	Training time/s	Testing time/s
PCA-SSH	18.7	168.4	10.41
BGA	23.3	172.6	12.51
C	43.7	266.2	18.61
DAE-GAN	6.4	98.4	9.27

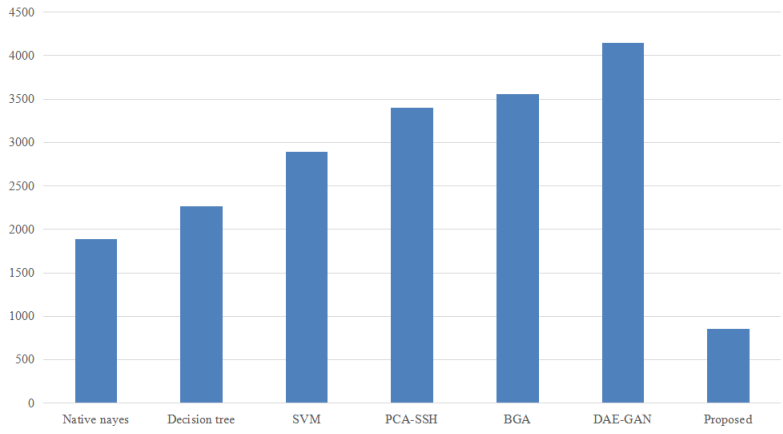


Figure 3: Bandwidth comparison of different schemes

To validate the advantages of deploying distributed detection units over centralized IDS, the paper also compares the performance of different schemes in terms of resource consumption and time overhead. The bar chart shows the resource consumption between different schemes, as shown in Figure 3. According to the data, it is not difficult to see that the minimum bandwidth consumption of the proposed scheme is only 856kb/s. Compared with other mainstream methods, the resource consumption is significantly reduced, because distributed anomaly detection does not need to transmit data to the IDS of the cloud server for calculation, and each detection unit reduces the resource consumption of centralized IDS with less bandwidth occupation.

Table 2 shows the number of training parameters and running time of the proposed scheme and other comparison schemes, using GPU to speed up the training of all models. It can be seen that by improving the message passing module of traditional RBF, the scheme in this paper has achieved a good improvement in time overhead, and the training time and training speed have been reduced. At the same time, there are few trainable parameters in the algorithm, which can realize efficient parallel computation.

4 Conclusions

In this paper, a distributed abnormal traffic detection scheme is proposed based on the complex characteristics of device nodes in robot environment and the requirement of low delay and high precision detection. The RBF convolutional neural network is optimized to replace the existing messaging module with an improved multi-layer perceptron to learn, making the model more suitable for iot environment. On this basis, combined with numerous characteristics of robot nodes, node RBF and edge RBF are designed to implement distributed traffic anomaly detection, realize localized abnormal traffic detection, and introduce attention mechanism to further improve the detection effect of the model. The experimental results show that the proposed scheme not only improves the detection accuracy effectively, but also reduces the overhead of network communication and speeds up the detection speed. The next step will be to perform graph structure analysis on more types of traffic datasets to train and test the model in a wider range of scenarios.

Acknowledgments

This work was supported by the Applied Research Program from Henan Provincial Department of Education, project No.23B460011. Key scientific research project plan of colleges and universities in Henan Province, project No.23B480003. Key scientific research project of Higher Education Department of Henan Province "Research on Obstacle Avoidance Detection and Safety Control of Industrial Robots" project No. 21B460017 and research on Vibration Analysis and Control of Chemical Fiber High-speed Winding Head Based on Particle Damping Project No. 182102210553. The authors gratefully

acknowledge the anonymous reviewers for their valuable comments.

References

[1] A. Aboah, M. Shoman, V. Mandal, S. Davami, Y. Adu-Gyamfi and A. Sharma, "A Vision-based System for Traffic Anomaly Detection using Deep Learning and Decision Trees," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Nashville, TN, USA*, pp. 4202-4207, 2021.

[2] A. Ahmim, L. Maglaras, M. A. Ferrag, M. Derdour and H. Janicke, "A Novel Hierarchical Intrusion Detection System Based on Decision Tree and Rules-Based Models," in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), Santorini, Greece*, pp. 228-233, 2019.

[3] A. Alsaleh and W. Binsaeedan, "The Influence of Salp Swarm Algorithm-Based Feature Selection on Network Anomaly Intrusion Detection," *IEEE Access*, vol. 9, pp. 112466-112477, 2021.

[4] C. Ai, L. Jia, M. Hong and C. Zhang, "Short-Term Road Speed Forecasting Based on Hybrid RBF Neural Network With the Aid of Fuzzy System-Based Techniques in Urban Traffic Flow," *IEEE Access*, vol. 8, pp. 69461-69470, 2020.

[5] A. Bansal, S. Mahapatra, "A comparative analysis of machine learning techniques for botnet detection," *Proceedings of the 10th international conference on security of information and networks*, pp. 91-98, 2017.

[6] L. Chen, J. Cao, K. Wu, Z. Zhang, "Application of generalized frequency response functions and improved convolutional neural network to fault diagnosis of heavy-duty industrial robot," *Robotics and Computer-Integrated Manufacturing*, vol. 73, pp. 102228, 2022.

[7] S. Chowdhury, M. Khanzadeh, R. Akula, F. Zhang, S. Zhang, H. Medal, M. Marufuzzaman, L. Bian, "Botnet detection using graph-based feature clustering," *Journal of Big Data*, vol. 4, pp. 1-23, 2017.

[8] T. H. Feng, W. T. Li, M. S. Hwang, "False data report filtering scheme in wireless sensor networks: A survey", *International Journal of Network Security*, vol. 17, no. 3, pp. 229-236, 2015.

[9] T. H. Feng, N. Y. Shih, M. S. Hwang, "A safety review on fuzzy-based relay selection in wireless sensor networks", *International Journal of Network Security*, vol. 17, no. 6, pp. 712-721, 2015.

[10] L. Fu, Q. Tang, P. Gao, J. Xin, J. Zhou, "Damage identification of long-span bridges using the hybrid of convolutional neural network and long short-term memory network," *Algorithms*, 2021, 14(6): 180.

[11] M. Hajiabbasi, E. Akhtarkavan and B. Majidi, "Cyber-Physical Customer Management for Internet of Robotic Things-Enabled Banking," *IEEE Access*, vol. 11, pp. 34062-34079, 2023.

[12] W. P. Hu, C.-B. Lin, J.-T. Wu, C.-Y. Yang, and M. S. Hwang, "Research on Privacy and Security of Federated Learning in Intelligent Plant Factory Systems," *International Journal of Network Security*, vol. 25, no. 2, pp. 377-384, 2023.

[13] C. H. Ling, W. F. Hsien, M. S. Hwang, "A double circular chain intrusion detection for cloud computing based on adjointVM approach", *International Journal of Network Security*, vol. 18, no. 2, pp. 397-400, 2016.

[14] L. C. Huang, M. S. Hwang, "Study of intrusion detection systems", *Journal of Electronic Science and Technology*, vol. 10, no. 3, pp. 269-275, 2012.

[15] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548–555, Jan. 2000.

[16] Z. Jiang, Z. Wang and E. -H. Kim, "Noise-Robust Fuzzy Classifier Designed With the Aid of Type-2 Fuzzy Clustering and Enhanced Learning," *in IEEE Access*, vol. 11, pp. 8108-8118, 2023.

[17] C. H. Lee, M. S. Hwang, W. P. Yang, "A novel application of the phone card and its authentication in mobile communications", *Journal of Information Science and Engineering*, vol. 15, no. 4, pp. 471-484, 1999.

[18] C. H. Lee, M. S. Hwang, W. P. Yang, "Phone card application and authentication in wireless communications", in *The International Federation for Information Processing*, pp. 323-329, 1996,

[19] C. T. Li, M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks",*Information Sciences*, vol. 181, no. 23, pp. 5333-5347, 2011.

[20] C. T. Li, M. S. Hwang and Y. P. Chu, "An efficient sensor-to-sensor authenticated path-key establishment scheme for secure communications in wireless sensor networks", *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 8, pp. 2107-2124, Aug. 2009.

[21] C. T. Li, C. C. Yang, M. S. Hwang, "A secure routing protocol with node selfishness resistance in MANETs", *International Journal of Mobile Communications*, vol. 10, no. 1, pp. 103-118, 2012.

[22] H. Li, H. Ge, H. Yang, J. Yan and Y. Sang, "An Abnormal Traffic Detection Model Combined BiIndRNN With Global Attention," *IEEE Access*, vol. 10, pp. 30899-30912, 2022.

[23] R. Li, Y. Yang and Q. Zhang, "Neural Network Based Adaptive SMO Design for TCS Fuzzy Descriptor Systems," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 10, pp. 2605-2618, 2020.

[24] W. T. Li, T. H. Feng, and M. S. Hwang, "Distributed detecting node replication attacks in wireless sensor networks: A survey," *International Journal of Network Security*, vol. 16, no. 5, pp. 323–330, 2014.

[25] W. T. Li, C. H. Ling, M. S. Hwang, "Group rekeying in wireless sensor networks: A survey", *International*

Journal of Network Security, vol. 16, no. 6, pp. 400-410, 2014.

[26] Z. Li, S. Chen, H. Dai, D. Xu, C. -K. Chu and B. Xiao, "Abnormal Traffic Detection: Traffic Feature Extraction and DAE-GAN With Efficient Data Augmentation," *IEEE Transactions on Reliability*, vol. 72, no. 2, pp. 498-510, 2023.

[27] I. C. Lin, H. H. Ou, M. S. Hwang, "A user authentication system using back-propagation network", *Neural Computing & Applications*, vol. 14, pp. 243-249, 2005.

[28] C. H. Ling, C. C. Lee, C. C. Yang, and M. S. Hwang, "A secure and efficient one-time password authentication scheme for WSN", *International Journal of Network Security*, vol. 19, no. 2, pp. 177-181, Mar. 2017.

[29] M. C. Lucas-Esta, B. Coll-Perales and J. Gozalvez, "Redundancy and Diversity in Wireless Networks to Support Mobile Industrial Applications in Industry 4.0," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 311-320, 2021.

[30] Y. Ma, Q. Yang, Y. Gao, "An internet of things intrusion detection method based on cnn-fdc," in *2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS). IEEE*, pp. 174-177, 2021.

[31] M Maimaitimin, K Watanabe, S Maeyama, "Stacked convolutional auto-encoders for surface recognition based on 3d point cloud data," *Artificial Life and Robotics*, vol. 22, pp. 259-264, 2017.

[32] A. Massing, M. G. Larson, A. Logg, "Efficient implementation of finite element methods on nonmatching and overlapping meshes in three dimensions," *SIAM Journal on Scientific Computing*, vol. 35, no. 1, pp. C23-C47, 2013.

[33] M. Putra, D. Hostiadi, T. Ahmad, "Botnet dataset with simultaneous attack activity," *Data in Brief*, vol. 45, pp. 108628, 2022.

[34] G. Stafford and L. L. Yu, "An Evaluation of the Effect of Spam on Twitter Trending Topics," in *2013 International Conference on Social Computing, Alexandria, VA, USA*, pp. 373-378, 2013.

[35] L. Tao, Z. Xie, D. Xu, K. Ma, Q. Qiu, S. Pan, B. Huang, "Geographic Named Entity Recognition by Employing Natural Language Processing and an Improved BERT Model," *ISPRS International Journal of Geo-Information*, vol. 11, no. 12, pp. 598, 2022.

[36] Z Wang, D Han, M Li, H Liu, M Cui, The abnormal traffic detection scheme based on PCA and SSH," *Connection Science*, vol. 34, no. 1, pp. 1201-1220, 2022.

[37] D. Wu, Y. Zhang, M. Ourak, K. Niu, J. Dankelman and E. V. Poorten, "Hysteresis Modeling of Robotic Catheters Based on Long Short-Term Memory Network for Improved Environment Reconstruction," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2106-2113, 2021.

[38] H. J. Wu, Y. H. Chang, M. S. Hwang, I. C. Lin, "Flexible RFID location system based on artificial neural networks for medical care facilities", *ACM SIGBED Review*, vol. 6, no. 2, pp. 1-8, 2009.

[39] Y. Wu, M. Li, G. Li and Y. Savaria, "Persistence Region Monitor With a Pheromone-Inspired Robot Swarm Sensor Network," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12093-12110, 2022.

[40] S. Yin, H. Li, A. A. Laghari, S. Karim, A. K. Jumani, "A Bagging Strategy-Based Kernel Extreme Learning Machine for Complex Network Intrusion Detection," *EAI Endorsed Transactions on Scalable Information Systems*, vol. 21, no. 33, e8, 2021.

[41] S. Yin, L. Wang, M. Shafiq, L. Teng, A. A. Laghari and M. F. Khan, "G2Grad-CAMRL: An Object Detection and Interpretation Model Based on Gradient-Weighted Class Activation Mapping and Reinforcement Learning in Remote Sensing Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 3583-3598, 2023.

[42] Q. Zhu and X. Zu, "A Softmax-Free Loss Function Based on Predefined Optimal-Distribution of Latent Features for Deep Learning Classifier," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 3, pp. 1386-1397, 2023.

Biography

Jintao Liu biography. Jintao Liu is with College of Mechanical Engineering & Zhengzhou University of Science & Technology. Research interests are Mechanical design, robotics, data security analysis.

Zhenxing Hao biography. It is required by the Zhenxing Hao is with College of Mechanical Engineering & Zhengzhou University of Science & Technology. Research interests are Mechanical design, robotics, data security analysis.

Jiyue Wang biography. Jiyue Wang is with College of Mechanical Engineering & Zhengzhou University of Science & Technology. Research interests are Mechanical design, robotics, data security analysis.

Xi Zhang biography. Xi Zhang is with College of Mechanical Engineering & Zhengzhou University of Science & Technology. Research interests are Mechanical design, robotics, data security analysis.

Effective Data Encryption of Sports Documents Based on Graph Neural Network and Data Fusion

Yuhang Li

(Corresponding author: Yuhang Li)

Art and Sports Department, Henan Technical College of Construction
No.51 Industrial Road, Erqi District, Zhengzhou City, 450064 China
Email: liyuhang@hnjs.edu.cn

(Received July 20, 2023; Revised and Accepted Oct. 1, 2023; First Online Oct. 10, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

The number of sports documents is huge, and there are many categories. The parallel use of valid data of mixed storage sports documents increases the difficulty of data encryption, resulting in a long encryption time. Therefore, an effective data encryption of sports documents based on graph neural network and data fusion is proposed in this paper. It uses a graph neural network to identify encrypted and unknown data and a hidden and visible layer unit to classify valid data. Combining a symmetric encryption scheme with a searchable algorithm generates a symmetric searchable encryption algorithm against information leakage to realize effective data encryption simulation of sports documents. The experimental results show that the proposed method can avoid adversarial attacks and greatly improve the classification performance on public data sets.

Keywords: Data Encryption; Data Fusion; Graph Neural Network; Sports Documents

1 Introduction

Under the background of continuous innovation of data processing technology, sports information construction has incorporated university document management into modern management work, generated massive electronic sports documents, and stored files of different time and athletes in a more convenient way. At present, the integration construction of sports documents and archives in universities is an important link of university digital management. In order to strengthen the security management of university Intranet documents, and at the same time reduce management costs and realize resource sharing, the encryption scheme is optimized, the encryption level is further strengthened by researching a new retrieval mechanism, and the similarity degree between data is calculated by dynamic confusion parameter adjustment, so as to achieve efficient encryption [8, 10, 28]. Due to the redundant characteristics of data such as duplication and overlap, some encryption schemes take the optimization of lightweight threshold re-encryption mechanism [7, 14, 16, 27] as the research idea, and use different levels of data encryption to protect medical documents from leakage [19, 31]. However, these research ideas only meet the requirements of sports document encryption at a certain stage. With the expansion of the university scale, the increase of the number of athletes, and the new goal of digital management requirements, it will increase the difficulty of document data encryption [4, 13].

To this end, a hybrid encryption protection scheme is also proposed for data security, and different data blocks are further encrypted by dividing data into fine granularity to improve encryption efficiency [28]. However, there are not only a large number of college sports data, but also a large number of categories, so we study the effective data encryption method of documents based on graph neural network.

2 ZIP Document Compression Key Structure Design

Sports document format includes DOC format, DOCX format, PDF format and TXT format four main types. When it parses the DOC document, some fields in the file header and some data in the information block belong to the key attributes of the DOC format. Random numbers are used to replace the original data, and ZIP compression is performed by extracting the location [12]. Compressed source file data area, directory area, and directory end

identifier three completely different parts, together make up the ZIP compressed file. However, because there is no important information in the first 30 bytes of the file header, the location extraction is performed after 30 bytes are skipped. The start identifier of the directory area is fixed to 4 bytes, and about 34 bytes is not important information, so the random location of the identifier area is extracted from the start of 34 bytes to the end of the directory [2, 9, 26]. According to the location extraction results of the above two links, the extracted data is first, and the key attributes are last. Since all subsequent files are compressed in ZIP format, the ZIP format of the compressed DOC document is represented by "1". Documents in DOCX format are a combination of XML documents and ZIP documents. Therefore, partial data and key attributes of DOCX documents are extracted, and random locations are extracted according to the same operation as the previous document. "2" is used here to represent the compressed ZIP format of DOCX documents [29]. A PDF document is composed of different objects, where the cross-reference table is the key attribute of the document, and the same substitution method is used to extract this data.

The document starts and ends with "trailer" and "%%EOF", and the next line of numbers provided by "startxref" is used as the offset address of the data table. Therefore, the PDF document is compressed three times according to the result, and the first time is compressed. Extract part of the compressed data and key data attributes related to the target data from the ZIP compressed document for the second time, and compress the ZIP document again. The third time extracts again according to the last compression result to obtain the final ZIP document of the new attribute, which is marked as "3". Finally, TXT document compression, the document is a text format without special structure, so it can be compressed according to the DOC document compression process, the document format is set to "4". By synthesizing the above four document formats, different ZIP document compression key structures are obtained [3,20].

According to the attributes of the above four different documents, different key structures are designed and the documents are compressed according to the structure. However, due to different encryption requirements, there are some categories of data to be encrypted and other categories of unknown data in the document, so the data of different documents are preprocessed before the compression of the document.

3 Data Processing Based on Graph Neural Network

3.1 Topology Construction

The proposed topology construction algorithm firstly divides the original traffic according to the session granularity, and the flow is defined as a data stream consisting of a series of packets with the same quintuple [5,6,21]. Session flow refers to all packets composed of bidirectional flows, that is, the source port and destination port, and the source IP address and destination IP address in the quintuple are interchangeable. Compared with unidirectional flow, researchers usually choose to use session flow for traffic classification, because the information contained in session flow is richer than that in unidirectional flow. In this article we will take advantage of the direction and order information used for packet interactions in the session flow. First, we extract the load length of each packet in the session flow, and then we can obtain the direction information of the packet according to the IP address of the packet in the session flow. We randomly select a session flow as an example to illustrate several elements of the topology.

1) Node. Given a specific session flow, each node in the topology formed by the session flow corresponds to each packet in the session flow, and the load length of the packet is used as the node feature. In order to reflect the direction of packet transmission, we retain the symbol of packet load length. We set the transmission direction of the first packet in the session flow to positive, that is, the load length of this packet and subsequent packets in the same direction is set to positive, and the load length of packets in the opposite direction is set to negative. In this paper, data packets continuously transmitted in the same direction are called clusters, even if only one data packet meets the conditions, it is also called clusters. The three nodes of -83, -41 and -393 corresponding to the three data packets continuously sent to the sender in Figure 1(a) are a cluster.

2) Edge. There are two types of edge types in the topology graph formed by the proposed algorithm: edge inside the cluster and edge outside the cluster. The inner edge of the cluster is successively connected to the continuous nodes in each cluster, and the outer edge of the cluster is connected to the two adjacent clusters before and after, that is, the first node of the two adjacent clusters is connected, and the last node is connected to the corresponding first node and the last node of the last cluster. When there is only one packet in a cluster, this packet is the first and last point, and only one edge can be added between the connected nodes. a topology consisting of information from the specific flow in Figure 1(a) is shown in Figure 1(b).

The method in this paper classifies traffic based on the differences in packet interaction characteristics of session flow between different types of traffic. The packet interaction characteristics adopted include the following four types: load length, direction, packet sequence and cluster characteristics of packets in session flow. The above features are shown in the interactive topology diagram of

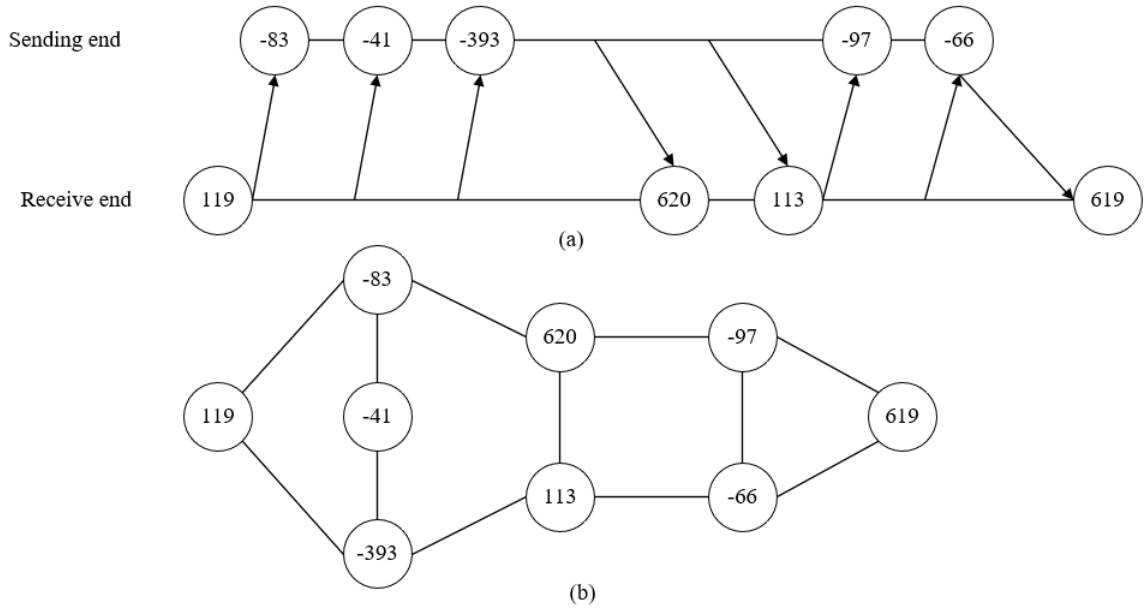


Figure 1: Topology based data flow representation. (a) Session flow packet interaction diagram; (b) Traffic interaction topology.

data flow conversion, as shown in Figure 1(b). The numbers on the nodes in the topology indicate the load length of the packet, and the positive and negative numbers represent the different directions of the packet transmission. Nodes in the topology corresponding to a series of packets continuously sent in the same direction in the session flow reflect the cluster characteristics of the session flow, and a series of nodes in each box in Figure 1(b) are represented as a cluster in Figure 1(a). The positions of nodes in the topology reflect the characteristics of packet sequence. For packets of different clusters, the horizontal positions of nodes in the topology represent the sequence of packet sending, and the sequence of packet sending represented by nodes on the left is earlier than that on the right. For packets in the same cluster, the packet sending order in the packet sequence is represented by the vertical position relationship of the nodes in the topology diagram, and the packet sending order represented by the upper node precedes the lower node.

3.2 Graph Classification Based on GNN

Given a set of topology graphs $G_1, \cdots, G_N \subseteq \Im$ and tags $y_1, \cdots, y_N \subseteq \Re$, the purpose of training the GNN model is to learn to predict the representation vector H_G of each topological label, i.e. $y_N = g(H_G)$. The GNN-based classification model consists of a graph convolution layer [24,30,33,35], which is used to extract features from the topological graph for training, and a fully connected layer, which is used for classification. The model uses the cross entropy loss function to measure the difference information between the actual label and the predicted label.

1) Graph convolution layer. In graph topology, each node is related to each other. The method of graph convolution operation is to aggregate the state information of neighbor nodes when updating node state information, that is, to aggregate the features of neighbor nodes to the central node. As follows:

$$H_{G_i}^{(l+1)} = \sigma(\tilde{D} - 0.5\tilde{A}\tilde{D} - 0.5H_{G_i}^{(l)}W^{(l)}). \quad (1)$$

Where $\tilde{A} = A + I_N$ is the adjacency matrix of an undirected graph $\Im$ introduced by the self-loop. I_N is the identity matrix. $\tilde{D}$ is the degree matrix of $\tilde{A}$. $\tilde{D}^{-0.5}\tilde{A}\tilde{D}$ is a symmetric normalization of $\tilde{A}$. $H_{G_i}^{(l)}$ is the output of the previous convolution layer. The initial eigenmatrix is set to $H_{G_i}^{(0)} = X_i$. X_i is the embedded feature of the node. $W^{(l)}$ is the weight matrix of layer l. $\sigma(\cdot)$ is a nonlinear activation function, and $ReLU(\cdot) = max(0, \cdot)$ is used in this paper.

2) Fully connected layer. GNN uses a linear function after the graph convolution layer to linearly transform the output data of the convolution layer and uses dropout functions to avoid overfitting. The output of the linear function is represented by the vector H_{G_i} as a feature of each topology graph G_i, and H_{G_i} needs to be mapped to a new potential space $H_{G_i} \in R^C$ to facilitate the following prediction process. C is the number of different elements in $\Re$, that is, the number of traffic categories to be classified. The softmax function is then used to obtain the predicted probability vector $\hat{y}_{ic}$, representing the likelihood that G_i belongs to each traffic type. As

follows:

$$\hat{y}_{ic} = softmax(H_{G_i}). \tag{2}$$

3) Loss function. The GNN in the classification model uses the cross entropy function as a loss function, which is mainly used to measure the difference information between two probability distributions, and is used to measure the similarity between the real label and the predicted label of the trained model when it is used as a loss function in deep learning. The advantage of using the cross entropy function as a loss function is that the cross entropy function [17, 34] is more convergent than the mean square error loss function. The formula of cross entropy function is as follows:

$$L = -\frac{1}{|N|}\sum_{i=1}^{|N|}\sum_{c=1}^{C} y_{ic} log(\hat{y}_{ic}). \tag{3}$$

Where $|N|$ is the number of trained samples and y_{ic} is the actual label.

4) Optimizer. The GNN in the classification model uses the Adam optimizer. Adam is a first-order optimization algorithm that can replace the traditional stochastic gradient descent process [11,18]. It can update the weight of neural network iteratively based on training data, which is simple and efficient and requires less memory.

3.3 Symmetric Searchable Encryption Based on Data Fusion

In order to reduce the problem of information leakage in the search process, the designed encryption structure and the symmetric searchable encryption algorithm against information leakage are fused to realize the encryption of valid data in sports documents. A symmetric encryption scheme is designed using negligible function and pseudo-random function [25, 32]. When the ignorable function faces the condition $f : M^* \rightarrow M^*$, if there is a sufficiently large value r, then the positive polynomial $f(r) < \frac{1}{p(r)}$, then the function f on r can be ignored. If the function $p(r)$ has a random condition such as $f : 0,1^m \times 0,1^r \rightarrow 0,1^n$, then the function can always be evaluated in polynomial time. To convert all polynomials generated by the above function to a pseudo-random permutation function requires that the function can be computed in polynomial time. Symmetric encryption The symmetric encryption scheme is set to SKE according to the above polynomial correlation conditions. Assume that the obtained input security parameters of the compressed key structure and the returned private key are p and H, respectively, represented by Gen; Set the input key and plaintext message to R and r respectively, which are represented by Enc. When the key is the same as the key generated during ciphertext generation, the decrypted message is set to G and represented by Dec. According to the above assumptions, the symmetric encryption scheme is:

$$S = (Gen, Enc, Dec). \tag{4}$$

According to the above scheme, a symmetric searchable encryption against information leakage is generated by incorporating the anti-information leakage search technology. The algorithm satisfies general searchable encryption security and encrypts compressed documents to ensure data privacy security. Using this algorithm to encrypt documents, build a secure index, and store them in the server. When the user needs to query a certain sports information, the server uses this algorithm to search the target document on the index structure. According to the above two parts, set the access mode of symmetric searchable encryption algorithm against information leakage, set the sports data word dictionary as λ, and set a random set of documents in the dictionary as W, then the access mode of historical data Z for k queries can be described by two-dimensional matrix $\partial(K)$, the formula is:

$$\partial(K) = [y_{1,1}, \cdots, y_{k,k}]^T. \tag{5}$$

Where $y_{k,k}$ represents the number of documents returned by the two queries. If the encryption function is defined as μ, the final encryption result is obtained according to the fusion design.

$$\mu(K) = \partial(K), lnfS(K). \tag{6}$$

Through the above calculation process, the symmetric searchable encryption result of anti-information leakage based on symmetric encryption scheme is obtained, and the effective data encryption of sports documents is realized.

4 Experimental Analysis and Results

Set up a simulation test environment, select Intel Core i3 VM with 1G memory and 30 GB hard disk, and select RedHat Linux6.4 64-bit server edition to verify the application effect of the graph-based neural network method for effective document data encryption. Sports documents from a university from July to December were selected as experimental test objects to analyze the application effect of the graph neural network, the overall encryption time and efficiency of the encryption method when encrypting valid data.

4.1 GNN Application Effect Test

The graph neural network is used to identify the valid data in the sports document. It sets the value of batch_size to 128 and Epochs to 10, that is, to update the parameters of 128 sets of sports data, and then the

GNN conducts 10 training times in total. At this point each Epoch will derive the training accuracy of the training set and the test set. At the same time, each Epoch will also obtain the training loss rate of different sets through Equation (2). The test results of the two groups are shown in Figure 2 and Figure 3.

According to the above test results, it can be seen that the training accuracy is in an increasing trend, and the training accuracy of the 10 groups is relatively stable, about 0.96. At the same time, with the increase of training times, the test results of the loss rate of the two sets of sets decreased rapidly, and finally stabilized within 0.1. The data recognition of the graph neural network was reliable. The probability condition coefficient of a certain category of valid data set by simulation is 0.85, and the effect of data classification by the neural network is tested. If the probability test result is greater than 0.85, the classified data is considered to belong to the expected data category; if the probability test result is less than 0.85, the classified sports data is considered to be in the sports document. For other types of unknown data, the classification probability obtained is shown in Table 1.

Table 1: Classification probability (CP) test results

Testing group	CP	Testing group	CP
1	0.889	5	7.36×10^{-14}
2	1.29×10^{-13}	6	0.988
3	3.43×10^{-6}	7	4.64×10^{-11}
4	1.08×10^{-11}	8	1.96×10^{-14}

According to the test results shown in the table above, the probability of group 6 is the highest, followed by the probability of group 1, and the results of both groups exceed the 0.85 specified in the experiment. We believe that these two data are valid data of the same category and belong to the preset data category, so the data classification of convolutional neural network can be determined with high accuracy. Based on the above test results, it can be seen that the convolutional neural network has a good effect on data recognition and classification.

4.2 Encryption Time

Graph neural network processes different document data, although it guarantees the reliability of encrypted data, but whether it affects the file compression time and thus the encryption efficiency remains to be verified. According to the analysis of four sports documents by this research method, the compression time and encryption time of valid data in DOC format, DOCX format, PDF format and TXT format were tested. In order to ensure the diversity of test results, the method sets 5 different key attribute extraction rates, and the obtained results are shown in Table 2.

Table 2: Typical states of SEIR model

Document type	Document size/MB	Compression time/s	Encryption time/s
DOC	20.6	1.28	1.69
DOC	21.6	1.33	1.86
DOC	22.6	1.29	1.89
DOC	23.6	1.38	2.13
DOC	24.6	1.46	2.59
DOCX	16.3	1.13	1.36
DOCX	17.3	1.18	1.55
DOCX	18.3	1.26	1.72
DOCX	19.2	1.36	1.95
DOCX	20.3	1.44	2.43
PDF	90.6	2.08	2.44
PDF	91.7	2.26	2.69
PDF	92.7	2.68	2.97
PDF	93.7	3.37	3.84
PDF	94.6	3.39	4.53
TXT	9.9	0.76	0.82
TXT	10.9	0.78	0.99
TXT	11.9	0.83	1.15
TXT	12.9	0.86	1.37
TXT	13.9	0.88	1.55

According to the test results shown in Table 2, it can be seen that the encryption time is related to the extraction rate of key attributes of the valid data of the sports Institute document. When the extraction rate is large, the encryption time increases. However, according to the above test results of compression time and encryption time, although the proportion of compression time is large in the overall encryption time, compression does not affect the final encryption time, that is, after the compression is completed, this method can quickly complete effective data encryption. Note In the process of data processing by using graph neural network, the data encryption time is not greatly increased, and the basic encryption efficiency can be ensured within an appropriate range.

4.3 Encryption Efficiency

Encryption efficiency is one of the key indicators that affect the practicability of the design method. For example, although RSA algorithm has many advantages [15], its encryption efficiency is not high, which limits the scope of use of this algorithm. Sports files of 10GB and 50GB in HDFS are selected as test objects, and valid data of block size is encrypted. The acceleration ratio is calculated to analyze the encryption efficiency of the proposed method. The formula is as follows.

$$\varepsilon = \frac{T1}{T2}. \tag{7}$$

Where ε represents the calculation result of accelera-

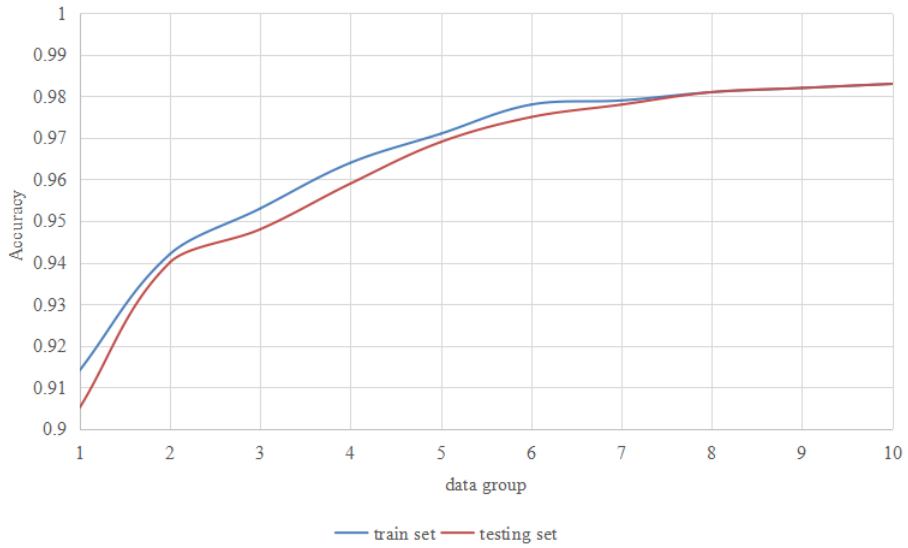


Figure 2: The training accuracy result of GNN.

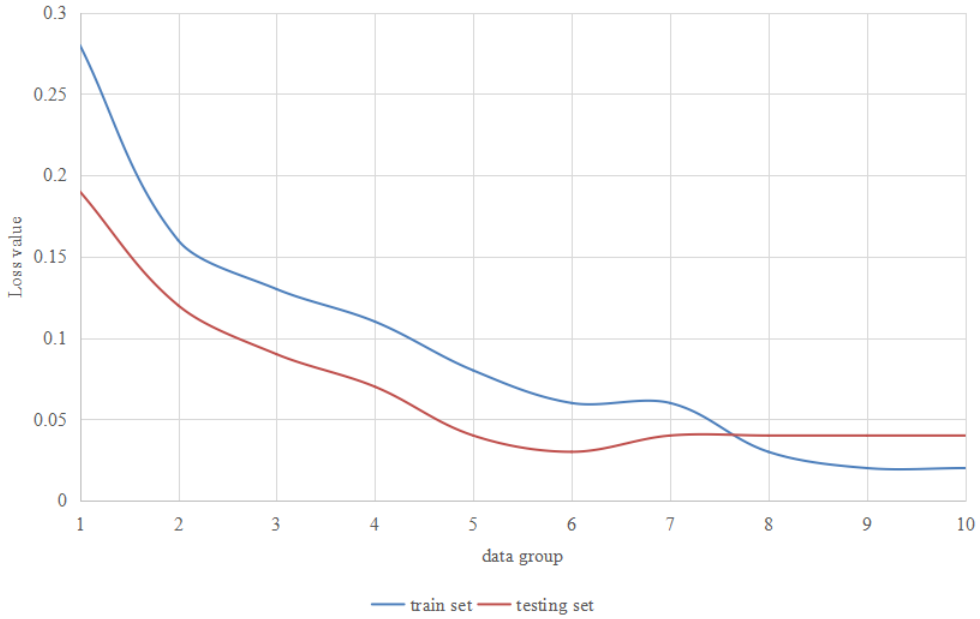


Figure 3: The training loss of GNN.

tion ratio; $T1$ and $T2$ indicate the encryption time and parallel encryption time on a single machine, respectively. According to the above test conditions, effective data encryption is carried out on sports documents, and the test results are shown in Table 3.

Table 3: Encryption time comparison result with different methods

Method	data size=10GB	data size=50GB
TDES [22]	4.12s	6.88s
CMPT [23]	2.38s	5.67s
DesPatNet25 [1]	1.24s	2.33s
Proposed	0.58s	0.67s

According to the test results shown in Table 3, when the total amount of data in sports documents increases, the encryption time increases, but the time consumed by the method in this paper is still relatively short. It shows that the encryption method in this paper can still have the advantage of parallel encryption in the face of massive medical data, and proves that the encryption efficiency of this method is high.

5 Conclusions

Aiming at the shortcomings of existing encryption technology, this study optimizes the use of graph neural network to pre-process valid data of documents, and solves the problem of low encryption efficiency caused by complex data in the past. The topology is entered into the GNN model for feature learning and classification. The node topology is used as the input of the model instead of the gray image, which avoids the influence of the adversarial attack on the gray image. In the future, the graph neural network can be optimized to further improve the efficiency of effective data encryption for sports documents.

Acknowledgments

The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] E. Akbal, P. Barua, S. Dogan, T. Tuncer, U. Rajendra Acharya, "DesPatNet25: Data encryption standard cipher model for accurate automated construction site monitoring with sound signals," *Expert Systems with Applications*, vol. 193, pp. 116447, 2022.

[2] I. Atzeni, J. Arnau and M. Kountouris, "Downlink cellular network analysis with LOS/NLOS propagation and elevated base stations," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 142-156, 2018.

[3] S. Chan, "Multi-attributes image analysis for the classification of web documents using unsupervised technique," in *International Conference on Intelligent Data Engineering and Automated Learning. Berlin, Heidelberg: Springer Berlin Heidelberg*, vol. 3578, pp. 78-85, 2005.

[4] C. C. Chang, M. S. Hwang, "Parallel computation of the generating keys for RSA cryptosystems", *Electronics Letters*, vo1. 32, no. 15, pp. 1365–1366, 1996.

[5] Q. Chen, G. Shou, Y. Hu and Y. Liu, "Topology construction of backbone network based on machine learning," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China*, pp. 1972-1976, 2018, doi: 10.1109/CompComm.2018.8780829.

[6] K. Chow, A. Sarkar, R. Elhesha, P. Cinaglia, A. Ay and T. Kahveci, "ANCA: Alignment-based network construction algorithm," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 2, pp. 512-524, 2021.

[7] P. S. Chung, C. W. Liu, M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments", *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[8] M. Fikri, S. Herdjunanto and A. Cahyadi, "On the performance similarity between exponential moving average and discrete linear Kalman filter," in *2019 Asia Pacific Conference on Research in Industrial and Systems Engineering (APCoRISE), Depok, Indonesia*, pp. 1-5, 2019, doi: 10.1109/APCoRISE46197.2019.9318810.

[9] A. Guo and M. Haenggi, "Spatial stochastic models and metrics for the structure of base stations in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5800-5812, 2013.

[10] C. Hou, Z. Li, J. Wu, "Unsupervised hash retrieval based on multiple similarity matrices and text self-attention mechanism," *Applied Intelligence*, vol. 52, pp. 7670-7685, 2022.

[11] X. Jia, X. Feng, H. Yong and D. Meng, "Weight decay with tailored adam on scale-invariant weights for better generalization," *IEEE Transactions on Neural Networks and Learning Systems*, 2022, doi: 10.1109/TNNLS.2022.3213536.

[12] W. Jiang, "Rate-distortion optimized image compression based on image inpainting," *Multimedia Tools and Applications*, vol. 75, pp. 919-933, 2016.

[13] C. T. Li, M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks",*Information Sciences*, vol. 181, no. 23, pp. 5333-5347, 2011.

[14] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245-255, 2003.

[15] Y. Liu, X. Shen, J. Liu, K. Peng, "Optical asymmetric JTC cryptosystem based on multiplication-division operation and RSA algorithm," *Optics & Laser Technology*, vol. 160, pp. 109042, 2023.

[16] E. J. L. Lu, M. S. Hwang, and C. J. Huang, "A new proxy signature scheme with revocation", *Applied Mathematics and Computation*, vol. 161, no. 3, PP. 799-806, Feb. 2005.

[17] L. Meng, L. Li, W. Xie, Y. Li, Z. Liu, "Time-sequential hesitant fuzzy entropy, cross-entropy and correlation coefficient and their application to decision making," *Engineering Applications of Artificial Intelligence*, vol. 123, pp. 106455, 2023.

[18] R. Mohapatra, S. Saha, C. A. C. Coello, A. Bhattacharya, S. S. Dhavala and S. Saha, "AdaSwarm: Augmenting gradient-based optimizers in deep learning with swarm intelligence," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 2, pp. 329-340, 2022.

[19] R. Precup, P. Angelov, B. Costa, M. Sayed-Mouchaweh, "An overview on fault diagnosis and nature-inspired optimal control of industrial process applications," *Computers in Industry*, vol. 74, pp. 75-94, 2015.

[20] R. Qiu, Y. Fu, J. Le, F. Zheng, G. Qi, C. Peng, Y. Zhang, Y. Li, and Y. Liu, "Deep security detection framework based on ATT&CK," *International Journal of Network Security*, vol. 25, no. 1, pp. 161-170, 2023.

[21] N. Radosavljević, P. Prvulović, D. Vujoević and A. Gavrić, "Traffic analysis of A3 topology construction protocol in wireless sensor networks," in *2022 21st International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, Bosnia and Herzegovina*, pp. 1-6, 2022, doi: 10.1109/INFOTEH53737.2022.9751272.

[22] M. N. Ramachandra, M. Srinivasa Rao, W. C. Lai, B. D. Parameshachari, J. A. Babu, and K. L. Hemalatha, "An efficient and secure big data storage in cloud environment by using triple data encryption standard," *Big Data and Cognitive Computing*, vol. 6, no. 4, pp. 101, 2022.

[23] X. Rong, D. Jiang, M. Zheng, X. Yu, X. Wang, "Meaningful data encryption scheme based on newly designed chaotic map and P-tensor product compressive sensing in WBANs," *Nonlinear Dynamics*, vol. 110, pp. 2831-2847, 2022.

[24] Y. Sui, X. Wang, J. Wu, M. Lin, X. He, T. S. Chua, "Causal attention for interpretable and generalizable graph classification," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1696-1705, 2022.

[25] X. Sun, S. Li, "Multi-sensor network tracking research utilizing searchable encryption algorithm in the cloud computing environment," *PeerJ Computer Science*, vol. 9, pp. e1433, 2023.

[26] L. Teng, H. Li and S. Yin, "IM-MobiShare: An improved privacy preserving scheme based on asymmetric encryption and bloom filter for users location sharing in social network," *Journal of Computers (Taiwan)*, vol. 30, no. 3, pp. 59-71. 2019.

[27] S. F. Tzeng, M. S. Hwang, C. Y. Yang, "An improvement of nonrepudiable threshold proxy signature scheme with known signers", *Computers & Security*, vol. 23, no. 2, pp. 174-178, Apr. 2004.

[28] X. Wang, S. Yin, M. Shafiq, A. A. Laghari, S. Karim, O. Cheikhrouhou, W. Alhakami, H. Hamam, "A new V-Net convolutional neural network based on four-dimensional hyperchaotic system for medical image encryption," *Security and Communication Networks*, vol. 2022, pp. 1-14, 2022.

[29] X. Wang, H. Zhang, Y. Tian and V. C. M. Leung, "Modeling and analysis of aerial base station-assisted cellular networks in Finite Areas Under LoS and NLoS Propagation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6985-7000, 2018.

[30] L. Wei, H. Zhao, Z. He, Q. Yao, "Neural architecture search for GNN-based graph classification," *ACM Transactions on Information Systems*, 2023. https://doi.org/10.1145/3584945.

[31] Q. Wei, Z. Liao, R. Song, P. Zhang, Z. Wang and J. Xiao, "Self-learning optimal control for ice-storage air conditioning systems via data-based adaptive dynamic programming," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 4, pp. 3599-3608, 2021.

[32] W. Wei, J. Wang, Z. Yan, W. Ding, "EPMDroid: Efficient and privacy-preserving malware detection based on SGX through data fusion," *Information Fusion*, vol. 82, pp. 43-57, 2022.

[33] B. Wu, X. Yang, S. Pan and X. Yuan, "Adapting membership inference attacks to GNN for graph classification: Approaches and implications," in *2021 IEEE International Conference on Data Mining (ICDM), Auckland, New Zealand*, pp. 1421-1426, 2021, doi: 10.1109/ICDM51629.2021.00182.

[34] J. Yu, H. Li, S. Yin, "Dynamic gesture recognition based on deep learning in human-to-computer interfaces," *Journal of Applied Science and Engineering*, vol. 23, no. 1, pp. 31-38, 2020.

[35] X. Zuo, H. Yuan, B. Yang, H. Wang, Y. Wang, "Exploring graph capsual network and graphormer for graph classification," *Information Sciences*, vol. 640, pp. 119045, 2023.

Biography

Yuhang Li biography. Yuhang Li is with Art and sports Department & Henan Technical College of Construction. His research interests include sports data analysis, big data security, and artificial intelligence.

Data Encryption Security of Track and Field Big Data Based on Deep Residual Network and Elliptic Curve Cryptography

Yiliang Chen, Jie Ren, and Chunlin Luo

(Corresponding author: Chunlin Luo)

College of Physical Education and Training, Harbin Sport University
No.1, Dacheng Street, Nangang District, Harbin City 150008, China
Email: zzll_201@foxmail.com

(Received July 21, 2023; Revised and Accepted Oct. 1, 2023; First Online Oct. 11, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

The traditional data privacy encryption protection algorithm mainly encrypts the data directly, ignores the classification between layers, and the complicated data encryption hierarchy calculation results in inaccurate data privacy encryption and easy leakage of sports data. Therefore, we propose a data encryption security method of track and field based on deep residual network and elliptic curve cryptography. An arithmetic coding model for sports data privacy protection in a hybrid cloud environment is constructed under an elliptic curve cryptosystem. Public key encryption is exhaustively searched based on optimizing the elliptic curve equation, and the master key of privacy protection data encryption is constructed. Based on a deep residual network, a threshold cryptosystem is established, and symbol frequency detection is carried out in the link layer of the hybrid cloud platform. A private key-sharing scheme is constructed, and the user's private key is recovered according to the bilinear cyclic shift method, which realizes privacy protection. Simulation results show that the improved method improves the security and robustness of data privacy encryption protection.

Keywords: Data Privacy Encryption Protection; Deep Residual Network; Elliptic Curve Cryptography; Track And Field Data

1 Introduction

In recent years, with the large-scale popularization of the Internet and the continuous development of the technical level of mobile devices, the scale of data volume is growing rapidly in an exponential form, and the types of data are also diverse and complicated [11, 19, 27, 30]. It can be seen from the characteristics of big data such as huge volume, wide range of sources, diversity and complexity that these big data contain a lot of valuable information. However, existing processing technologies are unable to effectively encrypt and protect data privacy in the big data environment [16, 17, 26, 34]. In the big data environment, the data privacy encryption protection optimization method can meet the data protection requirements of various types and conditions without disclosing the original data, which is the fundamental way to solve the above problems, and has attracted the attention of many experts and scholars [31, 36]. Because of the profound development significance of data privacy encryption protection optimization method in the big data environment, it has become the focus of research in the industry, has received wide attention, and there are many good methods.

Reference [22] proposed a data privacy encryption protection method in big data environment based on fault-tolerant privacy protection data aggregation algorithm. This method used elliptic curve encryption scheme to encrypt private data in big data environment. By calculating the topology of data privacy encryption track chart, the repair aggregate value of data encryption was improved, and the data privacy encryption protection algorithm was improved in big data environment. This method was robust, but it had the problem of poor scalability. Reference [3] focused on the method of data privacy encryption protection in big data environment based on Bloom filter algorithm. The method used Bloom filter to transform the data that needed to be encrypted into operator set elements for encryption, thus realizing the optimization of data privacy encryption protection algorithm in big data environment. The algorithm was short in time, but had the problem of poor data encryption stability. Reference [25] studied the data privacy encryp-

tion protection method in big data environment based on cloud computing algorithm. This method integrated the CES encryption scheme with the fuzzy retrieval of cloud data, and effectively improved the security and stability of data privacy encryption protection in the big data environment under the cloud computing platform. The cost of this method was low, but when the current algorithm was used for privacy encryption protection, the calculation was more complicated, and the data privacy encryption took a long time. Reference [4] proposed a method for data anonymization privacy protection based on differential privacy, adopted chaotic public key encryption scheme to construct a differential privacy anonymization privacy protection model, and used Laplace data clustering mechanism to classify and process data under different privacy protection budgets to improve the efficiency of data encryption. However, this method had poor anti-attack ability against high-intensity plaintext attacks, which easily leaded to information leakage. In reference [28], a big data security and privacy protection model based on data release and analysis was proposed. SuLQ framework was introduced to design Hash function for privacy protection communication data encryption in hybrid cloud environment, and encryption keys were constructed in a limited domain to improve the depth of data encryption. However, the calculation cost of this method was large, and the real-time performance of privacy protection calculation was not good.

In this paper, a track and field data privacy protection model based on deep residual network and elliptic curve encryption is proposed. Firstly, an arithmetic coding model for privacy protection in hybrid cloud environment is constructed under elliptic curve cryptosystem, and the exhaustive search of public key cryptography is carried out with the optimization of elliptic curve equation. Then the master key of privacy protection data encryption is constructed based on deep residual network, and the threshold cryptosystem is established to realize privacy protection. Finally, the simulation experiment of privacy protection and data encryption is carried out, and the validity conclusion is drawn.

2 Related Works

2.1 Elliptic Curve Cryptosystem

This paper adopts an encryption technology based on Elliptic Curve Cryptography (ECC) to protect information privacy in hybrid cloud environment [7–10,18,29,33]. First, the elliptic curve encryption model [20,21,35] is analyzed. The elliptic curve equation of user data encryption is generally defined as follows:

$$f(x)=\begin{cases} x/p & x\in[0,p] \\ (1-x)/p & x\in[p,1] \end{cases} \tag{1}$$

The elliptic curve equation adjusts the extended range of information encryption by the parameter p. When $p=0.5$, this maps to a standard normal distribution model.

For ease of description, in the mapping area of the elliptic curve, the length of the bit sequence of the privacy protection information to be encrypted is defined as n than that of the binary encoded sequence as $E=\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. Firstly, based on affine transformation $x_i=2\varepsilon_i-1$, the code elements expressed as binary coded bits are constructed to obtain a new sequence $X=x_1,x_2,\cdots,x_n$ that needs privacy protection in hybrid cloud environment, and the binomial and $S_n=x_1+x_2+\cdots,x_n$ of encrypted bit sequence X are calculated. The hashing value of the message is calculated by taking the extreme point $\frac{S_n}{\sqrt{n}}$ of the elliptic curve S_n and the point G belonging to E, which $\frac{S_n}{\sqrt{n}}$ is close to the standard normal n-state distribution function. When $\underset{n\to\infty}{limit}\,P(\frac{S_n}{\sqrt{n}}\le z)=\varphi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-u^2/2}du$. For cyclic point group $S_{obs}=\frac{S_n}{\sqrt{n}}$, the convergence function of elliptic curve encryption is obtained according to the normal distribution property.

2.2 Construction of Network Classification Model

Although ResNet residuals have changed the problems of degradation and gradient disappearance in the process of deepening the network layers of traditional convolutional neural networks, they are faced with problems such as information loss, excessive complexity and excessive computation. Based on the ResNet-50 network structure, the network model in this paper solves the above problems by changing the network structure to create a new network structure, and improves the accuracy of classification.

The ResNet-50 network architecture consists of an input backbone, four subsequent stages, and a final output layer [13,14,23]. The input trunk has a 7×7 convolution kernel and the output channel size is 64 with a step length of 2, followed by a 3×3 maximum pooling layer with a step length of 2. The input backbone reduces the width of the input image, quadruples the height of the input image, and increases the channel size to 64. The ResNet-50 network architecture starts with stage2, with each phase starting with a down-sampling module, followed by two residual modules. In the down-sampling module, there are two paths A and B, and path A contains two convolutional layers where the size of the convolution kernel is 1×1, 3×3 and 1×1, respectively. The first convolution layer has a step size of 2, which can reduce the width and height of the input image by half. The output channel of the last convolution layer is four times that of the first two convolution layers. This structure is called the bottleneck structure. Path B uses A 1×1 convolution kernel with step size 2 to convert the size of the output image to the same as path A. Summing the two paths can get the output of the down-sampling module.

The classification network model in this paper is a further improvement on the ResNet-50 network structure. By observing the ResNet-50 network structure, we can see that the input backbone consists of 7×7 convolution

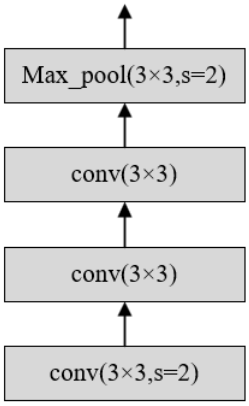


Figure 1: Input structure of the network model in this paper

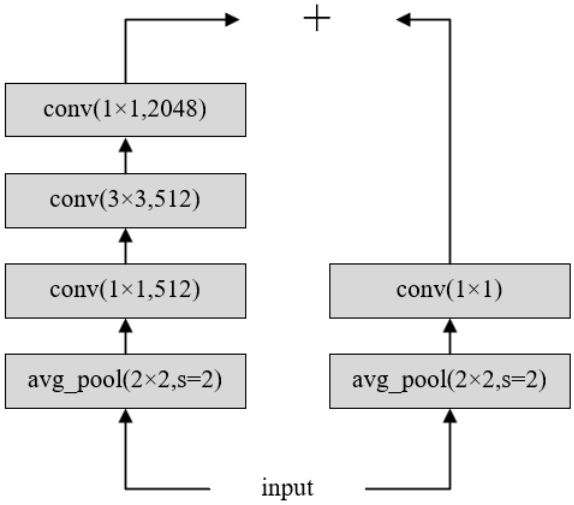


Figure 2: Structure chart of down-sampling module

nuclei, etc., and the calculation cost of the convolution layer is the square of the width of the convolution kernel or the height of the convolution kernel. Therefore, the 7×7 convolution kernel in the input backbone requires 6 times more computation than the 3×3 convolution kernel. The classification network model in this paper replaces the 7×7 convolution kernel in the input backbone with three 3×3 convolution nuclei, where the output channel size of the first and second 3×3 convolution layer is 32 and the step size is 2. The output channel size of the last convolutional layer is 64, which greatly reduces the calculation cost of classification network and the calculation amount of network model while ensuring the consistency with the previous output trunk information. The input backbone structure of the classification network model in this paper is shown in Figure 1.

At the same time, the classification network model changes the down-sampling module, deepens the number of network layers on the basis of the original, and adds average pooling layer. The down-sampling module of ResNet-50 network structure has two paths, path A and path B, and the convolutional layer in path A and path B has lost 75% of the information of the input feature map. Since the first convolution layer of both paths A and B has a step size of 2 and a convolution kernel of 1×1, both paths lose a large amount of input feature information. In this paper, the classification network model adds an average pooling layer before the first convolution layer of paths A and B. The convolution kernel size of the average pooling layer is 2×2, and the step size is 2. The integrity of the input feature information is ensured and the accuracy of classification is improved while the size of the output feature map is not changed. Moreover, the addition of the average pooling layer does not increase the calculation amount and calculation cost significantly. The structure of the down-sampling module of the classification network model in this paper is shown in Figure 2.

2.3 Code Frequency Detection and Privacy Protection

On the basis of establishing the threshold cryptosystem combined with the identity authentication scheme, the token frequency detection is carried out in the link layer of the hybrid cloud platform, the private key sharing scheme is constructed, the user private key is recovered according to the bilinear cyclic shift method, and the privacy protection optimization model is constructed [6]. The implementation steps of the improved model are described in detail as follows.

Step 1. Calculate $rsk_{ID_i} = sr_{i1} = g_1^{r-c/d}$ for each cyclic point group in the privacy protection sequence in the hybrid cloud computing environment, and record it in the $rsk - list$ list;

Step 2. Divide the privacy protection bit sequence of length n into $N = [\frac{n}{M}]$ bit blocks to form the required encryption sequence $x_1, x_2, \cdots, x_r$, detect the symbol frequency of continuous features in the bit sequence, and output the ciphertext sequence $c_i = t_i \oplus x_i$, $i = 1, 2, \cdots, r$;

Step 3. Construct an elliptic curve function conforming to the standard normal distribution, update each privacy protection key, and search the optimal value in the elliptic curve equation in the following way:

$$KC_1 = KC_1 \oplus t_j, t_{j+1}, \cdots, t_{j+m-2}. \tag{2}$$

$$KS_1 = KS_1 \oplus t_{j+m-2}. \tag{3}$$

Where $j = rmod128$, when $j > r-m$, then $j = j-m$, and if $j = 0$ then $j = m$. The in-block frequency of the bit sequence to be identified is counted, the initial value of the elliptic curve adjustment parameter is updated to $x_0 = x_{[r/16]+1}$, and the remaining privacy

data is encrypted according to the above steps to complete the encryption of privacy protection data.

2.4 Track and Field Data Privacy Encryption in Big Data Environment

In the process of data privacy optimization and encryption protection in the big data environment, the private data in the big data environment is converted into numerical values according to the theoretical idea of OPES +, and the arranged location data is divided into buckets to ensure a balanced distribution, and the number of points in each bucket is less than the given threshold. The protection sequence encryption algorithm is used to encrypt the data in the bucket, and keep the order of the encrypted values unchanged, and the data privacy encryption process is restricted to the solution of the encryption function. The specific steps are described below:

1) In the process of data privacy optimization and encryption protection in a big data environment, the unit of availability is used to represent the latitude and longitude of the data, and the theoretical idea of OPES+ is used to convert the private data in a big data environment into a numerical value for expression. If the original data value is expressed by availability ", minute and second ", the following formula can be obtained:

$$Decimal = Degrees + Min/60 + Sec/3600. \quad (4)$$

2) The arranged location data is divided into buckets to achieve a balanced distribution. Assuming that in the plaintext space, p represents the original spatial numerical data, p is divided into $p = B_1, B_2, \cdots, B_m$, then B_m needs to satisfy the division of any two buckets, and the above process is defined as bucket partitioning.

3) In the process of data privacy optimization and encryption protection in a big data environment, it is assumed that $SPlit(B_m)$ represents the partition function, and the function is to linearly divide B_m into two sub-buckets until the number of points in each bucket is less than the given threshold. Calculate the linear expected value represented by $p_i (p_i \in B_m)$, which can be expressed using the following formula:

$$p_i = \frac{A \times SPlit(B_m) - (p_k - p_j)}{k - j + 1}. \quad (5)$$

Where $A = \sqrt{p = B_1, B_2, \cdots, B_m}$, $(p_k - p_j)$ represents the range parameter of the data partition, $j+1$ represents the vector of the data location, and k represents the data range.

In the process of data privacy optimization and encryption protection in a big data environment, according to the above calculation, it can be concluded that the point farthest from the expected value needs to be selected for recursive division.

4) In the process of optimizing encryption protection for data privacy in a big data environment, the protection sequence encryption algorithm is used to encrypt the data in the bucket, and the size order of the encrypted values is kept unchanged. Calculate the encryption function represented by $M(P)$, if the ciphertext space represented by $C = c_1, c_2, \cdots, c_n$ is encrypted by the plaintext space represented by $p = p_1, p_2, \cdots, p_n$, and ϖ_i satisfies the following formula.

$$\varpi_i = M(P) - \frac{C \times p_i}{p}. \quad (6)$$

Where ϖ_i stands for barrel width.

In the process of data privacy optimization and encryption protection in a big data environment, it is assumed that ϖ_i^c represents the width of the encrypted ciphertext, so the following formula should be satisfied.

$$\varpi_i^c = M(P) - \frac{C \times \varpi_i}{p}. \quad (7)$$

5) In the process of data privacy optimization and encryption protection in a big data environment, for any value in bucket B_i represented by p_j, and c_j represents its encrypted ciphertext value, c_j can be calculated using the following formula.

$$c_j = \sum_{j=1}^{i-1} \varpi_j^c + (P_j - \sum_{i=1}^{j-1} \varpi_i). \quad (8)$$

6) In the process of data privacy optimization and encryption protection in a big data environment, the following formula is used to constrain the data privacy encryption process to solve the encryption function represented by $M(P)$, and make the distribution of $M(P)$ more balanced.

$$M(P) = Z(\frac{q}{2r} \mp c_j). \quad (9)$$

Where $\frac{q}{2r}$ represents the quadratic coefficient of data encryption, and Z represents the range parameter of $M(P)$.

7) In the process of data privacy optimization and encryption protection in the big data environment, according to the above conclusions, it can be concluded that the data privacy decryption process is opposite to the encryption process.

To sum up, it can be seen that In the process of data privacy optimization and encryption protection in the big data environment, OPES is the first basis The theoretical

idea of "+" converts private data in a big data environment into numerical values for representation, divides the arranged location data into buckets for balanced distribution, and promotes the number of points in each bucket to be less than the given threshold value. The data in the bucket is encrypted using the protection sequence encryption algorithm, and the size order of the encrypted values is kept unchanged, and the data privacy encryption process is restricted to the encryption function The solution of the number lays a foundation for the realization of data privacy optimization and encryption protection in the big data environment.

2.5 Loss Function

In order to better distinguish the feature information, a complex loss function soft-center loss is proposed, which can increase the spacing between classes and reduce the fit degree. softmax classifier in ResNet-50 network can not distinguish the extracted features well [15,32], which will cause the model to be overconfident and even cause the intra-class spacing to be larger than the inter-class spacing. In addition, from the perspective of clustering, large spacing between classes can ensure the recognition of extracted features. The central loss function can solve the problem of large intra-class spacing, so this paper combines the central loss function center-loss and softmax-loss together as the loss function of the classification network model in this paper.

softmax-loss is basically a standard method for single-class classification problems. For training sample set $(s_i, y_i)_{i=1}^{N}$, $s_i \in R^{m\times n}$, $y_i \in (1, 2, \cdots, c)$. c is the total number of categories of training samples, and the data set is mapped $y = p(y|x)$ through feature learning, $x_i \in R^m$, the loss function L_s is:

$$L_s = -\frac{1}{N}[\sum_{i=1}^{N} \frac{e^{x_j^T}\theta_i}{\sum_{j=1}^{c} e^{x_j^T}\theta_i}]. \tag{10}$$

Where N indicates the minimum batch size and θ indicates the model parameter. In the training process, the model parameter $\theta = (\theta_1, \theta_2, \cdots, \theta_c)$ is obtained by gradient descent, and the global optimal solution of the loss function is obtained.

The central loss function is defined as:

$$L_C = 0.5\sum_{i=1}^{m} ||x_i - c_{yi}||_2^2. \tag{11}$$

Where c_{yi} represents the feature center of the $y_i - th$ category and the average feature of all sample features of the category corresponding to sample i. By calculating the features of all samples of the same class and then calculating the average, this method is impractical, because the training samples in this paper are relatively large. In this paper, the average features of each class in the minimum batch are used to approximate the average features of all samples of different classes [5]. x_i represents the feature of the input fully connected layer. m indicates the minimum batch size. The loss function of the classification network model in this paper is defined as:

$$L = L_s + \lambda L_c. \tag{12}$$

Where the λ value is 0.5.

3 Experimental Results and Analysis

Experiments are needed to prove the effectiveness of data privacy encryption protection method in big data environment based on deep residual network. The main function of the experiment is to test the security and efficiency of the improved data privacy encryption protection algorithm under different big data environments. The simulation virtual environment of data privacy encryption protection experiment under big data environment is built on Windows11 operating system. The server used in the experiment is business Intelligence database server. The experimental development languages and tools are C# and Visual Studio professional 2016.

TDE [24], EMHT [12], HEM [2] and the proposed algorithms are used to carry out data privacy encryption protection experiments in big data environment. The data encryption time of the four algorithms is compared, and the comparison results are shown in Table 1 and Figure 3. Table 2 and Figure 4 are the comparison of data retrieval efficiency.

Table 1: Comparison of data encryption efficiency with different algorithms

Method	15	20	25	35	45
TDE	0.45	0.43	0.41	0.39	0.39
EMHT	0.31	0.29	0.28	0.27	0.27
HEM	0.25	0.23	0.22	0.22	0.22
Proposed	0.09	0.08	0.07	0.06	0.06

Table 2: Comparison of data retrieval efficiency with different algorithms

Method	15	20	25	35	45
TDE	0.37	0.35	0.31	0.31	0.31
EMHT	0.28	0.25	0.23	0.21	0.21
HEM	0.23	0.21	0.17	0.15	0.15
Proposed	0.09	0.08	0.06	0.06	0.05

As can be seen from Figure 3 and Figure 4 the efficiency of the improved algorithm in data privacy encryption protection experiment in big data environment is obviously better than that of the traditional algorithm,

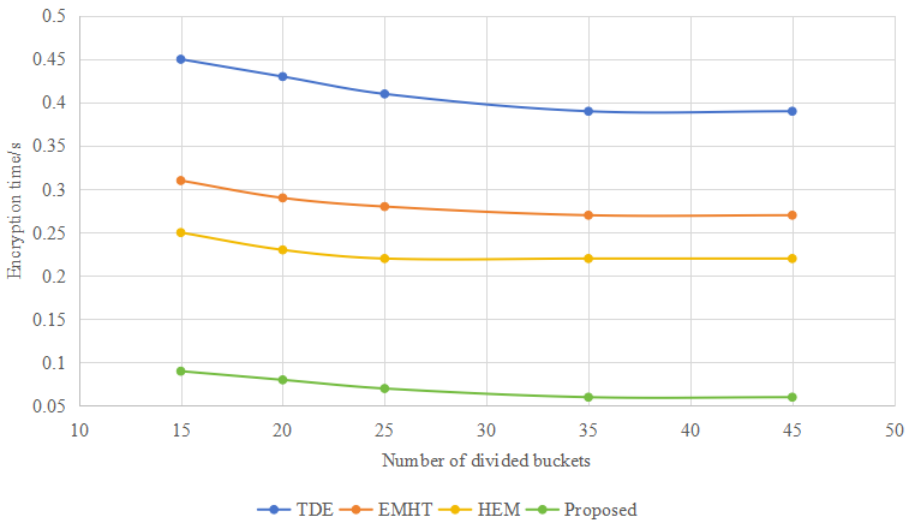


Figure 3: Comparison of data encryption efficiency with different algorithms

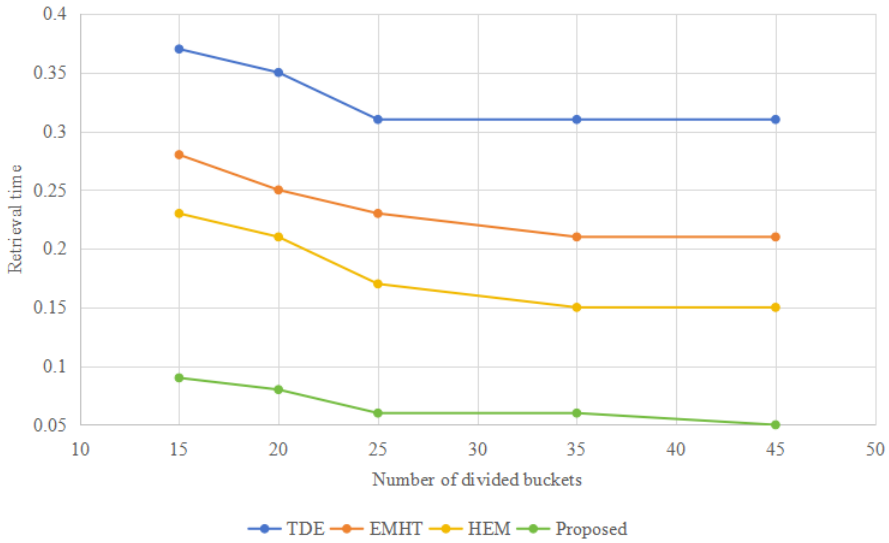


Figure 4: Comparison of data retrieval efficiency with different algorithms

mainly because the improved algorithm first converts the private data in big data environment into numerical values and expresses them according to the theoretical idea of OPES+ [1], and divides the arranged location data into buckets to ensure a balanced distribution. Make the number of points in each bucket less than the given threshold, use the protection order encryption algorithm to encrypt the data in the bucket, and keep the size order of the encrypted values unchanged, and restrict the data privacy encryption process to the solution of the encryption function, on this basis, the obtained encryption function is the core. The problem of data privacy encryption protection in big data environment is transformed into a linear optimization problem of constrained isomorphism, and the security of the problem is optimized by combining simplex algorithm, thus ensuring the high efficiency of the improved algorithm.

The results of security reliability, precision and error rate obtained by using the proposed encryption method are shown in Table 3.

As can be seen from Table 3, the accuracy rate, security reliability and error rate of the proposed algorithm in the data privacy encryption protection experiment in the big data environment are relatively good, which verifies that the improved algorithm can effectively prevent data privacy disclosure in the big data environment, and has good practical value. The simulation results show that the data privacy encryption protection method based on improved residual network encryption algorithm in big data environment improves the security of data privacy encryption protection and has strong robustness.

4 Conclusions

This paper studies the issues of track and field data encryption and privacy protection in hybrid cloud environment, and proposes a privacy protection model based on residual network and elliptic curve encryption. The arithmetic coding model of privacy protection in hybrid cloud environment is constructed under elliptic curve cryptosystem, the public key encryption is searched through the optimization of elliptic curve equation, the master key of privacy protection data encryption is constructed, the threshold cryptosystem is established in combination with residual network scheme, the private key sharing scheme is constructed, and the user private key is recovered by bilinear cyclic shift method. Privacy protection. The results show that the data encryption method adopted in this paper has strong anti-attack capability, high reliability and better performance.

Acknowledgments

This work was supported by Harbin Sport University Doctoral Talent Research Project (RCYJ-2115).

References

[1] C. Ahn, J. Hong, M. H. Kim, "Fermionic construction in the supersymmetric coset model," *International Journal of Modern Physics A*, vol. 37, no. 04, pp. 2250007, 2022.

[2] S. Das, S. Namasudra, "A novel hybrid encryption method to secure healthcare data in IoT-enabled healthcare infrastructure," *Computers and Electrical Engineering*, vol. 101, pp. 107991, 2022.

[3] V. Devmane, B. K. Lande, J. Joglekar, D. Hiran, "Preserving data security in cloud environment using an adaptive homomorphic blockchain technique," *Arabian Journal for Science and Engineering*, vol. 47, no. 8, pp. 10381-10394, 2022.

[4] R. Durga, E. Poovammal, K. Ramana, R. H. Jhaveri, S. Singh and B. Yoon, "CES Blocks: A Novel Chaotic Encryption Schemes-Based Blockchain System for an IoT Environment," *IEEE Access*, vol. 10, pp. 11354-11371, 2022.

[5] X. Fang, K. Jiang, N. Han, S. Teng, G. Zhou and S. Xie, "Average Approximate Hashing-Based Double Projections Learning for Cross-Modal Retrieval," *IEEE Transactions on Cybernetics*, vol. 52, no. 11, pp. 11780-11793, 2022.

[6] M. Q. Hong, P. -Y. Wang and W. -B. Zhao, "Homomorphic Encryption Scheme Based on Elliptic Curve Cryptography for Privacy Protection of Cloud Computing," in *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), New York, NY, USA*, pp. 152-157, 2016, doi: 10.1109/BigDataSecurity-HPSC-IDS.2016.51.

[7] L. C. Huang, M. S. Hwang, "Two-party authenticated multiple-key agreement based on elliptic curve discrete logarithm problem", *International Journal of Smart Home*, vol. 7, no. 1, pp. 9-18, 2013.

[8] M. S. Hwang, E. F. Cahyadi, C. Y. Yang, S. F. Chiou, "An improvement of the remote authentication scheme for anonymous users using an elliptic curve cryptosystem", in *IEEE 4th International Conference on Computer and Communications (ICCC'18)*, pp. 1872-1877, 2018.

[9] M. S. Hwang, C. C. Lee, J. Z. Lee, C. C. Yang, "A secure protocol for bluetooth piconets using elliptic curve cryptography", *Telecommunication Systems*, vol. 29, no. 3, pp. 165-180, 2005.

[10] M. S. Hwang, S. F. Tzeng, C. S. Tsai, "Generalization of proxy signature based on elliptic curves," *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 73-84, 2004.

[11] M. S. Hwang and W. P. Yang, "Conference key distribution protocols for digital mobile communication systems", *IEEE Journal on Selected Areas in Communications*, vo1. 13, no. 2, pp. 416-420, Feb. 1995.

[12] J. S. Jayaprakash, K. Balasubramanian, R. Sulaiman, M. K. Hasan, B. D. Parameshachari, C.

Table 3: Overall performance with proposed method

Number of divided buckets	15	20	25	35	45
Security reliability/%	96	97	96	97	97
Precision/%	97	97	95	97	95
Error rate/%	0.001	0.001	0.001	0.001	0.001

Iwendi, "Cloud Data Encryption and Authentication Based on Enhanced Merkle Hash Tree Method," *Computers, Materials & Continua*, vol. 72, no. 1, 2022.

[13] R. Kumar, D. Singh, A. Chug and A. P. Singh, "Evaluation of Deep learning based Resnet-50 for Plant Disease Classification with Stability Analysis," in *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India*, pp. 1280-1287, 2022, doi: 10.1109/ICICCS53718.2022.9788207.

[14] B Li, D Lima, "Facial expression recognition via ResNet-50," *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 57-64, 2021.

[15] J. Liu, L. Xu, Y. Xie, T. Ma, J. Wang, Z. Tang, W. Gui, "Toward Robust Fault Identification of Complex Industrial Processes Using Stacked Sparse-Denoising Autoencoder With Softmax Classifier," *IEEE Transactions on Cybernetics*, vol. 53, no. 1, pp. 428-442, 2023.

[16] C. W. Liu, W. F. Hsien, C. C. Yang, M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900-916, 2016.

[17] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing", *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[18] J. W. Lo, C. C. Lee, M. S. Hwang, Y. P. Chu, "A secure and efficient ECC-based AKA protocol for wireless mobile communications", *International Journal of Innovative Computing, Information and Control*, vol. 6, no. 11, pp. 5249-5258, 2010.

[19] X. Lv, M. Li, "Application and research of the intelligent management system based on internet of things technology in the era of big data," *Mobile Information Systems*, vol. 2021, pp. 1-6, 2021.

[20] R. Ma and L. Du, "Attribute-Based Blind Signature Scheme Based on Elliptic Curve Cryptography," *IEEE Access*, vol. 10, pp. 34221-34227, 2022.

[21] M. A. Mehrabi, C. Doche and A. Jolfaei, "Elliptic Curve Cryptography Point Multiplication Core for Hardware Security Module," *IEEE Transactions on Computers*, vol. 69, no. 11, pp. 1707-1718, 2020.

[22] U. Narayanan, V. Paul, S. Joseph, "A novel system architecture for secure authentication and data sharing in cloud enabled Big Data Environment," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 6, pp. 3121-3135, 2022.

[23] M. K. Panda, A. Sharma, V. Bajpai, B. N. Subudhi, V. Thangaraj, V. Jakhetiya, "Encoder and decoder network with ResNet-50 and global average feature pooling for local change detection," *Computer Vision and Image Understanding*, vol. 222, pp. 103501, 2022.

[24] M. N. Ramachandra, M. S. Rao, W. Lai, B. D. Parameshachari, J. A. Babu, K. L. Hemalatha, "An efficient and secure big data storage in cloud environment by using triple data encryption standard," *Big Data and Cognitive Computing*, vol. 6, no. 4, pp. 101, 2022.

[25] K. B. Sarmila and S. V. Manisekaran, "A Study on Security Considerations in IoT Environment and Data Protection Methodologies for Communication in Cloud Computing," in *2019 International Carnahan Conference on Security Technology (ICCST), Chennai, India*, pp. 1-6, 2019, doi: 10.1109/CCST.2019.8888414.

[26] K. Shahid, G. Tong, J. Li, A. Qadir, U. Farooq, and Y. Yu, "Current Advances and Future Perspectives of Image Fusion: A Comprehensive Review," *Information Fusion*, vol. 90, pp. 185-217, 2023.

[27] Z. A. Shaikh, A. A. Khan, L. Teng, A. A. Wagan, A. A. Laghari, "BIoMT Modular Infrastructure: The Recent Challenges, Issues, and Limitations in Blockchain Hyperledger-Enabled E-Healthcare Application," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1-14, 2022.

[28] P. K. Tysowski and M. A. Hasan, "Hybrid Attribute- and Re-Encryption-Based Key Management for Secure and Scalable Mobile Applications in Clouds," *IEEE Transactions on Cloud Computing*, vol. 1, no. 2, pp. 172-186, 2013.

[29] S. F. Tzeng, M. S. Hwang, "Digital signature with message recovery and its variants based on elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 61-71, Mar. 2004.

[30] S. Wang, "Artificial intelligence applications in the new model of logistics development based on wireless communication technology," *Scientific Programming*, vol. 2021, pp. 1-5, 2021.

[31] X. Wang, S. Yin, M. Shafiq, A. A. Laghari, S. Karim, O. Cheikhrouhou, W. Alhakami, H. Hamam, "A New V-Net Convolutional Neural Network Based on Four-Dimensional Hyperchaotic System for Medical Im-

age Encryption," *Security and Communication Networks*, vol. 2022, pp. 1-14,, 2022.

[32] P. K. Yadav, T. Burks, Q. Frederick, J. Qin, M. Kim, M. A. Ritenour, "Citrus disease detection using convolution neural network generated features and Softmax classifier on hyperspectral image data," *Frontiers in Plant Science*, vol. 13, pp. 1043712, 2022.

[33] C. C. Yang, T. Y. Chang, M. S. Hwang, "A new anonymous conference key distribution system based on the elliptic curve discrete logarithm problem," *Computer Standards & Interfaces*, vol. 25, no. 2, pp. 141-145, 2003.

[34] M. Yang, I. Tjuawinata, K. -Y. Lam, T. Zhu and J. Zhao, "Differentially Private Distributed Frequency Estimation," *IEEE Transactions on Dependable and Secure Computing*, 2022. doi: 10.1109/TDSC.2022.3227654.

[35] L. -Y. Yeh, P. -J. Chen, C. -C. Pai and T. -T. Liu, "An Energy-Efficient Dual-Field Elliptic Curve Cryptography Processor for Internet of Things Applications," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 9, pp. 1614-1618, 2020.

[36] S. Yin, H. Li, S. Karim, and Y. Sun, "ECID: Elliptic Curve Identity-based Blind Signature Scheme," *International Journal of Network Security*, vol. 23, no. 1, pp. 9-13, 2021.

Biography

Yiliang Chen biography. Yiliang Chen, currently a lecturer at Harbin Sport University, holds a Ph.D. in Physical Education (currently studying), mainly focusing on physical education teaching, physical function training, and sports training. Published multiple research papers in academic journals in the aforementioned research fields, participated in the development of multiple textbooks and works, and obtained multiple invention patents.

Jie Ren biography. Jie Ren received her M.S. degree from Harbin Engineering University. She is a lecturer with Harbin Sport University. She has been engaged in the measurement and evaluation of physical activity, the measurement method of sports efficiency, smart Sports.

Chunlin Luo biography. Chunlin Luo, currently working at Harbin Sport University, is an associate professor with a PhD in physical education. His main research areas and directions include physical education teaching, mass fitness, and competitive sports. I have published multiple papers domestically and internationally, led and participated in multiple fund projects, and have profound achievements in the above fields.

A Study on Detection of Malware Attacks Using Machine Learning Techniques

Daojing Yang
(Corresponding author: Daojing Yang)
Department of Economic Statistics, The College of Accounting and Finance, Chengdu Jincheng College
No. 1 Xiyuan Avenue, Chengdu, Sichuan 611731, China
Email: oudao16986636848@yeah.net
(Received Sept. 4, 2022; Revised and Accepted July 28, 2023; First Online Oct. 11, 2023)

Abstract

Accurate identification of malware attacks can enhance network security. This paper used the XGBoost algorithm in machine learning to identify and detect malicious software attacks. To improve recognition performance, the term frequency-inverse document frequency (TF-IDF)-chi was used to filter features after using the N-gram method to extract sample feature sets. Finally, the performance of the XGBoost algorithm under three feature extraction methods (N-gram, TF-IDF, and TF-IDF-chi) was tested in simulation experiments and compared with support vector machine and back-propagation neural network algorithms. The results showed that the XGBoost algorithm with TF-IDF-chi feature extraction had the best performance; whether it was binary or multi-classification of malware attacks, the XGBoost algorithm always exhibited the best performance while taking less time for recognition.

Keywords: Machine Learning; Malware; Recognition; TF-IDF-Chi

1 Introduction

With the rapid development of computer technology, network security issues are becoming increasingly prominent. Malware attacks have become one of the main threats to cybersecurity in cyberspace today [8, 9]. The types of malware attacks often go unnoticed and pose a serious threat to individuals and businesses' privacy information and network security [10, 14]. In order to address such threats, detecting malware attacks has become an important task [18, 21].

Malware, also known as malicious software or harmful code, refers to programs intentionally inserted into a system or application with the aim of disrupting, interfering with, or controlling computer systems or networks. It can be divided into viruses, worms, trojans and phishing software [13, 20]. These malicious programs often use social engineering techniques, system vulnerabilities and forced installations to spread their attacks. There are various methods for detecting malware attacks such as regular vulnerability scanning and log analysis [3].

Among these detection methods, machine learning-based algorithms are an efficient means. This method uses machine learning technology to mine patterns in sample data features and construct classifiers that can automatically detect malware. The relevant studies are as follows. Huda *et al.* [11] proposed a signature-free malware detection method and a feature selection method based on hybrid package filters and validated them using statistical model selection criteria.

Eskandari *et al.* [7] obtained semantic signatures of malware by mining frequent subgraphs in a set of control flow graphs and used them to detect malware. The experimental results showed an improvement in the F-value of this method compared to classical graph-based methods. Lei *et al.* [4] introduced a regularized logistic regression model with probability discrimination for detecting Android malware. The experimental results showed that the model could produce probabilistic outputs with highly precise classification outcomes.

The XGBoost algorithm, a machine learning algorithm, was used in this article to detect and identify malicious software attacks. In order to improve the recognition performance, the features extracted from the samples using N-gram were further filtered using term frequency-inverse document frequency (TF-IDF)-chi. Finally, the performance of the XGBoost algorithms under three feature extraction methods (N-gram, TF-IDF, and TFIDF-chi) was tested in simulation experiments, and it was compared with support vector machine (SVM) and back-propagation neural network (BPNN) algorithms.

2 A Malware Attack Detection Algorithm Based on Machine Learning

There are various methods for detecting malware attacks, and the traditional approach is to conduct regular inspections to detect anomalies in software and software behavior data [15]. Regular inspections are relatively inflexible, and if they are done manually, efficiency will be lower. The emergence and development of machine learning algorithms have provided a flexible and efficient way to detect malicious software. In general, the basic principle of machine learning algorithms for detecting malicious software attacks is to extract features from sample data of such attacks and then use a classifier constructed from the features and classification labels of training samples to classify malware or attack behavior, thereby achieving detection of malware attacks [16].

Machine learning algorithms that can be used for malware attack detection include SVM [5], naive Bayes, decision trees, and others. Among them, the SVM algorithm maps the feature data of sample data to a high-dimensional space and then fits a "hyperplane" that can separate the space to classify the sample data in the space. The Naive Bayes algorithm uses prior conditional probabilities to calculate the probability of sample categories, and its training process calculates prior conditional probabilities of different features under different categories using a training set. The decision tree algorithm treats each feature of the sample data as a node [17], and then continuously splits the data based on these nodes until it cannot be further divided. Finally, in this paper, the XGBoost algorithm is chosen to detect malware attacks. The XGBoost algorithm is an ensemble learning model that combines multiple low-performance classification models. The detection process for malware attacks using the XGBoost algorithm is shown in Figure 1.

1) Malware and its behavior data are collected. The code of the malware and the log information of its dynamic behavior are structurally similar to text sequences, so they can be regarded as a special type of "text". The types of the "text" is identified using a classification algorithm.

2) The collected data are preprocessed, including word segmentation, case conversion [2], lemmatization, removal of non-text symbols, etc.

3) The N-gram algorithm is used to obtain the feature vector of the 'text'. Its basic step is as follows. A character window of length N slides with a step size of 1 on the sequence of the 'text'. This divides the 'text' into a sequence combination of N-character length fragments, where the term frequency of each fragment represents the N-gram feature vector for that particular 'text'. By combining multiple texts, we can create an N-gram feature vector matrix for the collection of texts [19].

4) Although the XGBoost algorithm can handle sparse high-dimensional data well and directly calculate N-gram feature vectors, the high dimensionality of sparse data still affects computation time and accuracy. Therefore, this paper uses TF-IDF [12] to extract key features from N-gram feature vectors. Considering the distribution differences of feature words between normal and malware behavior data, chi-square values are introduced into TFIDF. The corresponding calculation formula is:

$$\begin{cases} TFIDF_{chi_{i,j}} = FIDF_{i,j} \cdot chi_{i,t} \\ TFIDF_{i,j} = TF_{i,j} \cdot IDF_i \\ chi_{i,t} = \frac{(W+X+Y+Z)(WZ-XY)^2}{(W+Y)(X+Z)(W+X)(Y+Z)} \\ TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{i,k}} \\ IDF_i = \log \frac{|D|}{|d_j : v_i \in d_j|+1} \end{cases} \tag{1}$$

where $TFIDF_{chi_{i,j}}$ denotes the feature value of feature word i in text j, whose size can reflect the importance of the feature word, $TF_{i,j}$ is the word frequency of feature word i, IDF_i denotes the inverse text frequency of feature word i, $n_{i,j}$ is the occurrence frequency of feature word i in text j, $|D|$ denotes the total number of texts, $|d_j : v_i \in d_j|$ represents the total number of texts containing feature word i, $chi_{i,t}$ represents the chi-square value between feature word i and text category t, W is the number of texts containing feature word i in category t, X represents the number of texts containing feature word i in non-category t, Y represents the number of texts not containing feature i in category t, and Z is the number of texts not containing feature word i in non-category t.

5) The extracted text feature, $TFIDF_{chi_{i,j}}$, is input into the XGBoost algorithm for category prediction to obtain classification results. The corresponding formula is:

$$\hat{y_i} = \sum_{k=1}^{K} f_k(x_i), f_k \in F, \tag{2}$$

where $\hat{y_i}$ is the predicted category obtained through calculation, f_k is the base leaner, F is the set of base learners [1], totally K base learners, and x_i is the extracted text feature. The objective function used in the training of the XGBoost algorithm [6] is:

$$loss^{(t)} = \sum_{i=1}^{n} l(y_i, (\hat{y_i}^{t-1} + f_t(x_i)) + \Omega(f_t), \tag{3}$$

where $loss^t$ is the current predicted loss, y_i is the actual category corresponding to x_i, $\hat{y_i}^{t-1}$ is the category of x_i predicted by the first $(t-1)$ integrated

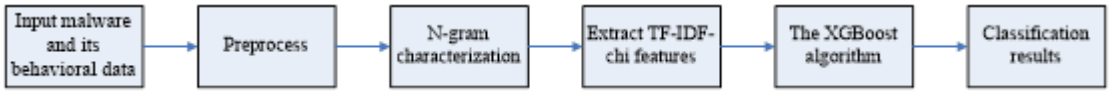


Figure 1: The improved XGBoost-based malware attack detection flow

learners, $f_t(x_i)$ is the category of x_i predicted by the current learner, and $\Omega(f_t)$ is the regular term of the current learner.

3 Simulation Experiment

3.1 Experimental Data

The simulation experiments conducted in this paper were carried out on a server in the laboratory, with Windows 10 operating system, 16 GB memory and i7 processor as the relevant configuration. The malware sample data required for the simulation experiment comes from the public dataset SoReL-20M of malware samples jointly released by network security companies Sophos and ReversingLabs. This dataset contains 20 million metadata of Windows portable executable (PE) files, with corresponding labels. There are twelve types of malware in the dataset, including Is_malware, Adware, Flooder, Ransomware, Dropper, Spyware, Packed, Crypto_miner, File_infector, Installer, Worm, and Downloader.

3.2 Experimental Setup

Before conducting the simulation experiment, the malware and normal software in the dataset were first tested in a separate intelligent sandbox, while recording corresponding running log data. Then, N-gram features of software running data were obtained using orthogonal testing method, with N set to 2 as determined by the orthogonal test.

The relevant parameters for the XGBoost algorithm were as follows: using a tree structure model as base learner, setting the maximum iteration times to 300, setting the maximum tree depth to 5, setting the learning rate to 0.05, and setting the minimum sample size of the leaf to 5.

In order to verify the performance improvement of the XGBoost algorithm, it was compared with the SVM and BPNN algorithms. The feature parameters required for malware attack detection using SVM and decision trees were consistent with the improved XGBoost algorithm. The relevant parameters for the SVM algorithm were: a sigmoid kernel function and a penalty factor set to 1. The relevant parameters for the BPNN algorithm were: two hidden layers with 64 nodes in each layer, and the Relu function used as the activation function.

Table 1: Confusion matrix

	Predicted as Positive Case	Predicted as Negative Case
True positive case	TP	FN
True negative case	FP	TN

3.3 Evaluation Criteria

The detection algorithm for malware attacks is a classification algorithm, so the performance can be evaluated using a confusion matrix (See Table 1). If only judging whether the detection algorithm can identify malware attacks, binary evaluation criteria should be used, with the formula as follows:

$$\begin{cases} P = \frac{TP}{TP+FP} \\ R = \frac{TP}{TP+FN} \\ F = \frac{2 \cdot P \cdot R}{P+R} \end{cases} \tag{4}$$

where P stands for the precision, R stands for the recall rate, and F stands for the harmonic mean of the precision and recall rate.

The dataset in this article contains 12 types of malware, so macro indicators are needed for multi-class evaluation. In the multi-class evaluation indicators, each type of malware is treated as a binary classification evaluation. The precision of identifying this type of malware can be calculated using Equation (4). The average precision of all types is the macro precision of the detection algorithm. Similarly, the macro recall rate can be calculated in the same way, and the macro F-value can be obtained by combining macro precision and macro recall rate.

3.4 Experimental Results

Figure 2 shows the detection performance of the XGBoost algorithm when selecting N-gram, TF-IDF, and TF-IDF-chi as feature extraction methods for malware attacks. It can be seen from Figure 2 that using TF-IDF-chi as the feature extraction method yielded the best performance for the XGBoost algorithm, followed by using TF-IDF, while using N-gram as the feature extraction method had the worst detection performance. This result indicated that utilizing the TF-IDF weight of key features could effectively enhance the detection performance of the algorithm, while introducing the chi-square value between

feature words and malware in TF-IDF-chi could further reflect correlation between feature words and categories to improve the detection performance of the algorithm.

Figure 3 displays the binary classification performance of three recognition algorithms for detecting malware attacks, i.e., the accuracy of identifying whether a software behavior is a malware attack. It can be seen from Figure 3 that the improved XGBoost algorithm had the highest recognition performance, followed by the BPNN recognition algorithm, while the SVM algorithm performed worst.

The performance and identification time of three recognition algorithms for the multi-classification of malware are shown in Table 2. It can be observed from Table 2 that, for different types of malware attacks, the improved XGBoost exhibits the highest multi-classification recognition performance, followed by the BPNN algorithm, while the SVM algorithm performed worst. In terms of identification time, the SVM algorithm took the longest time, followed by the BPNN algorithm, and the improved XGBoost algorithm required the least amount of time. Additionally, comparing the multi-classification performance and binary classification performance of these three recognition algorithms revealed a minimal difference in the improved XGBoost algorithm, a slightly larger difference in the BPNN algorithm, and a maximum difference in the SVM algorithm.

4 Conclusion

The XGBoost algorithm, a machine learning algorithm, was used in this article to detect and identify malware attacks. In order to improve the identification performance, the sample features were extracted using N-gram method, and then TF-IDF-chi was used to filter the features. Finally, simulation experiments were conducted to test the performance of the XGBoost algorithm under three feature extraction methods: N-gram, TF-IDF, and TF-IDF-chi. It was also compared with two other algorithms: the SVM and BPNN algorithms. The results are as follows.

1) When using TF-IDF-chi as the feature extraction method, the XGBoost algorithm performed the best, followed by TF-IDF. N-gram exhibited the weakest results.

2) In terms of binary classification for malware attacks, the improved XGBoost algorithm demonstrated superior recognition performance, compared to the BPNN and SVM algorithms.

3) In terms of the multi-classification performance and recognition efficiency for malware attacks, the improved XGBoost algorithm remained the top performer, followed by the BPNN algorithm, and the SVM algorithm exhibited the worst performance. Moreover, when comparing the multi-classification recognition performance to that of binary classification, the improved XGBoost algorithm showed little change while the BPNN algorithm decreased slightly and the SVM algorithm significantly decreased.

References

[1] R. Ahmad, "E-learning automated essay scoring system menggunakan metode searching text similarity matching text," *Journal Penelitian Enjiniring*, vol. 22, no. 1, pp. 38-43, 2019.

[2] M. A. Albahar, "A modified maximal divergence sequential auto-encoder and time delay neural network models for vulnerable binary codes detection," *IEEE Access*, vol. 8, pp. 14999-15006, 2020.

[3] D. Canali, A. Lanzi, D. Balzarotti, C. Kruegel, M. Christodorescu, E. Kirda, "A quantitative study of accuracy in system call-based malware detection," in *Proceedings of the International Symposium on Software Testing & Analysis*, pp. 122-132, 2015.

[4] L. Cen, C. S. Gates, L. Si, N. Li, "A probabilistic discriminative model for android malware detection with decompiled source code," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 400-412, 2015.

[5] S. Chen, M. Xue, L. Fan, S. Hao, L. Xu, H. Zhu, B. Li, "Automated poisoning attacks and defenses in malware detection systems: An adversarial machine learning approach," *Computers & Security*, vol. 73, pp. 326-344, 2018.

[6] J. Contreras, S. Hilles, Z. B. Abubakar, "Automated essay scoring using ontology generator and natural language processing with question generator based on blooms taxonomy's cognitive level," *International Journal of Advanced Technology and Engineering Exploration*, vol. 9, no. 1, pp. 2249-8958, 2019.

[7] M. Eskandari, H. Raesi, "Frequent sub-graph mining for intelligent malware detection," *Security and Communication Networks*, vol. 7, no. 11, pp. 1872-1886, 2015.

[8] A. Feizollah, N. B. Anuar, R. Salleh, A. W. A. Wahab, "A review on feature selection in mobile malware detection," *Digital Investigation*, vol. 13, pp. 22-37, 2015.

[9] L. C. Huang, C. H. Chang, M. S. Hwang, "Research on malware detection and classification based on artificial intelligence", *International Journal of Network Security*, vol. 22, no. 5, pp. 717-727, 2020.

[10] L. C. Huang, M. S. Hwang, "Study of intrusion detection systems", *Journal of Electronic Science and Technology*, vol. 10, no. 3, pp. 269-275, 2012.

[11] S. Huda, J. Abawajy, M. Alazab, M. Abdollalihian, R. Islam, J. Yearwood, "Hybrids of support vector machine wrapper and filter based framework for malware detection," *Future Generation Computer Systems*, vol. 55, pp. 376-390, 2016.

[12] Z. Li, D. Zou, J. Tang, Z. Zhang, M. Sun, H. Jin, "A comparative study of deep learning-based vulnerability detection system," *IEEE Access*, vol. 7, pp. 103184-103197, 2019.

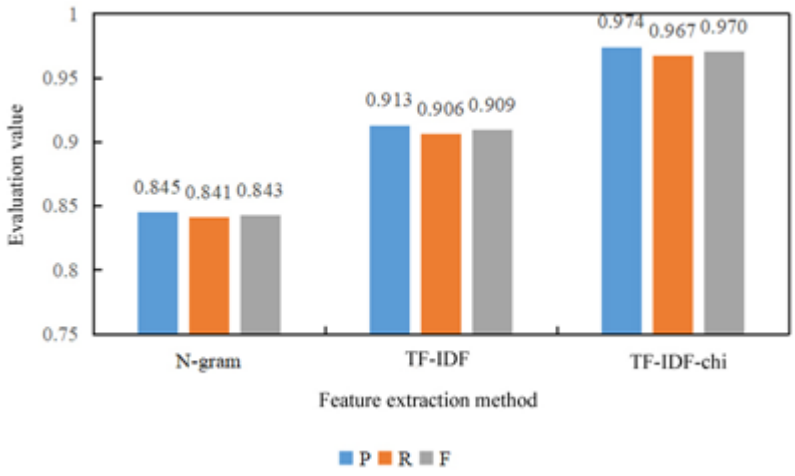


Figure 2: The detection performance of the XGBoost algorithm under different feature extraction methods

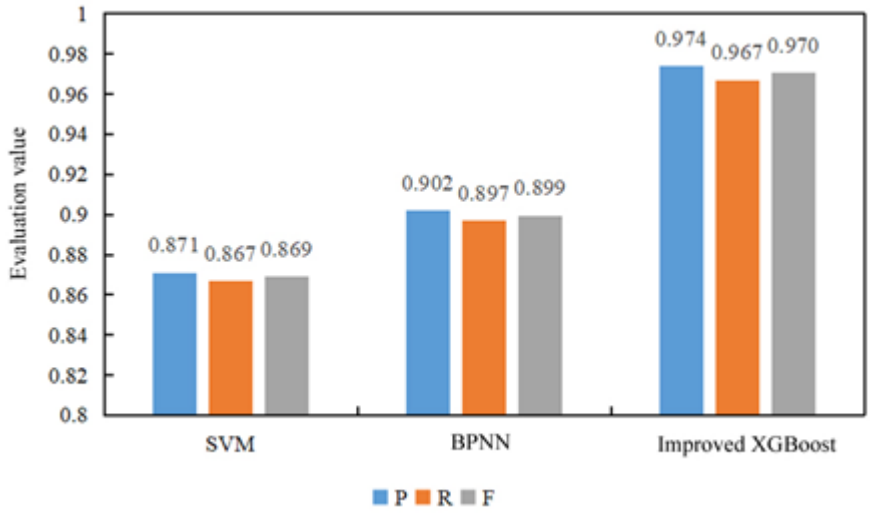


Figure 3: The binary classification performance of three recognition algorithms

Table 2: The multi-classification performance and recognition time of three recognition algorithms

	The SVM algorithm	The BPNN algorithm	The improved XGBoost algorithm
Macro precision	0.611	0.852	0.965
Macro recall rate	0.601	0.841	0.958
Macro F value	0.601	0.841	0.961
Time consumption/s	331	212	88

[13] I. C. Lin, Y. L. Chi, H. C. Chuang, M. S. Hwang, "The novel features for phishing based on user device detection", *Journal of Computers*, vol. 11, no. 2, pp. 109-115, 2016.

[14] C. H. Ling, W. F. Hsien, M. S. Hwang, "A double circular chain intrusion detection for cloud computing based on adjointVM approach", *International Journal of Network Security*, vol. 18, no. 2, pp. 397-400, 2016.

[15] N. Monire, S. Alireza, S. Z. Majid, "A data mining classification approach for behavioral malware detection," *Journal of Computer Networks and Communications*, vol. 2016, pp. 1-9, 2016.

[16] M. Narouei, M. Ahmadi, G. Giacinto, H. Takabi, A. Sami, "DLLMiner: Structural mining for malware detection," *Security & Communication Networks*, vol. 8, no. 18, pp. 3311-3322, 2016.

[17] P. V. Shijo, A. Salim, "Integrated static and dynamic analysis for malware detection," *Procedia Computer Science*, vol. 46, pp. 804-811, 2015.

[18] J. R. Sun, C. T. Huang, M. S. Hwang, "A SYN flooding attack detection approach with hierarchical policies based on self-information", *ETRI Journal*, vol. 44, no. 2, pp. 346-354, 2022.

[19] G. Tang, L. Yang, S. Ren, L. Meng, F. Yang, H. Wang, "An automatic source code vulnerability detection approach based on KELM," *Security and Communication Networks*, vol. 2021, no. 1, pp. 1-12, 2021.

[20] H. W. Yang, L. C. Huang, M. S. Hwang, "Research on detection and prevention of mobile device botnet in cloud service systems", *International Journal on Network Security*, vol. 23, no. 3, pp. 371-378, 2021.

[21] Y. Ye, T. Li, D. Adjeroh, S. S. Iyengar, "A survey on malware detection using data mining techniques," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1-40, 2017.

Biography

Yang Daojing, born in 1982, graduated from Guilin university of electronic technology in July 2006 with a Master's degree. She is an associate professor and is now working in Chengdu Jincheng College, China.She is interested in applications of data mining.

Query-based Local Black-box Adversarial Attacks

Jing Shi[1], Xiaolin Zhang[1], Enhui Xu[2], Yongping Wang[1], and Wenwen Zhang[1]
(Corresponding author: Xiaolin Zhang)
School of Information Engineering, Inner Mongolia University of Science and Technology[1]
Baotou 014010, China
China Nanhu Academy of Electronics and Information Technology, China[2]
Email: zhangxl6161@163.com
(Received Feb. 13, 2023; Revised and Accepted Sept. 24, 2023; First Online Oct. 11, 2023)

Abstract

Most of the current adversarial attacks perturb the entire area of the image, which will not only have a significant impact on the visual quality by perturbing the smooth background but also increase the complexity of the interaction between the attacker and the target model. To address the problems of high query cost and easy distinction between the adversarial samples and clean images, our method (LBQA) utilizes an attention mechanism to generate local perturbation and improve the image quality of the adversarial samples, and use the initial adversarial sample as a new starting point for local black-box attack. In addition, we also propose an improved attack method (E-LBQA) using the concatenation of significant regions of multiple models. Extensive experiments demonstrated that our methods could maintain high success rates of the attack while guaranteeing the image's visual quality. The experimental results also show that our attacks perform well in terms of image quality and query efficiency.

Keywords: Adversarial Example; Black-box Attack; Deep Neural Networks; Local Perturbation

1 Introduction

In recent years, deep neural networks (DNNs) have made unprecedented breakthroughs in areas of artificial intelligence such as image classification [30], speech recognition [32], face recognition [27], sentiment analysis [2], semantic segmentation [13] and medical data processing [16]. However, studies have demonstrated that attackers can produce adversarial samples by introducing some subtle, imperceptible perturbations to clean images in order to misclassify the target model. More crucially, adversarial samples could endanger our lives when used in safety-related tasks or circumstances(e.g., autonomous driving) because they are difficult to identify from clean images. On the other side, adversarial attacks also help researchers to design DNN models that are more robust and perform better. Therefore, research focusing on adversarial samples is receiving increasing attention [17, 20, 29].

Szegedy *et al.* [26] first pointed out the vulnerability of deep neural networks and introduced the definition of adversarial samples. Depending on the attacker's knowledge of the model, adversarial attacks can be divided into white-box and black-box attacks. In black-box attacks, only the input and output of the target model can be read, and all internal data are kept secret [1], which shows that black-box attacks are much more difficult to succeed than white-box attacks, and the attack scenarios are more realistic. The majority of existing methods create adversarial samples by globally perturbing the original image. If all pixel points are continuously modified in the same way, this causes redundant perturbations, which ultimately results in an extensive number of black-box queries and large differences between the original and adversarial images.

Since global attacks have the disadvantage of generating redundant perturbations, local attack methods have also been proposed by numerous scholars to enhance the quality of the adversarial samples. One such method is the Jacobi matrix-based attack (JSMA), which was developed by Papernot *et al.* [24], which uses forward derivatives and adversarial significant mapping to select the features that have the greatest impact on the output target class by searching for the two most effective pixel points in the input and iteratively changing the pixel values to generate the adversarial samples, but the quality of its generated images is not well. After that, a better method was then put forth, known as OJSMA [31], which accurately chooses and updates the input features utilizing an adversarial saliency map that mixes increasing and decreasing features with an objective function that optimizes the L_2 norm of perturbations of selected features. It has improved in terms of success rate and ingestion but sacrificed some time. Square Attack [3] finds the optimal attack position by constructing a square around each pixel point of the input image, which leads to a very high computational cost, especially on high-resolution images.

Inspired by the above problem, this paper proposes a local black-box query adversarial attack (LBQA) based on perceived color distance, which can have a high attack

success rate without being detected. The local black box attack starting with an initial adversarial sample reduces the complexity of the interaction between the attacker and the target model and improves the query efficiency. Specifically, given a clean image as input, a saliency map is first generated using the well-liked model interpretation method which is less intrusive to clean images. The discriminative zone is then perturbed depending on the saliency map. Numerous experiments have demonstrated that this method can reduce redundant perturbations while preserving attack capabilities. Our contributions can be divided into the following points:

- We propose a new effective local adversarial attack (LBQA) for black-box that uses the disturbance based on perceptual color difference to modify local pixels, guaranteeing the visual quality of the adversarial samples.
- Based on LBQA, we further propose E-LBQA which exploits the effectiveness of multi-model significant region merging sets to improve the performance of black-box attacks.
- Compared with classical approaches, we verify that the proposed approaches can improve the success rate of black-box attacks and significantly reduce the number of queries required, which also confirms the effectiveness of the local black-box attack framework.

2 Related Works

Adversarial samples can be crafted by adding invisible perturbations to the original image, and existing adversarial attack methods can be classified into white-box and black-box attacks. The white-box attack can obtain the structure and parameters of the target model, which is further divided into global and local attacks according to the number of attack pixels. Black-box attacks can be classified into query-based attacks and transfer-based attacks. In this section, we will briefly introduce the more popular adversarial attack methods.

2.1 Global Attacks

The current mainstream white-box global adversarial attack methods can be classified into gradient-based and optimization-based attacks, as well as decision-based attacks. For example, the FGSM [14] attack is one of the most representative gradient-based adversarial attacks, whose main idea is to quickly generate adversarial samples by maximizing the gradient of the loss function in one step to misclassify the model. After that, numerous scholars improved it. BIM [18] decomposes the operation of increasing the loss function of the target attack model by one step in FGSM into increasing the loss function by multiple small steps iteratively, MI-FGSM [11] iteratively adds a momentum factor to enhance transferability and stabilize the update direction of the adversarial perturbation. Additionally, PGD [22] generates adversarial samples through additional iterations, projecting the adversarial perturbations into a specified range at each iteration. Gradient-based attacks also include TI-FGSM [12], NI-FGSM [21], VMI-FGSM [28] and others. The CW attack [5], one of the most popular optimization-based white-box attacks, transforms the process of generating adversarial samples into an optimization problem by limiting the L_∞, L_2, and L_0 norm. DeepFool [23] is a decision-based attack that iteratively computes the minimum adversarial perturbation under the constraint of parameters, and gradually pushes the samples located within the decision boundary out of the decision boundary so that the target attack model misclassifies the adversarial samples.

Although the aforementioned global adversarial attacks can generate adversarial samples, they vary equally for all pixel points, which will not only produce large gaps with the original samples under the constraint, but also redundantly perturb smooth background regions that are not important for the target class impact, and more seriously blur the visual features of the target image and destroy the spatial semantic information. For instance, FGSM does not take into account the problem that the generated perturbations can destroy the overall structure of the image, which directly affects the threatening property of the adversarial samples. It follows that it is necessary to limit the range of adversarial perturbations and reduce invalid perturbations.

2.2 Local Attacks

Unlike the all-pixel attack, some researchers have also argued that perturbing a portion of the salient regions is more effective for generating adversarial samples, and JSMA uses Jacobi matrices to achieve a high success rate of the attack by computing the saliency map from the input to the output, modifying only 4.02% of the pixels in the original image. The Jacobi matrix consists of the partial derivatives of the output for the input, and the saliency map for the target adversarial attack is shown below Equation (1).

$$S(X,t)[i] = \begin{cases} 0, \text{ if } \frac{\partial F_t(X)}{\partial X_i} < 0 \text{ or } \sum\limits_{j \neq t} \frac{\partial F_j(X)}{\partial X_i} > 0 \\ (\frac{\partial F_t(X)}{\partial X_i}) \cdot |\sum\limits_{j \neq t} \frac{\partial F_j(X)}{\partial X_i}|, \text{ otherwise} \end{cases} \quad (1)$$

where X_i denotes an input feature, t denotes the target class, and S is the obtained saliency map. However, the drawback is that the Jacobi matrix is more complicated to calculate and resource consumption is high.

In addition, the one-pixel attack [25] is also an optimization-based local perturbation method, which restricts the adversarial perturbation to a small region and perturbs only one or a few pixel points to achieve a good attack effect, and the use of a differential evolution algorithm can further improve the finding efficiency of the

attacked pixel points. Following that, further local attack strategies have been put forth, SGA [10] can use the generation method of superpixels to achieve a local smoothing attack, and Finefool [6]uses a double attention mechanism to generate the adversarial samples.

2.3 Black-box Attacks

In contrast to white-box attacks, which perform an effective search by gradient descent, black-box attacks have limited knowledge of the target model and no information such as hyperparameters or training data to perform the attack. Therefore, ZOO [8] was proposed as the first gradient estimation-based method that uses zero-order optimization to improve the speed of gradient estimation but does not optimize confidence acquisition. Later Bhagoji [4] *et al.* explored two strategies to reduce the confidence cost: random feature grouping and query reduction by principal component analysis (PCA). In addition, Dong [11] *et al.* used alternative models of multiple target models to attack black boxes based on transferability, but the expense of training alternative models is high and the resulting adversarial samples do not have excellent transferability when the attacker does not have sufficient knowledge of the training dataset.

For the above approaches, queries on the target model are essential, but since access to the victim model is usually restricted, queries on it can be costly in terms of time and money. Although recent work manages to reduce the total number of queries from millions to thousands, its query efficiency is still less than optimal. For example, Sign-opt [9] transforms the decision-based attack into a problem of determining the direction of the shortest distance to the decision boundary by using a symbolic function to quickly acquire the estimated direction from the sample to the model. Then HSJA [7] improved it based on the decision boundary using the gradient direction and geometric level of the binary information to update the adversarial sample. SimBA [15] views adversarial perturbations to alter the model's prediction as a discrete optimization problem, repeatedly choosing a random direction from a pre-specified set of orthogonal search directions, determining whether it is pointing toward or away from the decision boundary using the confidence level, and perturbing the image by adding or removing vectors from the image.

Unlike attacking the global region, Li *et al.* [19] argue that the foreground is more predictive than the background, separating the foreground from the background and searching only the foreground region randomly by semantic segmentation model, using the segmentation model will again consume a lot of computational resources. Although the number of queries can be reduced by local perturbation, it still requires a large number of queries to deceive the target model. Therefore, a query-efficient black-box attack method remains an open problem. The attacker is given the additional goal of minimizing the number of black-box queries while successfully constructing imperceptible adversarial perturbations.

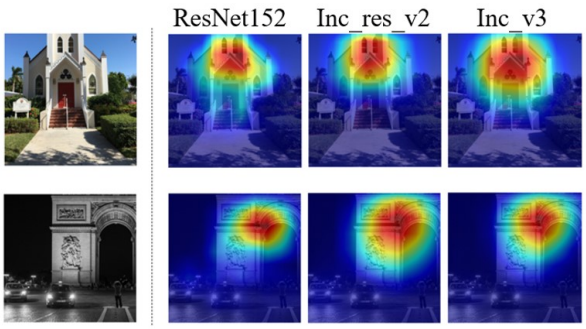


Figure 1: Visualization of significant regions for three models (ResNet_152, Inc_Res_v2, Inc_v3).

3 Methodology

3.1 Problem Definition

Given a classifier f with parameters θ and a clean image x whose true label is y, and a perturbation δ with a size constraint ϵ, attackers expect the image x_{adv} to mislead the target model, and the query-based attack can usually be represented as follows.

$$\begin{aligned} &\min_{\delta} f(x+\delta;\theta) \neq f(x;\theta) \\ &s.t. \quad \|\delta\|_p \leq \epsilon and\, queries \leq Q. \end{aligned} \tag{2}$$

where Q is the maximum number of queries allowed. Our approach is based on two types of transferability, the first of which is the transferability of model interpreters, as depicted in Figure 1. Although different models have different architectures and parameters, there is already a large overlap between models, indicating that there is transferability between different model interpreters, feature representations are common in neural networks, and the adversarial sample can be transferred after identifying the important regions with high impact on the target class of the model. Another feature is that the adversarial samples are transferable and the learned model can be successfully manipulated by the adversarial perturbations generated by attacking different models, the property that facilitates finding local pre-perturbations. Our approach combines transfer-based methods with query-based methods to improve the query efficiency of black-box attacks.

3.2 Attack Strategy

In this paper, we propose a new black-box attack method called Locally black-box Query Adversarial Attack (LBQA) based on Perceived Color Distance. This section describes the proposed attack method in detail.

To improve the query efficiency of the black-box model and reduce unnecessary perturbations, the initial adversarial sample is first generated on the white-box and used as a fresh starting point for the local black-box attack.

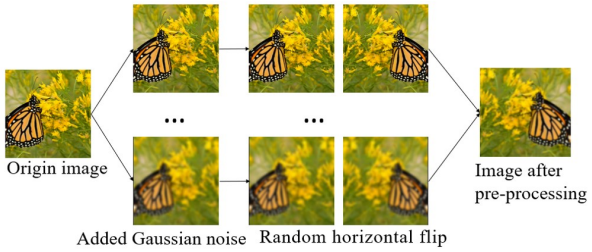


Figure 2: Image pre-processing process.

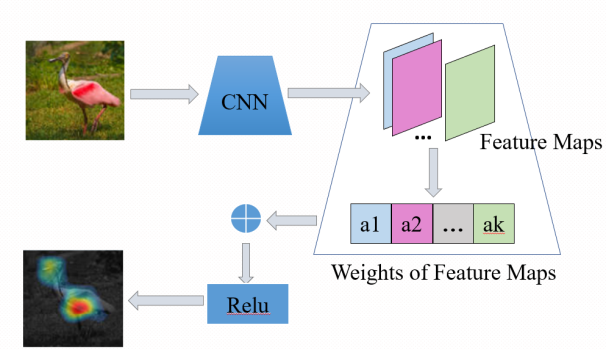


Figure 3: Finding significant areas.

Image enhancement operations are performed on images with Gaussian noise and random horizontal flips to increase image diversity and obtain more useful data. Following data preprocessing, the images are then input to an alternative model, where attention-seeking is used to find the significant region that affects the model's correct classification and binarize it under the premise of ensuring the success rate of the attack. By causing disturbances in this salient region based on the perceived color difference, the initial adversarial sample x' is produced. Given that the pre-perturbed sample is closer to the target model's decision boundary than the unperturbed one, the process of generating the initial adversarial sample can be seen as a preparation process for the black-box attack. The transferability of the adversarial sample can be utilized to locally perturb x' instead of x in a black-box manner. The attention technique dramatically reduces the dimensionality of choosing factors, which makes subsequent attacks easier. The approach suggested in this work may improve the redundancy of current global adversarial attack methods, and Figure 4 depicts the whole procedure.

3.2.1 Image Preprocessing

The pixel transformation approach with Gaussian noise and the geometric transformation method with horizontal flip are utilized to improve the transferability of the adversarial samples. Some of the pixels in the noise-added image may cross the decision boundary because the pixel values of the adversarial samples fall within a certain range. As a result, normalizing allows for the preservation of noisy information while still guaranteeing the original data distribution. This procedure is depicted in Figure 2.

$$x^{mix} = R(\frac{x + noise}{\|x + noise\|_{\infty}}) \quad s.t. \quad noise \sim N(0, \sigma^2) \tag{3}$$

where x^{mix} denotes the image obtained by the random horizontal flip operation with noise added, $R(.)$ denotes the random horizontal flip, and noise denotes the noise that obeys the normal distribution.

3.2.2 Finding Significant Regions

To derive the importance of different features of the input image on the model decision. The corresponding weights are first calculated for each feature mapping in the convolutional layer of the model, then they are multiplied and accumulated with the feature maps to obtain the weighted sum, and the influence of insignificant pixel points is filtered out by applying the function (See Figure 3). Grad-CAM identifies locally salient areas by emphasizing pixels that are strongly associated with the class. On the one hand, since these pixels can more accurately reflect spatial semantic information, attacking the filtered salient maps may be more successful than attacking the entire image. The complexity of decision variables may be decreased in high-resolution images by filtrating pixels, which is advantageous for black-box optimization. It is later transformed into a binary map with the help of the binarization factor k, and the binarization can be expressed as Equation (4) with u_{mn} as the pixel value at its location.

$$B_{mn} = \begin{Bmatrix} 0, & \mu_{mn} < k \\ 1, & \mu_{mn} \geq k \end{Bmatrix} \tag{4}$$

3.2.3 Local Perturbation Based On Perceived Color Distance

In order to ensure that an initial adversarial sample with a high success rate of the attack is generated under circumstances that are not easily perceptible by the human eye, an adversarial attack method based on the perceived color distance is first constructed. The attack function is illustrated below.

$$minimize \left\| \Delta E_{00}(x^{mix}, x^{'}) B_{mn=1} \right\|_2 + \alpha f(x^{'}) \tag{5}$$

The former of Equation (5) denotes the addition of perceptual color difference-based perturbations within a well-determined salient region (i.e., $B_{mn=1}$), where the perceptual color distance $CIEDE2000$ is used, as shown in Equation (6) below, where $\Delta L^{'}$, $\Delta C^{'}$, and$\Delta H^{'}$ are the distances between the three-channel pixel values of L (luminance), C (chromaticity), and H (hue) in CIELCH space, which has been experimentally shown to better match the human visual perception. The latter term $f(x') = max(max\left\{Z(x^{'})_i : i \neq t\right\} - Z(x^{'})_t, 0)$ is a common function that can guarantee the success rate of the attack in CW at present.

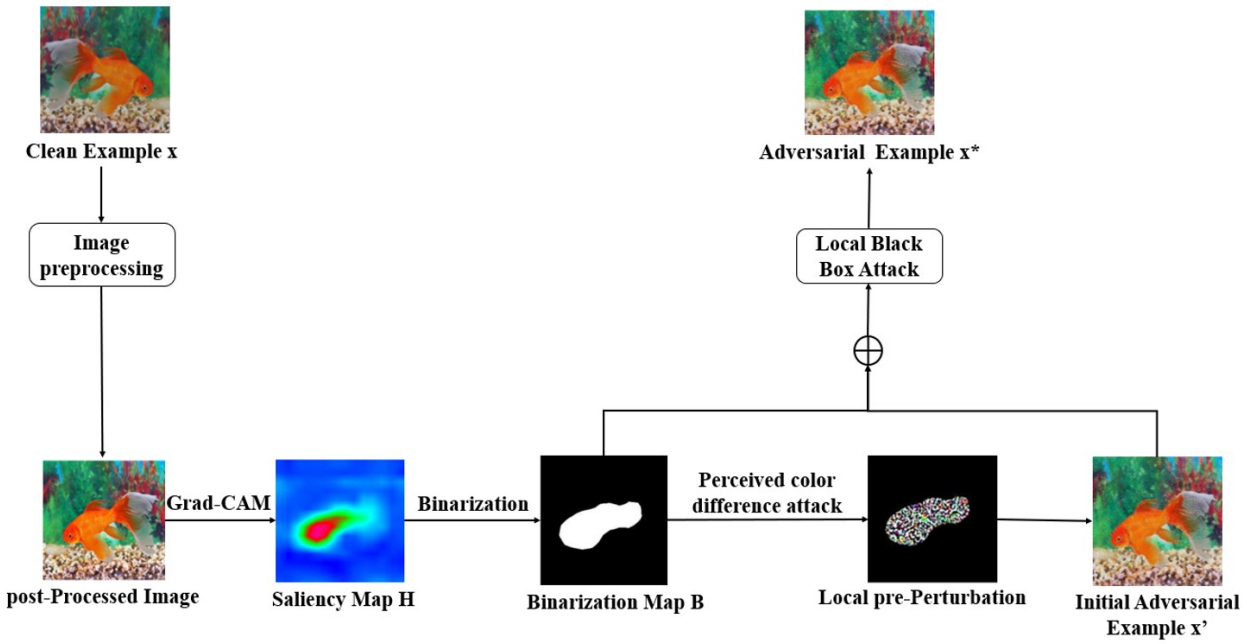


Figure 4: The overall framework of the attack.

In conclusion, the primary distinction between the method in this paper and the CW attack is the local attack, and secondly, the perceived color distance is used to guarantee the visual quality.

$$\Delta E_{00} = \sqrt{(\frac{\Delta L'}{k_L S_L})^2 + (\frac{\Delta C'}{k_C S_C})^2 + (\frac{\Delta H'}{k_H S_H})^2}$$
$$where \quad \Delta R = R_T(\frac{\Delta C^{'}}{k_C S_C})(\frac{\Delta H^{'}}{k_H S_H}) \tag{6}$$

3.2.4 Local Black-box Attacks

Since the initial adversarial sample is closer to the decision boundary of the target model than the original sample, in order not to waste query cost, we choose to estimate the gradient for pixels in the significant region (i.e.,$B_{mn=1}$) and use a random grouping strategy similar to that in [10] to improve query efficiency. All calculations depend only on the output of the true category y, and if the top-1 category is not y, it means we find an adversarial sample. The output probability of the true category is denoted by F_y and h is a small constant (set it to 0.0001 in the experiment). After obtaining the overall estimate reflecting the growth direction of F_y according to Equation (7), the probability is reduced by updating $x^{'}$ along the opposite direction. Eventually, the final adversarial sample is generated by combining with the BIM method, starting from $x_0^{'} = x^{'}$ continuously iteratively, and Algorithm 1 shows the specific process.

$$\nabla_{x'} = \begin{cases} \frac{F_y(x^{'}+hB_{mn}) - F_y(x^{'}-hB_{mn})}{2h}, & B_{mn=1} \\ 0, & B_{mn=0} \end{cases} \tag{7}$$

$$x_{t+1}^{*} = clip(x_t^{'} + \epsilon sign(-\nabla_{x'} F_y)) \tag{8}$$

3.2.5 E-LBQA

Given that the significant regions generated by different models are not exactly the same, we propose an improved approach (E-LBQA) to further improve query efficiency by employing multiple reference models, in which the adversarial samples are more portable, making the initial adversarial samples more likely to cheat the target model. Only part of the details of Algorithm 1 needs to be changed to accommodate multiple reference models. First, we apply Grad-CAM to each of the selected several alternative models separately, take the concatenation of all white regions of the binary maps to determine the final significant discriminative regions, which may contain the complete local region on which the target model depends, and generate the initial adversarial sample by Equation (9), followed by a local black-box attack.

$$minimize \left\| \Delta E_{00}(x^{mix}, x^{'}) \bigotimes U_{B_i} \right\|_2 + \alpha f(x^{'}) \tag{9}$$

$$H = \sum_{i}^{U} J(x^{'}, y; \theta_{f_i}) \tag{10}$$

$$U_{B_i} = \sum_{i}^{U} B_i \tag{11}$$

where U is the number of alternative models, U_{B_i} is the concatenation of binary maps of significant regions of several alternative models, θ_{f_i} and i_{th} is the model with parameter θ.

Algorithm 1 LBQA (Local Black-box Query Attack)

Input:
benign image x; real label y; reference model f; target model F; perturbation ϵ; iteration numbers T.

Output:
final adversarial example x^*.

Begin
update x^{mix} by Eq.(3)
input to reference model f
obtain Saliency Map H
obtain Binary Map B by Eq.(4)
local perturbation by Eq.(5), obtain initial adversarial sample $x^{'}$
input $x^{'}$ to target model F
for $t = 0$ to $T - 1$ **do**
  $x^*_{t+1} = clip(x^{'}_t + \epsilon sign(- \bigtriangledown_{x'} F_y))$
  if $x^{'}$ succeeds mislead F **then**
    end attack with x^* as the final adversarial sample.
  else
    return None
  end if
end for
return $x^* = x_t$

4 Experiment

In this section, we first introduce the experimental setup, including the datasets, the models, and the evaluation metrics of the adversarial samples. Then the performance of the proposed approach is investigated and compared with local attack-based versus decision-based attack approaches. In addition, it is demonstrated that attacks on significant areas of the model are more successful than attacks on non-significant regions by contrasting two particular sets of adversarial samples. Finally, ablation experiments are performed to verify the effectiveness of using the K-value and model significant graph merging sets on the attack performance.

4.1 Experimental Setup

4.1.1 Datasets

The MNIST and CIFAR-10 datasets are used in the experiment. The former dataset contains a total of 10 classes of size 28×28 with categories 0-9, divided into 60,000 training images and 10,000 test images. There are 60,000 images classified into 10 categories in the CIFAR-10 datasets, with 6,000 images in each category having a 32×32×32 pixel image size. Since the images of the MNIST are gray, there is no need to binarize them. The images in these two datasets also have fewer pixel points, and the model interpreter's visual output is not particularly excellent. Therefore, to improve the attack ability on high-resolution images, we conducted experiments on the ImageNet-compatible dataset containing 1000 color images with a resolution of 224 × 224 and scaled their image pixel points to the range [0,1].

Table 1: Classification accuracy of different datasets trained on different models.

Model/Datasets	MNIST	CIFAR-10	ImageNet
VGG16	99.22%	93.52%	88.97%
VGG19	99.28%	93.44%	88.93%
Inc_v3	99.62%	94.48%	96.23%
ResNet_50	99.36%	93.49%	94.93%
ResNet_101	99.55%	93.93%	96.33%
ResNet_152	99.62%	93.88%	95.51%
Inc_Res_v2	99.64%	93.62%	98.84%
DenseNet_121	99.61%	94.85%	94.63%

4.1.2 Models

Table 1 displays the classification accuracy results from various models trained on various datasets, including the high-resolution dataset ImageNet. These models include VGG16, VGG19, Inc_v3, ResNet_50, ResNet_152, and so on for each dataset.

The learning rate is set to 0.01 for the C&W attack, and the optimization process is stopped early if the loss is not decreased after 1000 iterations. The perturbation was set to 0.05 for all other comparison methods, using the same hyperparameter settings as the original authors.

4.1.3 Assessment Metrics

In this paper, we examine the quality of the adversarial samples by attack success rates, average L_p norm, and the number of queries. The percentage of samples that successfully deceive the target model out of all generated adversarial samples is known to attack success rate(ASR). Average L_p norm used to measure the size of the perturbation. Here, L_2 distance can be chosen to measure the perturbation size and we use L_0 distance to measure the number of perturbed pixels.

$$\|\delta\|_p = (\sum_{i=1}^{n} |\delta_i|^p)^{-p} \tag{12}$$

where δ_i is the adversarial perturbation of the pixel of the adversarial sample. In addition, we measure them by numbers required to attack a successful adversarial sample within a given query budget, which is fed into the target model. Moreover, we introduce SSIM to measure the similarity between the original image x and the adversarial examples. The larger the SSIM value, the higher the similarity between the two images.

$$SSIM\,(x, x^*) = [l\,(x, x^*)]^{\alpha}\,[c\,(x, x^*)]^{\beta}\,[s\,(x, x^*)]^{\gamma} \tag{13}$$

where $\alpha, \beta, \gamma > 0$, $l\,(x, x^*)$ is brightness comparison, $c\,(x, x^*)$ is contrast comparison, $s\,(x, x^*)$ is structure comparison.

4.2 Attack Success Rate & Image Quality

Since our attack process is partially similar to the CW attack, Table 2 compares the average attack success rate, L_2 distance, and SSIM values of LBQA with CW attack on three datasets using the alternative model as VGG16 and target model as VGG19, Inc_v3, ResNet_50, and ResNet_101, respectively. (*) denotes white-box attack and (B) denotes black-box attacks. As can be seen from the first row, the CW attack continues to have a high attack success rate against the white box and introduces fewer adversarial perturbations. In contrast, the CW black-box attack based on the alternative model has a significantly lower attack success rate and reduced invisibility as a result of not knowing its internal information. For our proposed attack, it improves the performance of the black-box attack and the success rate of the attack also improves a bit, for example, when attacking ResNet_101, the success rate of CW(B) on the ImageNet dataset is 17.36%, while LBQA improves to 83.17%, and it is worth noting that the average L_2 distance of LBQA is not only lower than CW(B) but also its SSIM value on the three datasets also performs optimally, which shows that our method improves both the success rate of black-box attacks and ensures the quality of visual perception.

To evaluate the performance of the proposed LBQA, we further considered the attack success rate and image quality of several local attack methods, namely the classical algorithms JSMA, SGA, and FineFool. The alternative model used is VGG16 and the target model is VGG19 and ResNet_50 respectively for the experiments. As shown in Table 3, our LBQA always outperforms JSMA in terms of L_2 perturbation norm and attack success rate. Although the proportion of perturbed pixels to all images (i.e., L_0 distance) is very small for JSMA, its cost is relatively high, and its SSIM value on the CIFAR-10 dataset is only 0.713. Moreover, as the complexity of the experimental dataset increases, our attack success rate is much higher than the other three local attacks, and although the L_0 value on the MNIST is not the lowest, the L_0 value on the CIFAR-10 is lower than the SGA and FineFool attack, and the L_2 value on ImageNet responds to better image quality. More importantly, it can be seen from Table 3 that our attack achieves optimal results in terms of SSIM metrics, for example, the SSIM value on the ImageNet dataset reaches 0.972 when attacking VGG19, which means that the adversarial sample is closer to the original sample. As can be seen, LBQA has great applicability and visual quality despite not being the best attack method.

The generated adversarial samples from the MNIST and ImageNet datasets are shown in Figure 5 respectively. In Figure 5(a), the middle two columns in the MNIST dataset are the original images, and the two sides are the adversarial samples generated by LBQA, while in Figure 5(b), the first column in the figure is the original image, the second column is the adversarial samples generated by CW(B), and the third column is the image

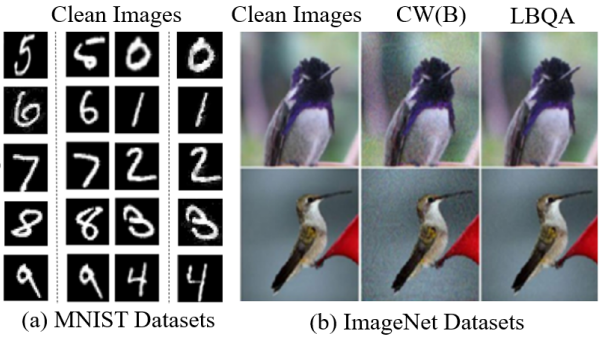


Figure 5: Visualization of the adversarial samples.

generated by LBQA. we can see that CW(B) has obvious perturbation, while LBQA generates images that are less likely to be perceived by the human eyes and somewhat smoother visually due to the consideration of spatial semantic structure.

4.3 Attack Success Rate & Number Of Queries

Additionally, Table 4 displays the performance of the attack method in this study using the query-based approach of ImageNet dataset (the alternative model is Inc_v3) when attacking the target models, which are respectively DenseNet_121, ResNet_50, and VGG16. We set the budget at 1000 queries to test the performance of the attack method under the constraint of L_∞ perturbation (set the perturbation range ϵ set to 0.05) under the attack success rate and the average number of queries required for a successful attack. Compared with the first two methods, our attack significantly improves its effectiveness, reaching a success rate of about 84%. Except for DenseNet_121, the number of queries required by LBQA on the other two models is lower than the first three methods, although the success rate is not as high as SimBA, the number of queries is also reduced.

4.4 The Effect of K-value

Here, we investigate how the binarization factor k affects an attack's capabilities on the ImageNet dataset. Intuitively, if k is too small, most of the image will be modified, and thus the significant map attack is meaningless. If k is too large, the effect is weak since just a small portion of the image is altered. As can be seen in Figure 6, The value of k is changed from 0 to 11/12 to demonstrate the performance difference, the success of the attack increases initially before decreasing. When roughly $k = 1/3$, the performance is at its peak, and when k is too large, good attack capability cannot be realized.

Table 2: Attack capability of CW and LBQA methods on different datasets

Target Models	Attack Approach	MNIST			CIFAR-10			ImageNet		
		ASR(%)	L2	SSIM	ASR(%)	L2	SSIM	ASR(%)	L2	SSIM
VGG16	CW(*)	100%	1.494	0.957	100%	0.181	0.952	98.42%	1.532	0.946
VGG19	CW(B)	33.20%	3.611	0.942	25.30%	2.991	0.933	19.31%	2.652	0.941
	LBQA	**85.60%**	**2.998**	**0.964**	**73.20%**	**2.511**	**0.974**	**82.25%**	**2.013**	**0.972**
Inc_v3	CW(B)	29.21%	3.456	0.951	24.35%	2.841	0.948	17.68%	2.352	0.943
	LBQA	**81.03%**	**2.753**	**0.973**	**74.24%**	**2.318**	**0.976**	**81.24%**	**2.183**	**0.973**
ResNet_50	CW(B)	32.74%	3.523	0.954	25.81%	2.811	0.942	18.65%	2.245	0.938
	LBQA	**84.31%**	**2.845**	**0.976**	**75.31%**	**2.437**	**0.978**	**83.16%**	**2.018**	**0.978**
ResNet_101	CW(B)	35.82%	3.162	0.923	26.20%	2.784	0.938	17.36%	2.438	0.945
	LBQA	**82.86%**	**2.658**	**0.968**	**79.81%**	**2.561**	**0.976**	**83.17%**	**2.021**	**0.971**

Table 3: Attack success rate and image quality of local attack methods

Target Models	Datasets	Attack Approach	ASR(%)	L0	L2	SSIM
VGG16→ VGG19	MNIST	JSMA	12.50%	4.12%	3.38	0.927
		SGA	42.30%	32.80%	2.15	0.873
		FineFool	84.60%	55.28%	0.93	0.786
		LBQA	**85.40%**	54.49%	2.32	**0.978**
	CIFAR-10	JSMA	8.40%	5.27%	3.61	0.713
		SGA	50.10%	85.10%	2.78	0.905
		FineFool	28.93%	74.73%	0.64	0.911
		LBQA	**73.20%**	**52.31%**	2.31	**0.963**
	ImageNet	JSMA	7.80%	5.12%	3.27	0.942
		SGA	56.20%	21.98%	2.86	0.975
		FineFool	46.80%	12.51%	2.85	0.963
		LBQA	**82.25%**	31.23%	2.01	**0.972**
VGG16→ ResNet_50	MNIST	JSMA	11.35%	4.19%	3.25	0.938
		SGA	43.56%	31.68%	2.23	0.854
		FineFool	83.26%	56.43%	0.96	0.763
		LBQA	**85.34%**	52.24%	2.12	**0.981**
	CIFAR-10	JSMA	8.65%	6.17%	3.54	0.683
		SGA	52.31%	85.19%	2.69	0.918
		FineFool	26.73%	73.34%	0.78	0.906
		LBQA	**74.26%**	**53.68%**	2.28	**0.976**
	ImageNet	JSMA	7.92%	5.32%	3.23	0.948
		SGA	56.29%	22.98%	2.74	0.967
		FineFool	45.83%	12.67%	2.89	0.959
		LBQA	**82.38%**	31.36%	2.09	**0.974**

Table 4: Attack success rate and the number of queries

Attack Approach	DenseNet_121		ResNet_50		VGG16	
	ASR(%)	Q	ASR(%)	Q	ASR(%)	Q
Sign-opt	5.38%	476.29	5.52%	456.21	10.41%	478.9
HSJA	8.64%	167.38	12.21%	144.53	17.31%	160.23
SimBA	92.41%	334.12	90.26%	348.21	93.02%	319.14
LBQA	**82.62%**	172.65	**83.10%**	**112.34**	**85.32%**	**128.63**

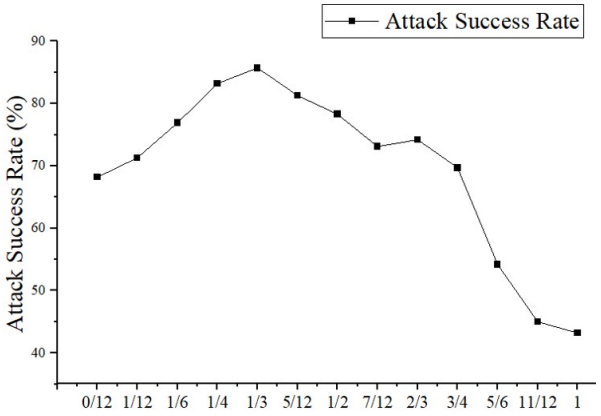


Figure 6: Effect of k-value on the success rate of the LBQA.

4.5 The Impact of Model Salient Regions on Classification

This section evaluates the impact of significant regions on the model's classification ability using two sets of special adversarial samples. In one set of images, the pixels in the significant region of the model are kept unchanged, and the rest of the pixels are set to 0, denoted as LBQA-S. In the other set of images, the pixels in the significant region are set to 0 and the rest of the pixels are kept unchanged, known as LBQA-NS. On CIFAR-10, 10,000 clean images are converted into these adversarial images using a threshold of 1/3. Figure 7 below shows the adversarial samples generated by LBQA-S and LBQA-NS. Figure 8 displays the success rate of the LBQA-S and LBQA-NS using the alternative model as VGG16 and the target model as VGG19, ResNet_50, ResNet_101 and ResNet_152, respectively. We can observe that the attack success rate of LBQA-NS is just nearly 20%, while LBQA-S performs much better than LBQA-NS. This depicts that the significant regions, despite being smaller in size than the non-significant regions, are very important for categorization. Therefore, altering the pixel regions that are more important to the classification of the model rather than altering other regions would enable the adversary to trick the classifier more successfully.

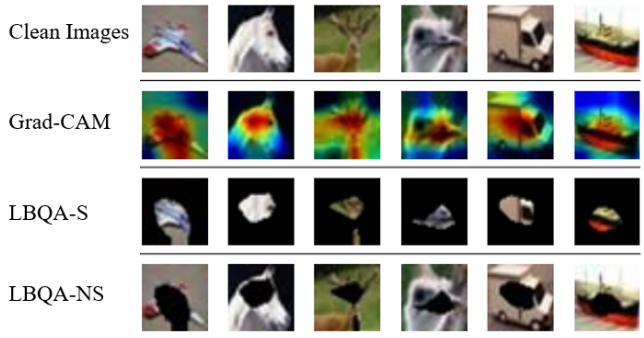


Figure 7: Visualization of LBQA-S and LBQA-NS on CIFAR-10.

Table 5: Validity of the model significance map concatenation

Attack	ASR	L0	L2	SSIM
LBQA	80.51%	45.63%	2.31	**0.962**
E-LBQA	89.74%	54.23%	1.86	**0.981**

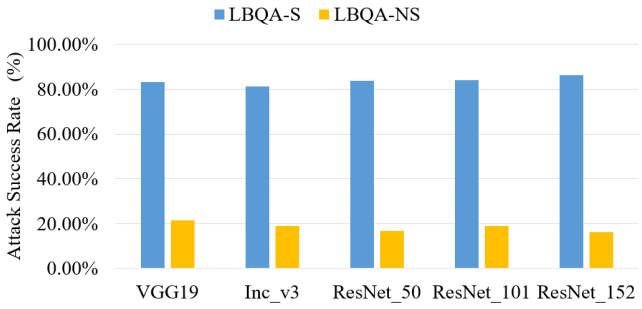


Figure 8: The impact of different target models on the success rate of LBQA-S and LBQA-NS.

4.6 Validity of Model Significance Map Concatenation

Finally, to verify the effectiveness of the model salient map merging on the quality of the generated adversarial samples, we used the ImageNet dataset, VGG16, Inc_v3, and ResNet_50 as alternative models, and the target model was Inc_res_v2 for the attack, as shown in Table 5, E-LBQA is higher than LBQA in L_0, but better than LBQA in L_2 values as well as SSIM, additionally, its success rate has also increased by almost 10%, demonstrating the effectiveness of the concatenation using model attention significant map.

5 Conclusion

By using the spatial semantic information of images to limit the attack zone, this work offers a new local black-box adversarial attack (LBQA) based on the perceived color distance that decreases the dimensionality of black-box attacks and increases query efficiency. Experimental findings demonstrate that LBQA and E-LBQA can effectively perform black-box adversarial attacks while maintaining the visual quality of the adversarial samples by combining transfer-based and query-based techniques. Our LBQA also shows its advantages and potential on high-resolution images (like ImageNet datasets) compared to conventional methods. Our future work will further improve the performance of adversarial attacks by performing genetic algorithms in clearly defined regions and utilizing pixel correlation to target only one out of every two surrounding pixels in a salient area.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grant 61562065 and the Inner Mongolia Natural Science Foundation Project under Grant 2019MS06001 and Grant 2023MS06012.

References

[1] N. Akhtar, A. Mian, N. Kardan, and M. Shah, "Advances in adversarial attacks and defenses in computer vision: A survey," *IEEE Access*, vol. 9, pp. 155 161–155 196, 2021.

[2] A. Almestekawy and M. Abdulsalam, "Sentiment analysis of product reviews using bag of words and bag of concepts," *International Journal of Electronics and Information Engineering*, vol. 11, no. 2, pp. 49–60, 2019.

[3] M. Andriushchenko, F. Croce, N. Flammarion, and M. Hein, "Square attack: a query-efficient black-box adversarial attack via random search," in *16th European Conference on Computer Vision (ECCV'20)*, pp. 484–501, 2020.

[4] A. Bhagoji, W. He, B. Li, and D. Song, "Practical black-box attacks on deep neural networks using efficient query mechanisms," in *Proceedings of the European conference on computer vision (ECCV'18)*, pp. 154–169, 2018.

[5] N. Carlini and D. Wagner, "Towards evaluating the robustness of neural networks," in *IEEE symposium on security and privacy (SP'17)*, pp. 39–57, 2017.

[6] J. Chen, H. Zheng, H. Xiong, R. Chen, T. Du, Z. Hong, and S. Ji, "Finefool: A novel DNN object contour attack on image recognition based on the attention perturbation adversarial technique," *Computers & Security*, vol. 104, p. 102220, 2021.

[7] J. Chen, M. Jordan, and M. Wainwright, "Hopskipjumpattack: A query-efficient decision-based attack," in *IEEE Symposium on Security and Privacy (SP'20)*, pp. 1277–1294, 2020.

[8] P. Chen, H. Zhang, Y. Sharma, J. Yi, and C.-J. Hsieh, "Zoo: Zeroth order optimization based black-box attacks to deep neural networks without training substitute models," in *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security*, pp. 15–26, 2017.

[9] M. Cheng, S. Singh, P. Chen, P.-Y. Chen, S. Liu, and C.-J. Hsieh, "Sign-OPT: A query-efficient hard-label adversarial attack," *arXiv preprint arXiv:1909.10773*, 2019.

[10] X. Dong, J. Han, D. Chen, J. Liu, H. Bian, Z. Ma, H. Li, X. Wang, W. Zhang, and N. Yu, "Robust superpixel-guided attentional adversarial attack," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12 895–12 904, 2020.

[11] Y. Dong, F. Liao, T. Pang, H. Su, J. Zhu, X. Hu, and J. Li, "Boosting adversarial attacks with momentum," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9185–9193, 2018.

[12] Y. Dong, T. Pang, H. Su, and J. Zhu, "Evading defenses to transferable adversarial examples by translation-invariant attacks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4312–4321, 2019.

[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, 2014.

[14] J. Goodfellow, I.J.and Shlens and C. Szegedy, "Explaining and harnessing adversarial examples," *arXiv preprint arXiv:1412.6572*, 2014.

[15] C. Guo, J. Gardner, Y. You, A. G. Wilson, and K. Weinberger, "Simple black-box adversarial attacks," in *International Conference on Machine Learning*, pp. 2484–2493, 2019.

[16] L. C. Huang, L. Y. Tseng, and M. S. Hwang, "A reversible data hiding method by histogram shifting in high-quality medical images," *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, 2013.

[17] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548-555, 2000.

[18] A. Kurakin, I. J. Goodfellow, and S. Bengio, "Adversarial examples in the physical world," in *Artificial Intelligence Safety and Security*, pp. 99–112, 2018.

[19] X. Li, S. Ji, M. Han, J. Ji, Z. Ren, Y. Liu, and C. Wu, "Adversarial examples versus cloud-based detectors: A black-box empirical study," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 4, pp. 1933–1949, 2019.

[20] I. C. Lin, H. H. Ou, M. S. Hwang, "A user authentication system using back-propagation network", *Neural Computing & Applications*, vol. 14, pp. 243-249, 2005.

[21] J. Lin, C. Song, K. He, L. Wang, and J. E. Hopcroft, "Nesterov accelerated gradient and scale invariance for adversarial attacks," *arXiv preprint arXiv:1908.06281*, 2019.

[22] A. Madry, A. Makelov, L. Schmidt, and T. D. V. A, "Towards deep learning models resistant to adversarial attacks," *arXiv preprint arXiv:1706.06083*, 2017.

[23] S. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, "Deepfool: a simple and accurate method to fool deep neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2574–2582, 2016.

[24] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. Celik, and A. Swami, "The limitations of deep learning in adversarial settings," in *IEEE European*

Symposium on Security and Privacy (EuroSP'16), pp. 372–387, 2016.

[25] J. Su, D. Vargas, and K. Sakurai, "One pixel attack for fooling deep neural networks," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 828–841, 2019.

[26] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," *arXiv preprint arXiv:1312.6199*, 2013.

[27] J. X. Tong, H. Li, and S. L. Yin, "Research on face recognition method based on deep neural network," *International Journal of Electronics and Information Engineering*, vol. 12, no. 4, pp. 182–188, 2020.

[28] X. Wang and K. He, "Enhancing the transferability of adversarial attacks through variance tuning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1924–1933, 2021.

[29] H. J. Wu, Y. H. Chang, M. S. Hwang, I. C. Lin, "Flexible RFID location system based on artificial neural networks for medical care facilities", *ACM SIGBED Review*, vol. 6, no. 2, pp. 1-8, 2009.

[30] L. Yan, X. Wang, and S. Yin, "Campus garbage image classification algorithm based on new attention mechanism," *International Journal of Electronics and Information Engineering*, vol. 13, no. 4, pp. 131–141, 2021.

[31] W. Zhang, X. Zhang, K. Hao, J. Wang, and S. Zhang, "Optimized jacobian-based saliency maps attacks," *International Journal of Network Security*, vol. 24, no. 6, pp. 1020–1030, 2022.

[32] X. Zhang, X. Wang, and S. Yin, "Multi-modal data transfer learning-based lstm method for speech emotion recognition," *International Journal of Electronics and Information Engineering*, vol. 13, no. 2, pp. 54–65, 2021.

Biography

Xiaolin Zhang was born in Baotou, China, in December 1966. She received the bachelor's degree in computer science and technology from Northeastern University, in 1988, the master's degree in automation from the Beijing University of Science and Technology, in 1995, and the Ph.D. degree in computer science and technology from Northeastern University, in 2006. Since 1988, she has been with the Inner Mongolia University of Science and Technology, where she is currently the Deputy Director of the Head of the Computer Science Department, the Professor Committee of the Information Technology College, and the Director of the Department of Computer Science. She has trained more than 70 master's degree students and now is training 12 master's degree students. She has published over 80 academic articles, including more than 30 articles in EI and 7 articles in SCI. She is responsible for many projects, such as the National Natural Science Foundation of China, the National Social Science Fund Project, the Chunhui Project of the Ministry of Education, the Natural Science Foundation of Inner Mongolia Project, and the Inner Mongolia Education Department Fund Project. Her current research interests include image processing, natural language processing, adversarial attacks of image and text, machine learning security, big data processing technology, and social network privacy protection technology. Dr. Zhang is a member of the Chinese Computer Society, the Information System Professional Committee, the China Computer Society, and the Director of the Inner Mongolia Autonomous Region Computer Society.

Jing Shi was born in Taiyuan, China in June 1997. She received her B.S. degree in Computer Science and Technology from Changzhi College in 2021. She is currently pursuing the M.S. degree in Computer Science and Technology at the Inner Mongolia University of Science and Technology. Her research interests include adversarial attacks based on image classification.

Enhui Xu was born in Xuzhou, China in January 1997. He received the B.S. degree in software engineering of Science and Technology in 2019 and the master's degree in computer science and technology in 2022 from the Inner Mongolia University. He is working in Nanhu Academy of Electronics and Information Technology, China. His research interests include machine learning security and natural language processing and adversarial samples.

Yongping Wang was born in Baotou, China, in 1984. She received the bachelor's degree in computer science and technology and the master's degree from the Wuhan University of Technology, in 2007 and 2010, respectively. Since 2010, she has been with the Inner Mongolia University of Science and Technology, where she is currently a lecturer with the School of Information Engineering. She has participated in a number of research projects of the Inner Mongolia Natural Science Foundation. Her main research interests include image processing, machine learning security and adversarial attacks on image.

Wenwen Zhang was born in Heze, China, in November 1996. She received the B.S. degree in computer science and technology from the Jining Medical University, in 2019. She is currently pursuing the master's degree in computer science and technology with the Inner Mongolia University of Science and Technology. Her research interests include adversarial attacks based on image classification.

A New Family of Universal Hash Functions for Quantum Key Distribution

Shuying Yang
School of Data and Computer Science, Shandong Women's University
No. 2399, University Road, Jinan 250300, Shandong, P. R. China
Email: ysystudy2005@163.com

(Received Jan. 10, 2023; Revised and Accepted Sept. 12, 2023; First Online Oct. 11, 2023)

Abstract

Universal hash functions are a very important family of functions used in quantum key distribution, which enjoys a higher level of security than traditional key establishment protocols. The polynomial hash functions are the most widely used universal hash functions. Recently, Bibak *et al.* generalized the polynomial hash functions to multivariate polynomial hash functions. In this paper, we propose a family of matrix polynomial hash functions that extend the classical polynomial hash functions in a new direction. Furthermore, to prove the proposed hash family is ϵ-almost-Δ-universal, we invent a technique that recursively transforms a matrix polynomial into a system of linear matrix equations. The technique may also be interesting in dealing with other problems about matrix polynomials.

Keywords: Message Authentication Code (MAC); Quantum Key Distribution (QKD); Universal Hash Function (UHF)

1 Introduction

Universal hash functions, due to Carter and Wegman [9], have many applications in cryptography, informatiom security, erro-correcting codes, pseudorandomness, complexity theory, randomized algorithms, data structures, parallel computing, etc. [2, 4–6, 11].

In the field of cryptography, a kind of basic problem is how to establish a common key over a public channel for two parties which are not in the same place. The tool to solve the problem is key establishment protocol which usually uses methods from number theory, like the classical Diffie-Hellman key exchange protocol [12]; see [7] for a comprehensive treatment of the theme.

However, the security of the traditional key establishment protocols relies on the computational hardness of certain problems in number theory, such as the discrete logarithm problem [12, 14] and the integer factorization problem [10, 17], which are shown to be solved on a sufficiently powerful quantum computer running Shor's algorithm [18].

Quantum key distributions (QKD) enjoy a higher level of security than such traditional key establishment protocols, as its security relies on the foundation of quantum mechanics rather than the computational hardness of certain mathematical problems. Thus, QKD has been proved to be secure even against an adversary with unbounded computational power [8].

QKD requires a quantum channel and a classical channel. First, through the quantum channel, some quantum states are obtained and measured. Then, through the classical channel, the parties determine which results of their measurements can generate secret bits. Next, they perform error correction and privacy amplification. In addition, to avoid man-in-the-middle attack, it is vital to authenticate these steps with a pre-shared secret [16, 19].

Universal hash functions, which are used in at least three steps, are one of the most important functions used in QKD. In the step of error correction, QKD fix any noise which may occur during communication with error correcting codes which, however, correspond to universal hash functions directly. In the step of privacy amplication, the raw key material is compressed using a shared secret universal hash function. In the step of authentication, the message authentication codes (MACs) proposed by Wegman and Carter [9] and its variants [5, 6] are usually used, all of which use universal hash functions in their constructions.

One of the most widely used universal hash function families is polynomial hash functions, such as Galois/Counter Mode [13] used in IPsec, SSH and TLS, and Poly 1305 [1] used in Google Chrome's TLS and OpenSSL. Recently, Bibak *et al.* proposed two other families of universal hash functions [6] and their generalization [3]. Li *et al.* also proposed a fast hash family for memory integrity [15]. But their proof is somewhat flawed, as we will point out at the end of this paper.

Our Contributions. In this paper, we generalize the classical polynomial hash functions using matrix theory over finite fields.

1) We first generalize the polynomial hash functions

over prime field $\mathbb{Z}_p$ to a matrix polynomial functions over $\mathbb{Z}_p$. The underlying main ideas are inspired by the following two observations: The first one is that message modularization will greatly improve the throughput and efficiency of hash functions. The most natural modularity is to group the message vectors into appropriately sized vectors. But the problem here is that vectors don't have powers, so we can't define polynomials. The second observation is square matrices have power operations, thus we can construct a matrix polynomial hash function using square matrices as variables and message vectors as coefficients.

2) The proof of the universal property of the matrix polynomial hash functions is far less straightforward and easy than that of the classical polynomial hash functions. The proof of the classical polynomial hash functions is directly based on the fundamental theorem of algebra which, however, does not apply to the matrix polynomial hash functions. In order to overcome the difficulty, we first discuss the case of a linear matrix equation, and then solve the problem by recursively transforming a matrix polynomial into a system of linear matrix equations. The new technique may be also useful in dealing with other problems about matrix polynomials.

2 Preliminaries

Notation. We use $\mathbb{Z}$ for the set of integers and $\mathbb{Z}_n$ for the ring of integers modulo n defined as $\mathbb{Z}_n = \{0, \ldots, n-1\}$. Elements of these sets are denoted by lowercase letters. We use uppercase letters in bold font (such as $\mathbf{M}$) to denote matrices, lowercase letters in bold font (such as $\mathbf{v}$) to denote vectors. Matrices enclosed by square brackets, such as $[\mathbf{M_1}\ \mathbf{M_2}]$, refer to an augmented matrix formed by horizontally joining the columns of $\mathbf{M_1}$ and $\mathbf{M_2}$, assuming the number of rows in these two matrices are the same. Matrices (or vectors) enclosed by parenthesis, such as $(\mathbf{v_1}, \mathbf{v_2})$, refer to an matrix or vector formed by vertically joining the rows of $\mathbf{v_1}$ and $\mathbf{v_2}$. For a set S, we write $s \leftarrow S$ to denote that s is chosen uniformly at random from S.

2.1 Universal Hash Functions

In this subsection, we review several related definitions of universal hash functions.

Definition 1. (Keyed hash functions) *Let $\mathcal{K}$, $\mathcal{M}$ and $\mathcal{T}$ be finite, non-empty sets. A keyed hash function $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ is a function that takes two inputs: a key k and a message m, and outputs a digest $t = H(k, x)$, where $\mathcal{K}$, $\mathcal{M}$ and $\mathcal{T}$ are called keyspace, message space and digest space, respectively.*

Definition 2. (Universal hash functions) *Let $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ be a keyed hash function. H is universal if for any two distinct $x, y \in \mathcal{M}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) = H(k, y)] \leq \frac{1}{|\mathcal{T}|}$. Furthermore, H is an ϵ-almost universal (ϵ-AU) hash function if for any two distinct $x, y \in \mathcal{M}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) = H(k, y)] \leq \epsilon$.*

Definition 3. (Δ-Universal hash functions) *Suppose $\mathcal{T}$ is a finite additive Abelian group. Let $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ be a keyed hash function. H is Δ-universal if for any two distinct $x, y \in \mathcal{M}$ and all $b \in \mathcal{T}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) - H(k, y) = b] = \frac{1}{|\mathcal{T}|}$, where '$-$' denotes the group subtraction operation. Furthermore, H is an ϵ-almost-Δ-universal (ϵ-AΔU) hash function if for any two distinct $x, y \in \mathcal{M}$ and all $b \in \mathcal{T}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) - H(k, y) = b] \leq \epsilon$. When $\mathcal{T} = \mathbb{Z}_2^n = \{0, 1\}^n$ for some integer n, the operation '$-$' can be replaced by '$\oplus$' (XOR), and H is also called ϵ-almost XOR universal hash function.*

Remark 1. Due to Definition 2 and 3, it's easy to see that Δ-Universal hash functions are also universal hash function, ϵ-AΔU hash functions are also ϵ-AU hash functions since $\mathcal{T}$ is a finite additive Abelian group in Definition 3 and b could be equal to 0.

Definition 4. *Let $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ be a keyed hash function. H is ϵ-balanced if for any nonzero $x \in \mathcal{M}$ and any $y \in \mathcal{T}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) = y] \leq \epsilon$.*

2.2 Polynomial Hash Functions

Universal hash functions constructed using polynomials modulo a prime is widely attributed to Wegman and Carter [9]. In this subsection, we review this construction.

Definition 5. *Given an integer n and a prime p. Let $\mathcal{K} = \mathbb{Z}_p$, $\mathcal{M} = \mathbb{Z}_p^{d+1}$, and $\mathcal{T} = \mathbb{Z}_p$. Define hash function $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ as*

$$H(k, m) = \sum_{i=0}^{d} m_i k^i (mod p),$$

for every message $\mathbf{m} = (m_0, m_1, \ldots, m_d) \in \mathcal{M}$ and every key $k \in \mathbb{Z}_p$.

Theorem 1. *[3] The hash function defined above is $\frac{d}{p}$-almost-Δ-universal.*

3 Proposed Universal Hash Functions

In this section, we propose a new family of universal hash function whose idea is inpired by the construction of polynomial hash functions.

Definition 6. *Given integers n, r and a prime p. Let $\mathcal{K} = \{$all n-by-n invertible matrices over finite field $\mathbb{Z}_p\}$, $\mathcal{M} = \mathbb{Z}_p^{nr}$ and $\mathcal{T} = \mathbb{Z}_p^n$. Define hash function $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ as*

$$H(\mathbf{K}, \mathbf{m}) = [\mathbf{K}^r\ \mathbf{K}^{r-1}\ \ldots\ \mathbf{K}]\mathbf{m}$$

where the multiplications and additions are done over $\mathbb{Z}_p$.

If we parse the messge vector $\mathbf{m}$ as r length-n blocks and write $\mathbf{m} = (\mathbf{m_1}, \mathbf{m_2}, \ldots, \mathbf{m_r})$, then the hash function defined above is equivalently represented by

$$H(\mathbf{K}, \mathbf{m}) = \mathbf{K}^r\mathbf{m_1} + \mathbf{K}^{r-1}\mathbf{m_2} + \cdots + \mathbf{K}\mathbf{m_r}.$$

This representation is very similar in form to the representation in definition 5. However, the proof of the universal property of this new hash function is far less straightforward and easy than the proof of the universal property of the function defined in definition 5. Here, let's first write down the conclusion that we're going to prove.

Theorem 2. *The hash function defined in Definition 6 is* $\frac{1}{p^n-1}$*-almost-*Δ*-universal.*

To prove the theorem, we first introduce and prove the following lemmas.

Lemma 1. *The hash function defined in definition 6 is linear. Specifically, for any two messages* $\mathbf{m}', \mathbf{m}'' \in \mathbb{Z}_\mathbf{p}^\mathbf{nr}$*, any key* $\mathbf{K}$ *and scalar* a*, it holds that* $H(\mathbf{K}, \mathbf{m}' + \mathbf{m}'') = \mathbf{H}(\mathbf{K}, \mathbf{m}') + \mathbf{H}(\mathbf{K}, \mathbf{m}'')$*, and* $H(\mathbf{K}, a\mathbf{m}') = aH(\mathbf{K}, \mathbf{m}')$.

Proof. Let $\mathbf{m}' = (\mathbf{m_1'}, \mathbf{m_2'}, \ldots, \mathbf{m_r'})$ and $\mathbf{m}'' = (\mathbf{m_1''}, \mathbf{m_2''}, \ldots, \mathbf{m_r''})$, where $\mathbf{m_i'}$ and $\mathbf{m_j''}$ are length-n blocks. Then

$$\begin{aligned}
&H(\mathbf{K}, \mathbf{m}' + \mathbf{m}'') \\
&= \mathbf{K^r}(\mathbf{m_1'} + \mathbf{m_1''}) + \mathbf{K^{r-1}}(\mathbf{m_2'} + \mathbf{m_2''}) + \cdots + \mathbf{K}(\mathbf{m_r'} + \mathbf{m_r''}) \\
&= (\mathbf{K^r}\mathbf{m_1'} + \mathbf{K^{r-1}}\mathbf{m_2'} + \cdots + \mathbf{K}\mathbf{m_r'}) \\
&\quad + (\mathbf{K^r}\mathbf{m_1''} + \mathbf{K^{r-1}}\mathbf{m_2''} + \cdots + \mathbf{K}\mathbf{m_r''}) \\
&= H(\mathbf{K}, \mathbf{m}') + \mathbf{H}(\mathbf{K}, \mathbf{m}'')
\end{aligned}$$

and

$$\begin{aligned}
&H(\mathbf{K}, a\mathbf{m}') \\
&= \mathbf{K^r}(a\mathbf{m_1'}) + \mathbf{K^{r-1}}(a\mathbf{m_2'}) + \cdots + \mathbf{K}(a\mathbf{m_r'}) \\
&= a(\mathbf{K^r}\mathbf{m_1'} + \mathbf{K^{r-1}}\mathbf{m_2'} + \cdots + \mathbf{K}\mathbf{m_r'}) \\
&= aH(\mathbf{K}, \mathbf{m}')
\end{aligned}$$

□

Lemma 2. *Let* $H : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ *be a linear hash function. Then* H *is a* ϵ*-almost-*Δ*-universal hash function if and only if* H *is a* ϵ*-balanced hash function.*

Proof. "$\Longrightarrow$" For any nonzero $x \in \mathcal{M}$ and any $y \in \mathcal{T}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) = y] = Pr_{k \leftarrow \mathcal{K}}[H(k, x - 0) = y] = Pr_{k \leftarrow \mathcal{K}}[H(k, x) - H(k, 0) = y] \leq \epsilon$, where the second equation is due to the linear property of H and the last inequality is since H is assumed as a ϵ-almost-Δ-universal hash function.

"$\Longleftarrow$" For any two distinct $x, y \in \mathcal{M}$ and all $b \in \mathcal{T}$, we have $Pr_{k \leftarrow \mathcal{K}}[H(k, x) - H(k, y) = b] = Pr_{k \leftarrow \mathcal{K}}[H(k, x - y) = b] \leq \epsilon$, where the second equation is due to the linear property of H and the last inequality is since H is assumed as a ϵ-balance hash function. □

Lemma 3. *Let* $n > 0$ *be an integer and* p *be a prime. For any nonzero* $\mathbf{m} \in \mathbb{Z}_p^n$ *and* $b \in \mathbb{Z}_p$*, the number of vectors* $\mathbf{k} \in \mathbb{Z}_p^n$ *such that* $\mathbf{k} \cdot \mathbf{m} = b$ *is* $p^n - 1$.

Proof. Regarding $\mathbf{K}$ as an unknown vector, the equation $\mathbf{k} \cdot \mathbf{m} = b$ is a system of n-element linear equations with only one equation. Let $\mathbf{k} = (k_1, \ldots, k_n)$, $\mathbf{m} = (m_1, \ldots, m_n)$. The coefficient matrix of this equation system is $\mathbf{m} = (m_1, \ldots, m_n)$. Since $\mathbf{m}$ is nonzero, the rank of the matrix is 1. Hence, the rank of the coefficient matrix $\mathbf{m}$ of this system is equal to the rank of the augmented matrix $(\mathbf{m}|b)$, which is equal to 1. Thus there must be a solution to this system of linear equations. The number of solutions of the system depends on the dimension of the solution space of the derived system $\mathbf{k} \cdot \mathbf{m} = 0$. By the theory of systems of linear equations, the dimension of the solution space of $\mathbf{k} \cdot \mathbf{m} = 0$ is $n - rank(\mathbf{m}) = n - 1$. Since the underlying number field is finite field $\mathbb{Z}_p$, the number of solutions of $\mathbf{k} \cdot \mathbf{m} = b$ is $p^n - 1$. □

Lemma 4. *Let* $n > 0$ *be an integer and* p *be a prime. For any nonzero* $\mathbf{m} \in \mathbb{Z}_p^n$ *and any* $\mathbf{b} \in \mathbb{Z}_p^n$*, the number of* $n \times n$ *invertible matrices* $\mathbf{K}$ *over* $\mathbb{Z}_p$ *such that* $\mathbf{Km} = \mathbf{b}$ *is exactly* $\frac{N}{p^n-1}$*, where* N *is the total number of* $n \times n$ *invertible matrices over* $\mathbb{Z}_p$*. In other words,* $Pr[\mathbf{Km} = \mathbf{b}] = \frac{1}{p^n-1}$ *where* $\mathbf{K}$ *is uniformly chosen from the set of* $n \times n$ *invertible matrices over* $\mathbb{Z}_p$.

Proof. If $\mathbf{b}$ is zero vector, the vector $\mathbf{m}$ must be zero since $\mathbf{K}$ is invertible. This contradicts the condition. That is, there is no $\mathbf{K}$ that satisfies $\mathbf{Km} = \mathbf{b}$. Hence, we only need to consider the case where $\mathbf{b}$ is nonzero.

Denote $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, where $m_i, b_j \in \mathbb{Z}_p$. Without loss of generality, assume that $b_1 \neq 0$, since otherwise we can simply permute the rows of $\mathbf{K}$ and $\mathbf{b}$ simultaneously such that $b_1 \neq 0$, without affecting the number of possible candidates for $\mathbf{K}$.

Since $\mathbf{m}$ is nonzero, from lemma 3, the number of vectors $\mathbf{k} \in \mathbb{Z}_\mathbf{p}^\mathbf{n}$ such that $\mathbf{k} \cdot \mathbf{m} = b_1$ is p^{n-1}. Due to $b_1 \neq 0$, all possible candidates for the first row $\mathbf{k_1}$ of $\mathbf{K}$ do not include the zero vector. So, all of the $\mathbf{k_1}$ obatined in this way must satisfy the requirement that matrix $\mathbf{K}$ is invertible.

Next we consider the second row $\mathbf{k_2}$ of $\mathbf{K}$. From lemma 3, the number of vectors $\mathbf{k}$ such that $\mathbf{k} \cdot \mathbf{m} = b_2$ is also p^{n-1}. However, these vectors no longer necessarily satisfy the requirement that matrix $\mathbf{K}$ be invertible. In order for the matrix to be invertible, the vector $\mathbf{k_2}$ has to be linearly independent of the vector $\mathbf{k_1}$. In other words, we have to subtract from all possible $\mathbf{k_2}$ the linear combinations of $\mathbf{k_1}$ that satisfy the equation $\mathbf{k} \cdot \mathbf{m} = b_2$. Let $\mathbf{v} = a\mathbf{k_1}$ be the possible linear combination of $\mathbf{k_1}$ with coefficient a. From $\mathbf{v} \cdot \mathbf{m} = b_2$, we have $\mathbf{v} \cdot \mathbf{m} = a\mathbf{k_1} \cdot \mathbf{m} = ab_1 = b_2$. Since $b_1 \neq 0$, $a = b_2 b_1^{-1} (mod p)$. That is, there is only one linear combination of $\mathbf{k_1}$ such that $\mathbf{v} \cdot \mathbf{m} = b_2$. Hence, the number of $\mathbf{k_2}$ is $p^{n-1} - 1$.

Similar to the case of the second row, now we consider the jth row $\mathbf{k_j}$ in general, where $3 \leq j \leq n$. From

lemma 3, the number of vectors $\mathbf{k}$ such that $\mathbf{k}\cdot\mathbf{m}=b_j$ is also p^{n-1}. However, these vectors no longer necessarily satisfy the requirement that matrix $\mathbf{K}$ be invertible. In order for the matrix to be invertible, it is required that $\mathbf{k_j}$ be not a linear combination of all the previous rows $\mathbf{k_1},\ldots,\mathbf{k_{j-1}}$. Let $\mathbf{v}=a_1\mathbf{k_1}+a_2\mathbf{k_2}+\cdots+a_{j-1}\mathbf{k_{j-1}}$ be the possible linear combination of $\mathbf{k_1},\ldots,\mathbf{k_{j-1}}$ with coefficients $a_1,a_2,\ldots,a_{j-1}$. From $\mathbf{v}\cdot\mathbf{m}=b_j$, we have $\mathbf{v}\cdot\mathbf{m}=a_1\mathbf{k_1}\cdot\mathbf{m}+a_2\mathbf{k_2}\cdot\mathbf{m}+\cdots+a_{j-1}\mathbf{k_{j-1}}\cdot\mathbf{m}=ab_1+ab_2+\cdots+ab_{j-1}=b_j$. Since $b_1\neq 0$, from lemma 3, the number of coefficients vector $(a_1,a_2,\ldots,a_n)$ such that $\mathbf{v}\cdot\mathbf{m}=b_j$ is p^{j-2}. Therefore, the number of possible vectors for $\mathbf{k_j}$ such that $\mathbf{k_j}$ is not a linear combination of $\mathbf{k_1},\mathbf{k_2}\ldots,\mathbf{k_{j-1}}$ and that $\mathbf{k_j}\cdot\mathbf{m}=b_j$ is $p^{n-1}-p^{j-2}$.

Let N_k be the total number of possible matrices for $\mathbf{K}$ such that $\mathbf{Km}=\mathbf{b}$, from above analyses, we have $N_k=p^{n-1}(p^{n-1}-1)(p^{n-1}-p)\ldots(p^{n-1}-p^{n-2})$. In addition, the total number of $n\times n$ invertible matrices over $\mathbb{Z}_p$ is $N=(p^n-1)(p^n-p)\ldots(p^n-p^{n-1})$. Hence, $N_k=\frac{N}{p^n-1}$ as claimed. When $\mathbf{K}$ is uniformly chosen from the set of $n\times n$ invertible matrices over $\mathbb{Z}_p$, the probability $Pr[\mathbf{Km}=\mathbf{b}]=\frac{N_k}{N}=\frac{1}{p^n-1}$. □

Now based on lemmas above, we give the proof of theorem 2.

Proof. (Theorem 2) Let $\mathbf{m}',\mathbf{m}''\in\mathbb{Z}_\mathbf{p}^\mathbf{nr}$ be distinct messages and $\mathbf{b}\in\mathbb{Z}_\mathbf{p}^\mathbf{n}$ be any hash value. Denote $\mathbf{m_\Delta}=\mathbf{m}'-\mathbf{m}''=(\mathbf{m_1},\mathbf{m_2},\ldots,\mathbf{m_r})$, where $\mathbf{m_i}(1\leq i\leq n)$ are n-length blocks. Obviously, $\mathbf{m_\Delta}\neq\mathbf{0}$.

Now, we present the proof by mathematical induction.

For $r=1$, from lemma 4, we have $Pr[H(\mathbf{K},\mathbf{m}')-\mathbf{H}(\mathbf{K},\mathbf{m}'')=\mathbf{b}]=Pr[H(\mathbf{K},\mathbf{m_\Delta})=\mathbf{b}]\leq\frac{1}{p^n-1}$. Hence, hash function H is $\frac{1}{p^n-1}$-almost-Δ-universal.

Next, assuming that H is $\frac{1}{p^n-1}$-almost-Δ-universal for $r\geq 1$, we show that H for $r+1$ is the same.

$$\begin{aligned}
&H(\mathbf{K},\mathbf{m}')-\mathbf{H}(\mathbf{K},\mathbf{m}'')=\mathbf{b}\\
\Longleftrightarrow\ &H(\mathbf{K},\mathbf{m_\Delta})=\mathbf{b}\\
\Longleftrightarrow\ &\mathbf{K^{r+1}m_1}+\mathbf{K^r m_2}+\cdots+\mathbf{Km_{r+1}}=\mathbf{b}\\
\Longleftrightarrow\ &\begin{cases}\mathbf{K^r m_1}+\mathbf{K^{r-1}m_2}+\cdots+\mathbf{Km_r} &=\mathbf{c}\\ \mathbf{Kc} &=\mathbf{b}\end{cases}
\end{aligned}$$

Now, we just have to count the number of matrices $\mathbf{K}$ that satisfy the last system.

Case 1: $\mathbf{b}=\mathbf{0}$.

If $\mathbf{c}=\mathbf{0}$, the second equation is a identity equation which has no constraint on $\mathbf{K}$; the first equation has been reduced to the assumption that H is $\frac{1}{p^n-1}$-almost-Δ-universal for r .

If $\mathbf{c}\neq\mathbf{0}$, the second equation is a contradictory equation. Therefore, no invertible key matrix $\mathbf{K}$ satisfies it. Hence, H is Δ-universal in this case.

Case 2: $\mathbf{b}\neq\mathbf{0}$.

If $\mathbf{c}=\mathbf{0}$, the second equation is a contradictory equation. Therefore, no invertible key matrix $\mathbf{K}$ satisfies it. Hence, H is Δ-universal in this case.

If $\mathbf{c}\neq\mathbf{0}$, from lemma 4, the number of possible $\mathbf{K}$ that satisfies the second equation is $\frac{N}{p^n-1}$, where N is the total number of invertible $n\times n$ invertible matrices over $\mathbb{Z}_p$. Hence, the number of possible $\mathbf{K}$ that satisfies both equations is at most $\frac{N}{p^n-1}$. That is, $Pr[H(\mathbf{K},\mathbf{m_\Delta})=\mathbf{b}]\leq\frac{1}{p^n-1}$. H is $\frac{1}{p^n-1}$-almost-Δ-universal. □

Corollary 1. *The hash function defined in Definition 6 is a $\frac{1}{p^n-1}$-balance hash function.*

Proof. The proof is easily derived from theorem 2, lemma 1 and 2, omitted here. □

Remark 3. The proofs of lemma 1-4 and theorem 2 only use the property of $\mathbb{Z}_p$ as a finite field, and does not involve the more specific property of $\mathbb{Z}_p$. Thus, we can extend the number field $\mathbb{Z}_p$ to a general finite field $\mathbb{F}_q$ where q is not necessarily a prime number, or more accurately, $q=p^l$ is a power of some prime number. In this way, the scope of the kind of hash functions has been greatly extended, and will be more widely used.

Remark 4. A similar hash function has been proposed in [15]. However, the underlying number field is just binary field $\mathbb{Z}_2$ there, not the pime fields $\mathbb{Z}_p$ or the more general finite field $\mathbb{F}_q$ as we have used here. Furthermore, the proof that the hash function is ϵ-universal and ϵ-balance is wrong. Specifically, in the last several rows of the proof of lemma 2 in [15], the construction of w_r depends on $\mathbf{K}$, which results in the probability $Pr[H_r(\mathbf{K},\mathbf{x})=H_r(\mathbf{K},\mathbf{w})]$ no longer the probability required in the definition of universal hash function and thus not illustrative.

4 Conclusions

In this paper, we first generalize the classical polynomial hash functions to a family of matrix polynomial hash functions. Then, in order to prove the proposed hash family is ϵ-almost-Δ-universal, we invent a technique that recursively transforms a matrix polynomial into a system of linear matrix equations. The technique may be also useful in dealing with other problems about matrix polynomials.

Acknowledgments

The author gratefully acknowledges the anonymous reviewers for their valuable comments.

References

[1] D. J. Bernstein, "The poly1305-aes message-authentication code," in *Proceedings of Fast Software Encryption (FSE'05)*, pp. 32–49, 2005.

[2] K. Bibak, *Restricted Conguences in Computing*. CRC Press, 2020.

[3] K. Bibak, "Quantum key distribution using universal hash functions over finite fields," *Quantum Inf. Process*, vol. 21, no. 121, 2022.
[4] K. Bibak, B. M. Kapron, and V. Srinivasan, "Authentication of variable length messages in quantum key distribution," *EPJ Quantum Technol.*, vol. 9, no. 8, 2022.
[5] K. Bibak and R. Ritchie, "Quantum key distribution with prf(hash, nonce) achieves everlasting security," *Quantum Inf. Process*, vol. 20, no. 228, 2021.
[6] K. Bibak, R. Ritchie, and B. Zolfaghari, "Everlasting security of quantum key distribution with 1k-dwcdm and quadratic hash," *Quantum Inf. Comput.*, vol. 21, no. 3&4, pp. 181–202, 2021.
[7] C. Boyd, A. Mathuria, and D. Stebila, *Protocols for authentication and key establishment.* Berlin: Springer, 2020.
[8] D. Bruss, G. Erdélyi, T. Riege, and J. Rothe, "Quantum cryptography: a survey," *ACM Comput. Surv.*, vol. 39, no. 2, p. 6, 2007.
[9] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *J. Comput. Syst. Sci.*, vol. 18, no. 2, pp. 143–154, 1979.
[10] C. C. Chang, M. S. Hwang, "Parallel computation of the generating keys for RSA cryptosystems", *Electronics Letters*, vo1. 32, no. 15, pp. 1365–1366, 1996.
[11] T. Y. Chang, M. S. Hwang, W. P. Yang, "A new multi-stage secret sharing scheme using one-way function", *ACM SIGOPS Operating Systems Review*, vol. 39. no. 1, pp. 48-55, 2005.
[12] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.
[13] F. Grasselli, H. Kampermann, and D. Bruß, "Conference key agreement with single-photon interference," *New J. Phys.*, vol. 21, no. 123002, 2019.
[14] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245-255, 2003.
[15] Q. Li and S. Sovio, "A fast hash family for memory integrity," *IACR Cryptology ePrint Archive*, vol. 2022, no. 1378, 2022.
[16] C. Portmann, "Quantum authentication with key recycling," in *Proceedings of the 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT 2017)*, pp. 339–368, Paris, France, 2017.
[17] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.
[18] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proceedings of 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, Santa Fe, NM, November 1994.
[19] M. Tomamichel and A. Leverrier, "A largely self-contained and complete security proof for quantum key distribution," *Quantum*, vol. 1, no. 14, 2017.

Biography

Shuying Yang received the B.S. and M.S. degree in Mathematics from Shandong Normal University, China, in 2004 and 2007, respectively. Currently, She is an associate professor in school of data and computer science at Shandong Women's University. Her research interests include information security and cryptology. Prof. Yang may be reached at ysystudy2005@163.com.

Data Encryption by AES: Security Guarantee for Network Communication Sensitive Information

Jizhou Shan and Hong Ma
(Corresponding author: Jizhou Shan)

Hainan College of Economics and Business
Haikou, Hainan 571127, China
Email: shanji9125180@yeah.net

(Received Sept. 4, 2022; Revised and Accepted July 28, 2023; First Online Oct. 11, 2023)

Abstract

Encryption of data is an effective means of securing network communication. This paper briefly introduced the advanced encryption standard (AES) algorithm. It utilized the Rivest, Shamir, and Adleman (RSA) algorithm for encrypting the AES key to enhance the encryption process's efficiency. Then, the encryption and decryption efficiency, transmission efficiency, and security of the RSA, AES, and RSA+AES algorithms were compared using the simulation experiments. It was found that the increase in the size of transmission files increased the encryption and decryption time and reduced the average transmission rate. In terms of security, all three encryption algorithms had encryption sensitivity, and the RSA+AES algorithm had the best security performance in encrypting sensitive information.

Keywords: Advanced Encryption Standard; Information Protection; Network Communication; RSA Cryptosystem

1 Introduction

With the advancement and widespread use of computers and the Internet, communication between individuals has transcended traditional face-to-face interactions and letter writing [15]. The Internet enables seamless sending and receiving of emails, and real-time communication has become easily achievable due to the improvement in Internet transmission speeds. The Internet has significantly facilitated interpersonal communication [17]. However, the openness of the Internet also brings forth the risk of information theft or tampering during data transmission. Thus, safeguarding sensitive information during network communication has become a critical concern. Encryption of communication data is a widely adopted method to protect sensitive information, as it uses a key to transform plaintext into ciphertext, ensuring that even if intercepted, the data remains challenging to decipher [13, 14].

Several related studies have explored encryption techniques to enhance security. For instance, by using the advanced encryption standard (AES) for data compression and increasing the number of rounds in the AES encryption and decryption operations to 16, Kumar *et al.* [11] improved the security of a system. Yang *et al.* [22] developed an enhanced AES encryption algorithm based on chaos theory to address security concerns. They verified its feasibility and security through simulations. Albahar *et al.* [2] introduced a new triple algorithm combining Rivest, Shamir, and Adleman (RSA) [5, 8], AES, and TwoFish to further enhance the security of Bluetooth [9, 21], whose latest versions solely rely on 128-bit AES encryption.

Experimental data confirmed that the new algorithm eliminated all known weaknesses to improve the Bluetooth's encryption security. In this context, this paper offers a concise introduction to the AES encryption algorithm and suggests incorporating the RSA algorithm to encrypt the AES key, thus optimizing the encryption process. Subsequently, the study compares the encryption and decryption efficiencies, data transmission efficiencies, and security of three encryption algorithms, namely RSA, AES, and RSA+AES, through simulation experiments.

2 AES-based Encryption for Network Communication

In asymmetric cryptography, the public key is made available to the public, and anyone can access it to perform encryption [6, 10, 12, 20]. However, the private key is kept confidential by the individual, making it more secure. Symmetric encryption technology remains one of the commonly used mainstream cryptographic approaches due to its advantages of fast encryption, strong scalability, flexibility, and compatibility [19].

The fundamental process of the symmetric encryption algorithm AES comprises three main parts: key expan-

sion, encryption, and decryption. During the encryption process, the data is divided into several blocks, each consisting of four bytes. The initial round involves performing a bitwise exclusive OR (XOR) operation on the plaintext and the first round key. Subsequently, following the AES encryption algorithm rules, several rounds of iterations are conducted using the output of the initial round as input, culminating in the final ciphertext after the last round of iteration [7]. The decryption process follows a similar procedure to encryption but utilizes the round keys in reverse order. The AES algorithm is more efficient in both encryption and decryption processes; however, it also poses a higher risk of key leakage [3]. While the RSA algorithm provides greater security, its longer key length leads to decreased encryption efficiency. Hence, this article proposes combining AES and RSA by utilizing RSA for encrypting the AES keys. This approach not only mitigates the risk of leaking AES keys but also prevents unnecessary prolongation of the encryption process.

The process when the RSA and AES algorithms are combined is shown in Figure 1.

1) An AES key with 128 bits is used to encrypt the plaintext. The basic principle is as follows. A 128-bit plaintext and the key are put into a matrix of 4×4 respectively. Then, the key matrix and the plaintext matrix are subjected to the XOR operation [4]. After that, the S-box of the fixed table of the AES algorithm is utilized for byte replacement in the matrix. After the byte replacement, the matrix shifts the elements of each row to the left in a circular manner based on the specified displacement amount, which is called row shifting. Then, the column mixing operation is performed. The above procedure [16] is cycled nine times, and the column mixing is replaced by XOR operation at the tenth cycle.

2) The RSA public key is employed to encrypt the AES key. The encryption formula is:

$$\begin{cases} C = m^e \bmod n \\ \gcd[e, \phi(n)] = 1 \\ \phi(n) = (p-1)(q-1) \\ n = pq \end{cases} \tag{1}$$

where C is the ciphertext, m is the plaintext, e is the public key, p and q are two confidential prime numbers with random 1024 bits, and n is the product of p and q [18], which can be made public.

3) The communication ciphertext is combined with the ciphertext of the AES key at the communication sender, and then the combination is sent to the communication receiver.

4) The communication receiver splits the received combined ciphertext into communication ciphertext and AES key ciphertext.

5) The AES key ciphertext is decrypted using the RSA private key. The decryption formula is:

$$\begin{cases} m = c^d \bmod n \\ ed = 1 \bmod \phi(n) \end{cases} \tag{2}$$

where d is the private key, which is computed by the public key and is not public.

6) The decryption process uses the round keys in reverse order.

3 Simulation Experiment

3.1 Experimental Environment

The simulation experiments were conducted in a server in the lab, where three servers were set up. Server 1 served as the communication sender, server 2 served as the communication receiver, and server 3 acted as a third party to capture the communication data.

3.2 Experimental Setup

In the simulation experiment, servers 1 and 2 were used to establish a server-client architecture. Server 1 initiated communication while server 2 received it. The basic process is as follows. Firstly, servers 1 and 2 performed a handshake [23] to establish a unified AES key for encryption, and then they communicated with each other. Server 1 read the file that needs to be sent and encrypted it using an AES key; then, RSA encrypted the key and sent both the encrypted file and key to server 2. Finally, server 2 received the ciphertext and decrypted it.

3.3 Experimental Projects

1) Data encryption and decryption efficiency
Files with sizes of 100, 200, 300, 400, and 500 MB were set to test the encryption and decryption time. Additionally, the encryption and decryption time of both AES and RSA schemes was also tested.

2) Data transmission rate test
Files with sizes of 100, 200, 300, 400, and 500 MB were set up respectively, and they were encrypted and sent to the client according to the communication flow described above. The average transmission rate was calculated from the time when the server started sending the file to the time when the client finished receiving the file. In addition, the average transmission rate under the AES and RSA schemes was also tested.

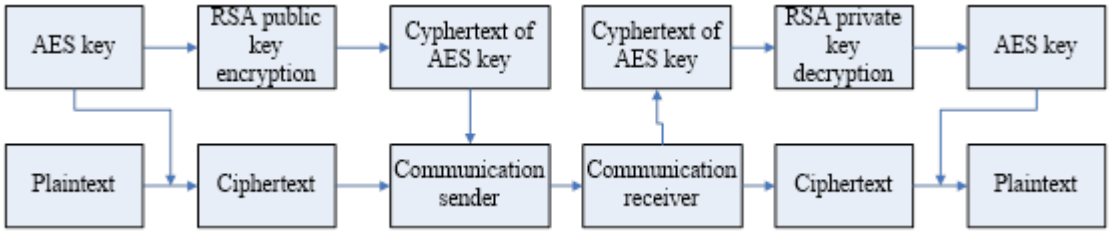


Figure 1: Communication encryption flow of the improved AES algorithm

3) Security of data encryption
In the process of communication between the server and client, the third-party server used the wiershark packet capture tool to simulate the listening condition and captured the communication data. In this test, the communication content was "hello", "hella", "nello", and the data captured by the packet capture tool was compared with the sent data.

4) Encryption effect on sensitive data
When encrypting communication data, sensitive information within it is also encrypted. Sensitive information includes the name, address and contact information of the communication user, so the encryption effect of the encryption scheme on the sensitive information was also tested. Firstly, 30 names, addresses and contact information were randomly generated, and then the sensitive information was transmitted under the three encryption schemes. The communication data was captured using a packet capture tool and violently decrypted for 60 min, and the decrypted text was compared with the original text to calculate the average completeness.

3.4 Experimental Results

The time required for encryption and decryption of the three encryption schemes is provided in Table 1. Comparing the encryption and decryption time consumed by the three encryption schemes under the same file size, it can be seen that the time consumed by the RSA algorithm was the most, the time consumed by the AES algorithm was the least, and the combination of the RSA and AES algorithms was slightly higher than the AES scheme. The RSA algorithm is a type of encryption algorithm that employs a public key to encrypt data and a corresponding private key for decryption purposes. The calculation method used in encryption and decryption is different, and moreover the calculation process both involves exponential operations. Therefore, it is the most time-consuming. The AES algorithm uses a key to encrypt and decrypt. The process of decrypting is the opposite of the process of encrypting, and the arithmetic process does not involve complex calculations. Therefore, it takes the least time. The RSA + AES algorithm encrypts plaintexts with the AES algorithm in nature, and the RSA algorithm is only applied to protect the AES key; therefore, the encryption and decryption by the combination algorithm is still faster than the RSA algorithm. However, encrypting the AES key with the RSA algorithm will take some time. As the length of the AES key is much smaller than the plaintext, the impact is not very large.

As can be seen from Figure 2, as the file size grew larger, the time for encryption at the server side and decryption at the receiver side in the communication process increased, making the average transmission rate decrease. Under the same file size, the average transmission rate of the RSA scheme was the smallest, the average rate of the AES scheme was the largest, and the average rate of the RSA+AES scheme was slightly lower than that of the AES scheme. The reason also lies in the difference of encryption efficiency of the encryption scheme. The RSA algorithm takes the longest time for encryption, which makes the average transmission rate low. The AES and combination algorithms take shorter time in encrypting and decrypting data, so the average transmission rate is high. Compared to the AES algorithm, the encryption and decryption of the combination algorithm take a little bit more time, which makes its average transmission efficiency slightly lower.

Under the three encryption schemes, the packet capture program was used to capture the transmitted data during the communication process. The captured data content, plaintext, and average completeness after brute-force decryption are demonstrated in Table 2. From Table 2, it is evident that the data bytes encrypted by the encryption algorithm were significantly different from the plaintext. In the same encryption algorithm, the difference of only one letter in the plaintext could make a significant difference in the ciphertext, and all three encryption algorithms had encryption sensitivity.

The cracking completeness of the three sensitive data such as name, address, and contact information after 60 min of brute force cracking under the three encryption schemes is displayed in Table 3. From Table 3, it can be seen that the average cracking completeness of the three sensitive data under the same encryption scheme did not differ much, and the comprehensive cracking completeness of the AES algorithm was the highest, the RSA algorithm was the second, and the RSA+AES algorithm was the lowest.

Table 1: The encryption and decryption time of three encryption schemes

Encryption scheme	100 MB	200 MB	300 MB	400 MB	500 MB
Time consumption of encryption and decryption by RSA/ms	740	950	1140	1320	1530
Time consumption of encryption and decryption by AES/ms	520	680	870	1020	1130
Time consumption of encryption and decryption by RSA+AES	580	730	890	1070	1190

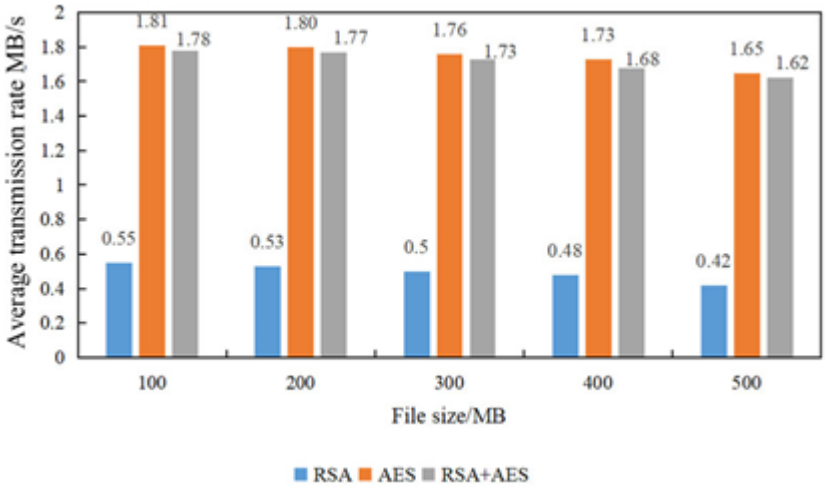


Figure 2: Average transmission rate of files of different sizes under three encryption schemes

Table 2: Data content captured under three encryption schemes

Encryption algorithm	Plaintext	Plaintext hexadecimal bytes	Captured data
RSA	hello	68656C6C6F	13551364da64d6355233
	hella	68656C6C61	20d5366d2gde2141da23
	nello	6E656C6C6F	635dg2e5f62314dde212
AES	hello	68656C6C6F	587a55a21ef56ac5dd54
	hella	68656C6C61	98a85d65ec6f54a55723
	nello	6E656C6C6F	7456fcea544f4dc44ef4d
RSA+AES	hello	68656C6C6F	3484ada8818461da4154
	hella	68656C6C61	7dfa2a45d5646654da11
	nello	6E656C6C6F	369a5f32587df1256564

Table 3: Brute force decryption completeness for different sensitive data types under three encryption schemes

Encryption scheme	Types of Sensitive Data	Average Cracking Completeness/%	Comprehensive Cracking Completeness/%
RSA	Name	3.2	3.2
	Address	3.1	
	Contact information	3.3	
AES	Name	4.2	4.1
	Address	4.1	
	Contact information	4.1	
RSA+AES	Name	2.1	2.1
	Address	2.2	
	Contact information	2.1	

4 Conclusion

This paper offers a brief overview of the AES encryption algorithm and incorporated the RSA algorithm to encrypt the AES key, aiming to enhance the encryption algorithm's performance. Subsequently, it conducted simulation experiments to compare the encryption and decryption efficiency, transmission efficiency, and security of three encryption algorithms: RSA, AES, and RSA+AES. The results are summarized as follows.

1) With the increase of the file size in the transmission, the encryption and decryption time for all three encryption schemes rose. When the file size was the same, RSA encryption and decryption required the most time, whereas AES encryption and decryption required the least amount of time. The RSA+AES algorithm showed slightly higher time consumption compared to the AES algorithm for the same file size.

2) As the number of transmitted files increased, the average transmission rate of the three encryption schemes showed a decreasing trend. Under the same file size, the RSA scheme had the smallest average transmission rate, while the AES scheme exhibited the largest average transmission rate. The average transmission rate of the RSA+AES scheme was slightly lower than that of the AES scheme.

3) All three encryption algorithms were cryptographically sensitive, meaning that even a minor change in the plaintext can result in a substantial change in the ciphertext.

4) Regarding the average completeness after 60 minutes of brute force decryption attempts, the AES encrypted ciphertext achieved the highest completeness upon decryption, followed by the RSA algorithm. The RSA+AES algorithm had the lowest completeness level.

References

[1] F. A. Abdulatif, M. Zuhiar, "Improve security of cloud storage by using third parity authentication, one time password and modified AES encryption algorithm," *International Journal of Informatics and Communication Technology*, vol. 7, no. 1, pp. 24-30, 2018.

[2] M. A. Albahar, O. Olawumi, K. Haataja, P. Toivanen, "Novel hybrid encryption algorithm based on AES, RSA, and Twofish for bluetooth encryption," *Information Security*, vol. 9, no. 2, pp. 168-176, 2018.

[3] N. Attar, H. Deldari, M. Kalantari, "AES encryption algorithm parallelization in order to use big data cloud Naser Attar, Hossein Deldari, Marzie Kalantari," *Computer and Information Science*, vol. 10, no. 3, pp. 23-28, 2017.

[4] A. Banushri, R. A. Karthika, "A survey on data security using file hierarchy attribute-based encryption in cloud computing environment," *Journal of Advanced Research in Dynamical & Control Systems*, vol. 2017, no. 4, pp. 144-149, 2017.

[5] C. C. Chang, M. S. Hwang, "Parallel computation of the generating keys for RSA cryptosystems", *Electronics Letters*, vo1. 32, no. 15, pp. 1365–1366, 1996.

[6] Y. Feng, Y. Liu, G. Zhao, M. Xia, "An improved AES encryption algorithm based on the Hénon and Chebyshev chaotic map," *International Journal of Simulation: Systems, Science & Technology*, vol. 17, no. 48, pp. 25.1-25.8, 2016.

[7] R. Hamamreh, E. Tabib, "Selective image compression-encryption algorithm using adaptive Huffman coding and AES," *Journal of Applied Information Technology*, vol. 99, no. 4, pp. 932-945, 2021.

[8] M. S. Hwang, K. F. Hwang, I. C. Lin, "Cryptanalysis of the batch verifying multiple RSA digital signatures", *Informatica*, vol. 11, no. 1, pp. 15-18, Jan. 2000.

[9] M. S. Hwang, C. C. Lee, J. Z. Lee, C. C. Yang, "A secure protocol for bluetooth piconets using el-

liptic curve cryptography", *Telecommunication Systems*, vol. 29, no. 3, pp. 165-180, 2005.

[10] M. S. Hwang, S. F. Tzeng, C. S. Tsai, "Generalization of proxy signature based on elliptic curves," *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 73-84, 2004.

[11] P. Kumar, S. B. Rana, "Development of modified AES algorithm for data security," *International Journal for Light and Electron Optics*, vol. 127, no. 4, pp. 2341-2345, 2016.

[12] C. C. Lee, M. S. Hwang, L. H. Li, "A new key authentication scheme based on discrete logarithms", *Applied Mathematics and Computation*, vol. 139, no. 2, pp. 343-349, July 2003.

[13] C. C. Lee, H. C. Tseng, C. C. Liu, H. J. Chou, "Using AES encryption algorithm to optimize high-tech intelligent platform," *WSEAS Transactions on Business and Economics*, vol. 18, pp. 1572-1579, 2021.

[14] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245-255, 2003.

[15] A. Mersaid, T. Gulom, "The encryption algorithm AES-RFWKIDEA32-1 based on network RFWKIDEA32-1," *Information Engineering*, vol. 4, no. 1, pp. 1-11, 2016.

[16] P. R. More, S. Y. Gaikwad, "An advanced mechanism for secure data sharing in cloud computing using revocable storage identity based encryption," *International Journal of Engineering Business Management*, vol. 1, no. 1, pp. 12-14, 2017.

[17] S. Oukili, S. Bri, "Hardware implementation of AES algorithm with logic S-box," *Journal of Circuits Systems & Computers*, vol. 26, no. 9, pp. 1-19, 2017.

[18] T. Paka, S. Divya, "Data storage security and privacy in mobile cloud computing using hierarchical attribute based encryption (HABE)," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 6, pp. 750-754, 2019.

[19] N. Rachmat, Samsuryadi, "Performance analysis of 256-bit AES encryption algorithm on Android smartphone," *Journal of Physics: Conference Series*, vol. 1196, 2019.

[20] S. F. Tzeng, M. S. Hwang, "Digital signature with message recovery and its variants based on elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 61-71, Mar. 2004.

[21] H. W. Yang, J. Z. Lee, M. S. Hwang, "A taxonomy of bluetooth security", *International Journal of Electronics and Information Engineering*, vol. 12, no. 2, pp. 43-65, 2020.

[22] Z. H. Yang, A. H. Li, L. L. Yu, S. J. Kang, M. J. Han, Q. Ding, "An improved AES encryption algorithm based on chaos theory in wireless communication networks," in *Third International Conference on Robot, Vision and Signal Processing (RVSP'15)*, pp. 159-162, 2015.

[23] H. Yue, X. Zheng, "Research on encrypting accounting data using des algorithm under the background of microprocessor system," *Microprocessors and Microsystems*, vol. 104061, 2021.

Biography

Yang Daojing, born in 1982, graduated from Guilin university of electronic technology in July 2006 with a Master's degree. She is an associate professor and is now working in Chengdu Jincheng College, China.She is interested in applications of data mining.

A New Secure Channel Free Public Key Encryption with Keyword Search Scheme Based on ElGamal Cryptosystems

Min-Shiang Hwang[1,2], Shih-Ting Hsu[3], and Cheng-Ying Yang[4]
(Corresponding author: Cheng-Ying Yang)

Department of Computer Science and Information Engineering, Asia University[1]
Fintech and Blockchain Research Center, Asia University[2]
Wufeng, Taichung 354, Taiwan (R.O.C.)
Department of Management Information Systems, National Chung Hsing University, Taiwan[3]
Department of Computer Science, University of Taipei[4]
Taipei, Taiwan (R.O.C.)
Email: cyang@utaipei.edu.tw
(Received May 11, 2023; Revised and Accepted Oct. 20, 2023; First Online Oct. 31, 2023)

Abstract

Public key encryption with keyword search (PEKS) scheme enables one to send the trapdoor, which involves the encrypted keyword in querying data without revealing the keyword. The trapdoor should transfer into the secret channel, which is costly and inefficient. Many articles proposed efficient, secure, channel-free public key encryption with keyword search (SCF-PEKS) schemes. After that, many schemes focus on "against the offline keyword guessing attack (OKGA)" by enhancing the model. However, most PEKS/SCF-PEKS schemes are constructed with bilinear pairing. However, the PEKS/SCF-PEKS based on bilinear pairing is susceptible to offline keyword-guessing attacks. In this article, we will propose a new SCF-PEKS based on ElGamal cryptography. The scheme is secure against offline keyword guessing attacks and improves efficiency.

Keywords: Designate Tester; ECC; ElGamal; Off-line Keyword Guessing; PEKS

1 Introduction

Since cloud computing has become the most popular issue in recent years, more and more Cloud services such as storage space, computing resources, and various software have been widely used worldwide. When people gradually use Cloud storage servers as the daily data storage space and replace hard discs in desktop computers, the problem of Cloud security has become an essential issue in the recent study. Moreover, people usually respect privacy; for example, when a user wants to search specific data in the Cloud with keywords, they do not want the query contents to be known by others, including the cloud systems [5, 13].

Public key encryption with keyword search (PEKS) scheme, the first proposed by Boneh *et al.* [2] in 2004, enables people to query data without revealing the keyword. PEKS scheme provides one of the alternatives for mail routing systems to route emails. For example, when a mail server receives mail with specific keywords in the title, such as "urgent" or "emergency," the server will route those emails to the receiver's mobile or tablet PC, and the receiver can read those emails immediately. The other emails will route to the receiver's desktop computer and read it later.

The PEKS scheme involves three roles: sender, receiver, and server. To avoid revealing the keywords, the sender encrypted them and appended them to the mail before uploading them to the mail server. Then, the server stores the emails and encrypted keywords. When the receiver wants to search specific emails related to the keyword w, they can send a trapdoor T_w, which involves the keyword w, to the server and get the corresponding emails after retrieving the encrypted keyword w of all the emails by the server. However, Boneh *et al.*'s scheme requires constructing a secure channel between the server and receiver for sending trapdoors. Therefore, it is unsuitable for some applications because the building cost is high [1].

In 2008, Baek *et al.* [1] proposed an efficient, secure, channel-free public key encryption with keyword search (SCF-PEKS) scheme that removes the secure channel assumption in the original PEKS scheme of Boneh *et al.* In Baek *et al.*'s scheme, only the designated server can test whether the trapdoor contains the specific keyword in the server so that the trapdoors can transfer to the public network. However, Byun *et al.* [3] pointed out

that [2] may be attacked by an offline keyword guessing attack. The attackers can capture trapdoors and have a chance to guess keywords because of the low entropy of keywords for searching documents. Keywords are chosen from much smaller spaces than passwords. In [1], trapdoors are transferred in the public network, and that means anyone can capture trapdoors. Yau *et al.* [28] demonstrated that outsider adversaries that capture the trapdoors sent in a public channel could reveal encrypted keywords by performing offline keyword guessing attacks.

In 2009 and 2010, Rhee *et al.* [19, 20] defined two criteria for the SCF-PEKS scheme: ciphertext indistinguishability and trapdoor indistinguishability. Furthermore, they constructed an SCF-PEKS scheme that satisfies these criteria. Fang *et al.* [6] think the random oracle model can only guarantee the existing schemes' security limitations. But, unfortunately, proof in the random oracle model has been shown to possibly not secure in the standard model [4]. Therefore, Fang *et al.* proposed a new and efficient scheme that does not require any secure channels, and its security does not use random oracles. Recently, Hu and Liu's scheme constructed an enhanced SCF-PEKS and extended the proposed scheme to secure decryptable encryption with a designated tester [11]. All the PEKS/SCF-PEKS schemes mentioned above are constructed in bilinear forms, susceptible to offline keyword guessing attacks [27]. Although both [20] and [11] can successfully against keyword-guessing attacks by outside attackers, they still can not resist the inside attack.

In 2010, Gu and Zhu proposed a PEKS scheme based on bilinear pairing [9]. No pairing operation is involved in the encryption process, so their PEKS scheme is efficient. In 2013, Hsu *et al.* outlined six existing security models of PEKS/SCF-PEKS schemes and summarized five security requirements that must be met to construct a secure PEKS/SCF-PEKS scheme [10, 14]. In 2015, Wang *et al.* proposed an improved decentralized fault-tolerant keyword search (DFKSSVS) scheme that supports verifiable search capabilities in the hybrid cloud. By improving the dictionary-based keyword creation scheme, fuzzy keyword sets are generated, and secure indexes are created to achieve efficient fuzzy search [22].

In 2016, Wang *et al.* proposed an ElGamal encryption scheme with fuzzy keyword search, which has the following advantages [24]: 1). External attackers cannot obtain any keyword-related information without knowing the server's private key. 2). It not only supports accurate keyword search encryption but also supports searches when the entered keywords have spelling errors or inconsistent formats, which significantly improves the usability of the system. 3). This scheme is built based on El Gamal encryption rather than bilinear pairing encryption, which significantly improves computing efficiency.

In 2017, Thiyagarajan and Ganesan proposed an architecture for multi-keyword search by using Bloom filters to create indexes and generate key pairs through a pseudo-random bit generator [21]. The Bloom filter's search time on large encrypted file systems is O(N) without decrypting the document, thus speeding up the process of user-side ciphertext retrieval. To achieve greater efficiency and security in information retrieval, Zou *et al.* introduced the concept of decentralized searchable asymmetric encryption in 2018, which is helpful for security and enables search operations on encrypted data [30].

In 2019, Hu *et al.* proposed a secure and efficient ranking keyword search for outsourced cloud data based on chaotic arithmetic coding and obfuscation [12]. Liu and Fan proposed an ingenious signcryption search scheme based on key policy attributes and a specific searchable attribute-based authentication encryption scheme [15]. In 2019, Feng *et al.* proposed a searchable CP-ABE privacy protection scheme that supports users to revoke it directly [8].

In 2020, Wang *et al.* proposed an improved homomorphic encryption method for multi-keyword retrieval [23]. Their scheme can effectively solve the privacy leakage problem of search keywords. In 2020, Liu *et al.* proposed an encryption scheme based on the ciphertext policy attribute, supporting data retrieval, result verification, and attribute revocation [16]. They use BLS signature technology to achieve attribute-based keyword search encrypted result verification.

In 2021, Liu *et al.* proposed a novel attribute-based IoT encryption scheme that can provide user revocation, multi-keyword search, and data integrity verification [17]. In 2021, Feng *et al.* proposed searchable encryption and proxy re-encryption [7]. They proposed a blockchain data-sharing scheme based on searchable proxy re-encryption, which solved the problems of conventional blockchain data privacy leakage and rapid query. In 2021, Wu *et al.* proposed a searchable encryption based on ciphertext policy attributes to eliminate these shortcomings [26]. Their scheme supports large attribute domains and multi-keyword searches in electronic medical record sharing. In 2021, Ren *et al.* proposed a decentralized multi-authority ABSE scheme based on access trees, which can resist keyword guessing (KG) attacks and key recovery attacks [18].

In 2022, Wang *et al.* proposed a forward-safe joint keyword search scheme for result-free pattern leakage [25]. Furthermore, they proposed a baseline joint keyword FSSE scheme by combining the resulting pattern-hidden joint keyword search scheme (HXT) with a single keyword FSSE.

To construct an efficient SCF-PEKS scheme, we propose a new SCF-PEKS scheme based on ElGamal cryptography. The following are some criteria that query hiding needs to achieve in a cloud environment:

1) Unforgeable of the trapdoor [29]: The receiver generates his trapdoor information based on his keyword and secret key. Others can get nothing from the trapdoor and cannot produce the same trapdoor.

2) Anonymous of the ciphertext [29]: No one could get the embedded keywords from the ciphertext. After the keyword is encrypted, the information is safe

without the private key.

3) Authorized identity protection: The data sender can produce the keyword ciphertext and authorize a specific user to search the data. Others cannot learn the authorized users from the keyword ciphertext.

4) User authentication: Although no one can know the authorized users' identities, the server still has to recognize whether the authorized user uploaded the trapdoor. Hence, the server should be able to authenticate the user's identity.

The paper is organized as follows: Section 2 reviews the concept of some relative works. In Section 3, we perform an offline keyword guessing attack on two SCF-PEKS schemes based on bilinear form. In Section 4, we list our proposed scheme. Then, discuss and analyze the security and performance comparison of the proposed scheme in Section 5. The conclusion is in Section 6.

2 Related Works

In this section, we first review the concept of a bilinear map and how PEKS/SCF-PEKS use it. Then, review the definition of ElGamal.

2.1 Bilinear Pairing

Let G_1 and G_2 be multiplicative cyclic groups of prime order p, and g be a generator of G_1. We say $e : G_1 \times G_1 \to G_2$ is a bilinear map if the following conditions hold.

1) $e(S^a, V^b) = e(S, V)^{ab}$ for all $a, b \in Z_p$ and $S, V \in G_1$.

2) $(g, g) \neq 1$.

3) There is an efficient algorithm to compute $e(S, V)$ for all $S, V \in G_1$.

2.2 Public Key Encryption with Keyword Search (PEKS) [2]

Global Parameter: G_1 and G_2 are two cyclic groups and $e : G_1 \times G_1 \to G_2$ is a bilinear map. g and h are two generator of G_1 and $e(g, g)$ is a generator of G_2. There are two hash function $H_1 : \{0,1\}^* \to G_1$ and $H_2 : G_2 \to \{0,1\}^{\log p}$. So, the global parameter $gp = (G_1, G_2, e, H_1(\cdot), H_2(\cdot), g, h)$, where $h \in G_1$ are random values.

$KeyGen_{Receiver}(\boldsymbol{gp})$**.** Input global parameter gp and choose a random value $\alpha \in Z_p^*$. It outputs public key $pk_R = [g, h = g^\alpha]$ and private key $pk_R = \alpha$.

$PEKS(pk_R, w)$**.** First compute $t = e(H_1(w), h^r) \in G_2$ where w is the keyword for a random value $r \in Z_p^*$. Output $C = [g^r, H_2(t)]$.

$Trapdoor(sk_R, w)$**.** Output $T_w = H_1(w)^\alpha \in G_1$.

$Test(pk_R, S, T_w)$**.** Let $S = [A, B]$. Test if $H_2(e(T_w, A)) = B$. If so, output 'yes'; if not, output 'no.'

2.3 Secure Channel Free Public Key Encryption with Keyword Search (SCF-PEKS)

Secure channel free public key encryption with keyword search, also named searchable keyword encryption with designated tester, was developed by Baek *et al.* [1] in 2008. The proposed scheme is based on bilinear pairing, and the algorithm is as follows:

Global Parameter: G_1 and G_2 are two cyclic groups and $e : G_1 \times G_1 \to G_2$ is a bilinear map where the order is q. P is a generator of G_1 and $e(g, g)$ is a generator of G_2. There are two hash function $H_1 : \{0,1\}^* \to G_1^*$ and $H_2 : G_2 \to \{0,1\}^p$. So, the global parameter $gp = (q, G_1, G_2, e, P, H_1(\cdot), H_2(\cdot), d_w)$, where d_w denotes a keyword space.

$KeyGen_{Server}(\boldsymbol{gp})$**:** Choose a random value $x \in Z_q^*$ and compute $X = xP$. Choose a random value $Q \in G_1^*$. Return public key $pk_S = (Q, X)$ and private key $sk_S = x$.

$KeyGen_{Receiver}(\boldsymbol{gp})$**:** Choose a random value $y \in Z_q^*$ and compute $Y = yP$. Return public key $pk_R = Y$ and private key $sk_R = y$.

$SCF - PEKS(pk_S, pk_R, w)$**:** Choose a random value $r \; in Z_q^*$ and compute $k = (e(Q, X)E(H_1(w), Y))^r$. Output ciphertext $C = (U, V) = (rP, H_2(k))$.

$Trapdoor(sk_R, w)$**:** Output $T_w = yH_1(w)$.

$Test(sk_S, C, T_w)$**:** Test if $H_2(e(xQ + T_w, U))$. If so, output 'yes'; if not, output 'no.'

2.4 ElGamal Public Key Cryptography

ElGamal public key cryptography system was created in 1985, and its security rests on the difficulty of solving the discrete logarithm problem. Compared to the RSA cryptography system, the same document can produce different ciphertexts as encrypting by using ElGamal cryptography. That is, the same data encrypted at different times are distinct thoroughly, so the attacker can not learn the plain text from gathering a large amount of ciphertext.

ElGamal Encryption

Suppose Bob wants to send a secret message M to Alice, he has to encrypt the message with Alice's public key, and Alice's private key can only decrypt the ciphertext. The operation is as follows:

Key Generation: Let G be a cyclic group of prime order P with generator g, and let Z_P be the field of integers modulo P. Alice randomly chooses a private key x and computes the public key $y = g^x \bmod P$.

Encrypt: First, Bob chooses a random value $r \in Z_P$ and compute

$$b = g^r \bmod P,$$
$$c = M \times y^r \bmod P.$$

Finally, Bob sends ciphertext b and c to Alice.

Decrypt: As Alice receiving ciphertext b and c, she uses her private key x to decrypt the ciphertext:

$$M = c \times (b^x)^{-1} \bmod P.$$

ElGamal Signature

Before Bob sends the ciphertext to Alice, he must also create his digital signature.

Sign If Bob's private key is x and the public key is (y, P) where $y = g^x \bmod P$. First, Bob choose a value k which $gcd(k, P-1) = 1$ and compute $r = g^k \bmod P$. Then, compute

$$s = k^{-1}(M - xr) \bmod (P-1).$$

Finally, Bob sends (M, r, s) to Alice.

Verify As Alice gets the ciphertext and Bob's signature, she can confirm Bob's identity using Bob's public key. She only has to confirm whether the following equation holds.

$$g^M = y^r r^s \bmod P.$$

3 Analysis of Two SCF-PEKS Scheme against KGA

In this section, we choose two SCF-PEKS schemes based on bilinear pairs and try to attack them with an offline keyword-guessing attack.

3.1 Attack on Scheme in [11]

The first scheme we review is Hu and Liu's scheme. The following is how Hu and Liu's scheme works:

Given a security parameter λ, it returns a global parameter $gp = (G_1, G_2, e, H_1, H_2, g)$, where $H_1 : \{0,1\}^* \to G_1, H_2 : G_2 \to \{0,1\}^*$.

$KeyGen_{Server}(gp)$: It randomly chooses $\alpha \in_R (Z_P)^*$, and $Q \in_R G_1$, and returns $sk_S = \alpha$ and $pk_S = (pk_1, pk_2) = (Q, g^\alpha)$ as a server's pair of private and public keys.

$KeyGen_{Receiver}(gp)$: It randomly chooses $x, t \in_R Z_P^*$, and returns $sk_R = (x, t)$ and $pk_R = (pk_{R_1}, pk_{R_2}, pk_{R_3}, pk_{R_4}) = (g^x, g^{tx^2}, g^{xt}, pk_{S_2})$ as a receiver's pair of private and public keys.

$SCF - PEKS(gp, pk_R, pk_S, w)$: This algorithm randomly picks a value $r \in_R Z_P^*$, and check if $e(pk_{R_1}, pk_{R_4}) = e(pk_{R_3}, pk_{S_2})$. If the equality is satisfied, then outputs $C = [A, B] = [(pk_{R_2})^r, H_2(e(pk_{R_4}, H_1(w)^r))]$.

$Trapdoor(gp, pk_S, sk_R, w)$: This algorithm randomly picks a value $r' \in_R Z_P^*$, and outputs $T_w = [T_1, T_2] = [(pk_{S_2})^{r'}, H_2(w)^{1/x^2} \cdot g^{r'}]$.

$Test(gp, C, sk_S, T_w)$: This algorithm computes $T = (T_2)^\alpha / T_1$ and checks if $B = H_2(e(A, T))$. If the equality is satisfied, output "1"; otherwise, output "0".

Now, we show Hu and Liu's scheme is insecure against offline keyword guessing attacks (OKGA) by inside attackers. The following parameters $(g, pk_{R_2}, pk_{R_4}, H_1)$ are public. The attacker can perform OKGA as follows:

1) The attacker compute $T = T_2^\alpha / T_1$.
2) The attacker guess a keyword w' and compute $H_1(w')$.
3) The attacker check if $e(pk_{R_2}, T) = e(pk_{R_4}, H_1(w'))$. If the equation holds, the guessed keyword w' is valid; otherwise, go to 2.

3.2 Attack on Scheme in [20]

Let G and G_T be bilinear groups of prime order p. Given a security parameter λ. First, we pick a random generator $g \in G$ and several random elements $u, \tilde{u} \in G$. Let $H : \{0,1\}^* \to G$, $H_1 : \{0,1\}^* \to G$ and $H_2 : G_T \to \{0,1\}^\lambda$ be hash functions that are modeled as a random oracle. The global parameter $gp = (p, G, G_T, e, H(\cdot), H_1(\cdot), H_2(\cdot), g, u, \tilde{u})$.

$KeyGen_{Server}(gp)$: Takes as an input gp. This algorithm chooses a random exponent $\alpha \in Z_p$, and sets $sk_S = \alpha$, and computes $pk_S = (pk_{S,1}, pk_{S,2}) = (g^\alpha, u^{1/\alpha})$. This algorithm outputs (pk_S, sk_S).

$KeyGen_{Receiver}(gp)$: Takes as an input gp. This algorithm chooses a random exponent $\beta \in Z_p$, and sets $sk_R = \beta$, and computes $pk_R = (pk_{R,1}, pk_{R,2}) = (g^\beta, \tilde{u}^\beta)$. This algorithm outputs $(pk)R, sk_R)$.

$SCF - PEKS(gp, pk_R, pk_S, w)$: Takes as inputs $emphgp$, the receiver's public key $pk_R = (pk_{R,1}, pk_{R,2})$, the server's public key $pk_S = (pk_{S,1}, pk_{S,2})$, and a keyword w. This algorithm chooses a random value $r \in Z_p$ and sets $A = pk_{R,1}^r$ and $B = H_2(e(pk_{s,1}, H_1(w)^r))$. This algorithm outputs a $dPEKS$ ciphertext $C = [A, B]$.

$Trapdoor(gp, pk_S, sk_R, w)$: Takes as inputs gp, the server's public key $pk_S = (pk_{S,1}, pk_{S,2})$, the receiver's secret key $sk_R = \beta$, and a keyword w. This algorithm chooses a random value $r' \in Z_p$ and computes $T_1 = g^{r'}$ and $T_2 = H_1(w)^{1/\beta} \cdot H(pk_{S,1}^{r'})$. This algorithm outputs a trapdoor $T_w = [T_1, T_2]$.

$Test(gp, C, sk_S, T_w)$: Takes as inputs gp, a $dPEKS$ ciphertext $C = [A, B]$, the server's secret key $sk_S = \alpha$, and a trapdoor $T_w = [T_1, T_2]$. This algorithm computes $T = T_2 / H_2(T_1^\alpha)$ and checks if $B = H_2(e(A, (T)^\alpha))$ holds. If the above equalities are satisfied, output 'yes'; otherwise, output 'no.'

Now, we show Rhee *et al.*'s scheme is insecure against offline keyword guessing attacks (OKGA) by inside attackers. The attacker can perform OKGA as follows:

1) The attacker compute $T = T_2/H(T_1^{sk_S}) = H_1(w)^{1/sk_R}$.
2) The attacker guess a keyword w' and compute $H_1(w')$.
3) The attacker check if $e(pk_{R_1}, T) = e(g, H_1(w'))$. If the equation holds, the guessed keyword w' is valid; otherwise, go to 2.

4 The Proposed Scheme

After reviewing the above concepts and scheme, we try to develop a new scheme that can against outside keyword guessing attacks. We apply the ElGamal cryptography system to improve the efficiency of SCF-PEKS. The algorithm is as follows:

Let G be a cyclic group of prime order P with generator g.

$KeyGen_{Server}$: First, select a value randomly x as private key sk_S and compute $X = g^x \bmod P$ as public key pk_S.

$KeyGen_{Receiver}$: Similar to the key generate steps of server. First, select a random value y as private key sk_R and compute $Y = g^y \bmod P$. Set public key $pk_R = Y$.

SCF-PEKS: In this step, data sender encrypted the keyword w with server's public key pk_S and appends them to the encrypted message. First, choose two random values $\alpha, \theta \in Z_P$ and compute

$$a_1 = g^\alpha \bmod P,\ b_1 = \theta H(w) \times pk_S^\alpha \bmod P.$$

Then, output the ciphertext of keyword $C = [C_1, C_2, C_3] = [\theta pk_R, a_1, b_1]$ and send C to the server.

$Trapdoor$: In this step, the authorized receiver produces a trapdoor with the keyword w' they want to search. There are two parts to generating Trapdoor $T_{w'}$: create a signature using sk_R and encrypt the $H(w')$ using pk_S. In the first part, the receiver chooses a value k which $gcd(k, P-1) = 1$. Then, compute

$$r = g^k \bmod P \text{ and } S = k^{-1} \bmod P.$$

Finally, output (H(w'), r, S). In the second part, the receiver chooses another random value $\beta \in Z_P$. Then, compute

$$a_2 = g^\beta \ mod\ P, b_2 = H(w) \times ps_S^\beta \ mod\ P.$$

After two parts, it will output the trapdoor $T_{w'} = [T_1, T_2, T_3, T_4] = [a_2, b_2, S, r]$ and send $T_{w'}$ to the server.

$Test$: After receiving the trapdoor $T_{w'}$, server start to find the corresponding keyword ciphertext. First, decrypts ciphertext C and trapdoor T_w with the server's private key sk_S as follows:

$$v = b_1 \times (a_1^{sk_S})^{-1}\ mod\ P = \theta H(w)$$
$$u = b_2 \times (a_2^{sk_S})^{-1}\ mod\ P = H(w')$$

Then, compute $Z = (C_1)^r \times r^S \ mod\ P$. Finally, check if $vg^u \ mod\ P = uZ \ mod\ P$. If it holds, it means $w = w'$.

5 Discussion and Analysis

5.1 Security Analysis

This section will analyze whether the proposed scheme meets the criteria in Section 1.

Unforgeable of the trapdoor: The receiver produces the trapdoor with the keyword they want to search for and their secret key. Since our scheme is based on ElGamal cryptography, an attacker can get nothing before solving the discrete logarithm problem. That is, an attacker cannot get any information from trapdoors.

On the other hand, the receiver will select two random values to produce the trapdoor every time they query the data, so the attacker cannot get any information from gathering it.

Anonymous of the ciphertext: As the data sender encrypts the keyword w before uploading to the server, they select a random value to protect $H(w)$. Although the server can decrypt C_3, he cannot learn the keyword w by performing an offline keyword guessing attack. That is, an attacker cannot get any information from the ciphertext.

Authorized identity protection: When the data sender encrypts the keyword, they will also authorize a user who can download the data. In the proposed scheme, the data sender will select a random value to protect $H(w)$ and the authorized receiver's public key so others cannot know the authorized user's identity.

User authentication: Since others cannot learn the keyword, authorized identity, and anyone's secret key, the server still can perform the $Test$ algorithm to match trapdoor with ciphertext that is stored in the server and send data to the authorized user. The proposed scheme can achieve user authentication without revealing any information.

5.2 Performance Evaluation

The proposed scheme is based on the ElGamal cryptography system. Here, we compare the efficiency with [11] and [20] on $SCF-PEKS$ and $Trapdoor$ algorithm.

From Table 1, we could find that our scheme only has one exponentiation computation and one hash function computation required by the data sender in the $SCF-PEKS$ algorithm, and the receiver requires three exponentiation computations and one hash function computation in the *Trapdoor* algorithm. Therefore, our scheme is more efficient than [11] and [20].

6 Conclusions

This paper presents a more efficient secure channel-free public key encryption with a keyword search (SCF-PEKS) scheme based on ElGamal cryptography. This scheme is secure against offline keyword guessing attacks and more efficient than other SCF-PEKS schemes. We also give out a simple analysis of the security criteria of our scheme. However, this scheme can extend to various query types, such as conjunctive keyword searches, range queries, etc.

Acknowledgments

The Ministry of Science and Technology partially supported this research, Taiwan (ROC), under contract no.: MOST 109-2221-E-468-011-MY3, MOST 108-2410-H-468-023, and MOST 111-2622-8-468-001 -TM1.

References

[1] J. Baek, R. Safavi-Naini, W. Susilo, "Public key encryption with keyword search revisited," in *Computational Science And Its Applications (ICCSA'08), Lecture Notes in Computer Science*, vol. 5072, pp. 1249–1259, 2008.

[2] D. Boneh, G. Di Crescenzom, R. Ostrovsky, G. Rersiano, "Public key encryption with keyword search," in *Advances in Cryptology (EUROCRYPT'04), Lecture Notes in Computer Science*, vol. 3027, pp. 506–522, 2004.

[3] J. W. Byun, H. S. Rhee, H. A. Park, and D. H. Lee, "Off-line keyword guessing attacks on recent keyword search schemes over encrypted data," in *Secure Data Management, Lecture Notes in Computer Science*, vol. 4165, pp. 75–83, 2006.

[4] R. Canetti, O. Goldreich, and S. Halavi, "The random oracle methodology, revisited," in *In: Proceedings of 30th ACM STOC*, pp. 209–218, New York, 2004.

[5] Z. Cao, C. Mao, L. Liu, W. Kong, J. Wang, "Analysis of one dynamic multi-keyword ranked search scheme over encrypted cloud data," *International Journal of Network Security*, vol. 20, no. 4, pp. 683-688, 2018.

[6] L. Fang, W. Susilo, C. Ge, and J. Wang, "A secure channel free public key encryption with keyword search scheme without random oracle," in *Cryptology and Network Security, Lecture Notes in Computer Science*, vol. 5888, pp. 248–258, 2009.

[7] T. Feng, H. Pei, P. Xie, and X. Feng , "Blockchain data sharing scheme based on searchable agent re-encryption," *International Journal of Network Security*, vol. 23, no. 3, pp. 535-544, 2021.

[8] T. Feng, X. Yin, Y. Lu, J. Fang, and F. Li, "A searchable CP-ABE privacy preserving scheme," *International Journal of Network Security*, vol. 21, no. 4, pp. 680-689, 2019.

[9] C. Gu and Y. Zhu, "New efficient searchable encryption schemes from bilinear pairings," *International Journal of Network Security*, vol. 10, no. 1, pp. 25-31, 2010.

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

[11] C. Hu and P. Liu, "A secure searchable public key encryption scheme with a designated tester against keyword guessing attacks and its extension," in *Advances in Computer Science, Environment, Ecoinformatics, and Education, Communications in Computer and Information Science*, vol. 215, pp. 131–136, 2011.

[12] M. Hu, H. Gao, T. Gao, "Secure and efficient ranked keyword search over outsourced cloud data by chaos based arithmetic coding and confusion," *International Journal of Network Security*, vol. 21, no. 1, pp. 105-114, 2019.

[13] F. G. Jeng, S. Y. Lin, B. J. Wang, C. H. Wang, T. H. Chen, "On the security of privacy-preserving keyword searching for cloud storage services," *International Journal of Network Security*, vol. 18, no. 3, pp. 597-600, 2016.

[14] C. C. Lee, S. T. Hsu, M. S. Hwang, "A study of conjunctive keyword searchable schemes," *International Journal of Network Security*, vol. 15, no. 5, pp. 321-330, 2013.

[15] Z. Liu, Y. Fan, "Provably secure searchable attribute-based authenticated encryption scheme," *International Journal of Network Security*, vol. 21, no. 2, pp. 177-190, 2019.

[16] Z. Liu, Y. Liu, J. Xu, B. Wang, "Verifiable attribute-based keyword search encryption with attribute revocation for electronic health record system," *International Journal of Network Security*, vol. 22, no. 5, pp. 845-856, 2020.

[17] Z. Liu, F. Yin, J. Ji, B. Wang, "Revocable and searchable attribute-based encryption scheme with multi-keyword and verifiability for internet of things," *International Journal of Network Security*, vol. 23, no. 2, pp. 205-219, 2021.

[18] J. Ren, L. Zhang, and B. Wang, "Decentralized multi-authority attribute-based searchable encryption scheme," *International Journal of Network Security*, vol. 23, no. 2, pp. 332-342, 2021.

[19] H. S. Rhee, J. H. Park, W. Susilo, and D. H. Lee, "Improved searchable public key encryption with designated tester," in *Proceedings of the 4th International Symposium on Information, Computer, and*

Table 1: Efficient Comparison

Algorithms	Hu and Liu [11]	Rhee *et al.* [20]	Our Scheme
$SCF-PEKS$	$2T+Exp{+}2T_H$	$2T_{Exp}{+}2T_H$	$1T_{Exp}+1T_H$
$Trapdoor$	$3T_{Exp}{+}1T_H$	$3T_{Exp}{+}2T_H$	$3T_{Exp}{+}1T_H$

T_{Exp}: The computing time of exponentiation.
T_H: The computing time of hash.

Communications Security (ASIACCS'09), pp. 376–379, 2009.

[20] H. S. Rhee, J. H. Park, W. Susilo, and D. H. Lee, "Trapdoor security in a searchable public-key encryption scheme with a designated tester," *Journal of Systems and Software*, vol. 83, no. 5, pp. 763–771, 2010.

[21] D. Thiyagarajan, R. Ganesan, "Cryptographically imposed model for efficient multiple keyword-based search over encrypted data in cloud by secure index using bloom filter and false random bit generator," *International Journal of Network Security*, vol. 19, no. 3, pp. 413-420, 2017.

[22] Jie Wang, Xiao Yu, and Ming Zhao , "Fault-tolerant verifiable keyword symmetric searchable encryption in hybrid cloud," *International Journal of Network Security*, vol. 17, no. 4, pp. 471-483, 2015.

[23] X. Wang, S. Yin, H. Li, L. Teng, and S. Karim, "A modified homomorphic encryption method for multiple keywords retrieval," *International Journal of Network Security*, vol. 22, no. 6, pp. 905-910, 2020.

[24] Y. Wang, W. Bao, Y. Zhao, X. Hu, Z. Qin, "An ElGamal encryption with fuzzy keyword search on cloud environment," *International Journal of Network Security*, vol. 18, no. 3, pp. 481-486, 2016.

[25] Y. Wang, Y. Zhu, and J. Wang, "Towards forward secure conjunctive searchable symmetric encryption with result pattern hidden," *International Journal of Network Security*, vol. 24, no. 2, pp. 273-285, 2022.

[26] Q. Wu, X. Ma, L. Zhang, and Y. Chen , "Expressive ciphertext policy attribute-based searchable encryption for medical records in cloud," *International Journal of Network Security*, vol. 23, no. 3, pp. 461-472, 2021.

[27] H. M. Yang, C. X. Xu, and H. T. Zhao, "An efficient public key encryption with keyword scheme not using pairing," in *2011 First International Conference on Instrumentation, Measurement, Computer, Communication and Control*, pp. 900–904, 2011.

[28] W. C. Yau, S. H. Heng, and B. M. Goi, "Off-line keyword guessing attacks on recent public key encryption with keyword search schemes," in *Autonomic and Trusted Computing, Lecture Notes in Computer Science*, vol. 5060, pp. 100–105, 2008.

[29] B. Zhang and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search," *Journal of Network and Computer Application*, vol. 34, no. 1, pp. 262–267, 2011.

[30] L. Zou, X. Wang, S. Yin, "A data sorting and searching scheme based on distributed asymmetric searchable encryption," *International Journal of Network Security*, vol. 20, no. 3, pp. 502-508, 2018.

Biography

Min-Shiang Hwang received M.S. in industrial engineering from National Tsing Hua University, Taiwan in 1988, and a Ph.D. degree in computer and information science from National Chiao Tung University, Taiwan, in 1995. He was a distinguished professor and Chairman of the Department of Management Information Systems, National Chung Hsing University, during 2003-2011. He obtained 1997, 1998, 1999, 2000, and 2001 Excellent Research Awards from the National Science Council (Taiwan). Dr. Hwang was a dean of the College of Computer Science, Asia University (AU), Taichung, Taiwan. He is currently a chair professor with the Department of Computer Science and Information Engineering, AU. His current research interests include information security, electronic commerce, database, data security, cryptography, image compression, and mobile computing. Dr. Hwang has published over 300+ articles on the above research fields in international journals.

Shih-Ting Hsu received his M.S. degree in Department of Management Information Systems from National Chung Hsing University, Taichung, Taiwan (ROC), in 2019. His research interests include network security, security and privacy of cloud computing, and applied cryptography.

Cheng-Ying Yang received the M.S. degree in Electronic Engineering from Monmouth University, New Jersey, in 1991, and Ph.D. degree from the University of Toledo, Ohio, in 1999. He is a member of IEEE Satellite & Space Communication Society. Currently, he is employed as a Professor at Department of Computer Science, University of Taipei, Taiwan. His research interests are performance analysis of digital communication systems, error control coding, signal processing and computer security.

Reviewers (Volume 25, 2023)

Dariush Abbasinezhad
Tarek Abbes
Ahmed Abd El-Rahiem Abd El-Latif
Slim Abdelhedi
Mohd Faizal Abdollah
Ahmed Mohammed Abdullah
Subrata Acharya
Sodeif Ahadpou
Tohari Ahmad
Muhammad Najmi Ahmad-Zabidi
Mohammad Reza Ahmadi
Asimi Ahmed
Ganesh V. Aithal
Mehrnaz Akbari Roumani
Abdul-Gabbar Tarish Al-Tamimi
Aws N. Al-Zarqawee
Monjur M Alam
Shahid Alam
Tanweer Alam
Dilip S Aldar
Sara Ali
Ali Mohamed Allam
Khalid Abdulrazzaq Alminshid
Seth Alornyo
Ali Mohammed Alsahlany
Richard Amankwah
Ruhul Amin
Rengarajan Amirtharajan
R. Anand
Karl Andersson
Ruo Ando
Benjamin Arazi
K. S. Arvind
Muhammad Asad
Travis Atkison
Hany Fathy Atlam
Cossi Blaise Avoussoukpo
Nancy A. Awad
Zulkhar Nain Bin Badruz
Anant M. Bagade
Amandeep Bagga
Nazrulazhar Bahaman
Saeed Bahmanabadi
Nischay Bahl
Anuj Kumar Baitha
Saad Haj Bakry
R. R. Balakrishnan
Kavitha Balu
Maram Y Bani Younes
Tamer Mohamed Barakat
Utpal Barman
Pijush Barthakur
Eihab Bashier Mohammed Bashier
Adil Bashir
Sunny Behal
Rydhm Beri
Taran Singh Bharati
Akashdeep Bhardwaj
Lathies T. Bhasker
Sugandh Bhatia
Sajal Bhatia
Dharmendra Bhatti
Krishna Bhowal
Li Bin
Sumitra Binu
Zhengjun Cao
Liling Cao
Chi-Shiang Chan
Eric Chan-Tin
Mohan Kumar Chandol
Yogesh Chandra
Arup Kumar Chattopadhyay
Nirbhay K. Chaubey
Ali M Chehab
Chi-Hua Chen
Chin-Ling Chen
Jan Min Chen
Qihong Chen
Tzung-Her Chen
Xi Chen
Yang Chen
Yi-Hui Chen
Yushuang Chen
Zhixiong Chen
Qingfeng Cheng
Kaouthar Chetioui
Mao-Lun Chiang
Shu-Fen Chiou

Tae-Young Choe
Kim-Kwang Raymond Choo
Christopher P. Collins
Joshua C. Dagadu
Ashok Kumar Das
Prodipto Das
Sanjoy Das
Debasis Das
Ranjan Kumar Dash
Subhrajyoti Deb
Abdelrahman Desoky Desoky
Mooramreddy Sree Devi
Sankhanil Dey
Subhasish Dhal
Jintai Ding
Jingnan Dong
Xiaoli Dong
Nishant Doshi
Ahmed Drissi
Crystal Wilson Dsouza
Qi Duan
Ashraf Diaa Elbayoumy
Abd Allah Adel Elhabshy
Ahmed A. Elngar
Edwin Engin Yaz
Aoxiong Fan
Arizona Firdonsyah
Xingbing Fu
Vladimir Sergeevich Galyaev
Rakesh C Gangwar
Juntao Gao
Tiegang Gao
Xinwei Gao
N. B. Gayathri
G. Geetha
Mohammad GhasemiGol
Madhumala Ghosh
Ramesh Gopalan
Poornima Ediga Goud
Krishan Kumar Goyal
Ke Gu
Avinash k Gulve
Sumalatha Gunnala
Shuai Guo
C. P. Gupta
Jatin Gupta
Pynbianglut Hadem
Charifa Hanin
Ali Hassan
Wien Hong
Tsung-Chih Hsiao
Chengyu Hu
Defa Hu
Xiong Hu
Yen-Hung Hu
Huajun Huang
Chin-Tser Huang
Jianmeng Huang
Munawar Hussain
Bala Venkateswarlu Isunuri
Grasha Jacob
Amit Jain
Yogendra Kumar Jain
Swati Jaiswal
Teena Jaiswal
Bappaditya Jana
V. S. Janani
N Jeyanthi
lin zhi jiang
Shaoquan Jiang
Rong Jiang
Rui Jiang
Zhengping Jin
Ashish Joshi
Li Su Juan
Omprakash Kaiwartya
Yoshito Kanamori
Nirmalya Kar
Gagandeep Kaur
Wongyos Keardsri
Omar Khadir
Vaishali D. Khairnar
Asif Uddin Khan
Md. Al-Amin Khandaker
Malik Sikander Hayat Khiyal
Dong Seong Kim
Kingsford Kissi Mireku
Vikas K Kolekar
P. Dhandapani Raman D. Kothandaraman
Anjan Krishnamurthy
Fengfei Kuang
Sajja Ratan Kumar
Manish Kumar
Naresh N Kumar
Saru Kumari
Yesem Kurt Peker
Owusu-Agyemang Kwabena
Albert Kofi Kwansah Ansah
Manmohan Lakhera
Then Lee
Cheng Li

Chun-Ta Li
Yanping Li
Zhaozheng Li
H. M. Lian
Changlu Lin
Chia-Chen Lin
Chih-Yang Lin
Iuon-Chang Lin
Yang-Bin Lin
Jiang Hong Ling
Desheng Liu
Li Liu
Shuang Gen Liu
Ting Liu
Ximeng Liu
Yanjun Liu
Yining Liu
K. Shantha Kumari Luke
Jayakumar
Zhiyong Luo
Ming Luo
Sagar Bhaskar Mahajan
Zahid Mahmood
Tanmoy Maitra
Doaa Mohsin Majeed
Arun Malik
Mahalinga V. Mandi
T. Manesh
Palvinder Singh Mann
Ali Mansouri
A. M. Meddeb-Makhlouf
Kamran Ali Memon
Bo Meng
Weizhi Meng
Yang Ming
Suhail Qadir Mir
Amit Mishra
Anuranjan Misra
Syed Shahul Hameed
Mohamed Ismail
Sirwan Ahmed Mohammed
Madihah Mohd Saudi
Guillermo Morales-Luna
Belmekki Mostafa
Alaa Moualla
Hamdy M. Mousa
Muhammad M. Muhammad
Kuntal Mukherjee
C. H. Mukundha
Bhagavathi Priya M
Muthumanikam
Ambika Nagaraj
Preeti Nagrath
K. Nandhini
Syed Naqvi
Kanagaraj Narayanasamy
Lakshmi Kannan Narnayanan
Prabir Kr Naskar
Sarmistha Neogy
Krishnamur G Ningappa
Sohail Noman
Chokri Nouar
Abdul Abiodun Orunsolu
Arezou Ostad Sharif
Nasrollah Pakniat
Dhiraj Pandey
S. K. Pandey
B. D. Parameshachari
Subhash S. Parimalla
Chintan J. Patel
Kailas Ravsaheb Patil
Suresh Kumar Peddoju
Hongmei Pei
Kanthakumar Pongaliur
A. Prakash
Krishna K. Prakash
Munivara Prasad
Hongquan Pu
Yudha Purwanto
Septafiansyah Dwi Putra
Murad Abdo Rassam Qasm
Qais Saif Qassim
Chuan Qin
Jiaohua Qin
Narasimhan Renga Raajan
Hashum Mohamed Rafiq
Abdul Hamid M. Ragab
V. Sampangi Raghav
Uma R. Rani
Ganga Rama Koteswara Rao
Golagani A.V.R.C Rao
Mohammad Maher Rasheed
V. Rathinasabapathy
Dhivya Ravi
Ramesh S Rawat
Siva Ranjani Reddi
Khaled Riad
Mohd Foad Rohani
Ou Ruan
Sanjay Kumar Sahay
Ashish Saini
Debabrata Samanta

Sabyasachi Samanta
Manju Sanghi
Arif Sari
Balamurugan K. S. Sathiah
Rajat Saxena
Michael Scott
Chandra Vorugunti Sekhar
Irwan Sembiring
Elena Sendroiu
Divyashikha Sethia
Vrutik M. Shah
Vrushank Shah
Kareemulla Shaik
Tarun Narayan Shankar
Udhayakumar Shanmugam
Rohith Shivashankar
Abhishek Shukla
Varun Shukla
Anuj Kumar Singh
Debabrata Singh
Jitendra Singh
Mahendra Pratap Singh
Mukesh Singh
Bala Srinivasan
Siva Shankar Subramanian
Karthikeyan Subramanian
T. SudalaiMuthu
K. S. Suganya
Guodong Su
Haiyan Sun
Fei Tang
Maryam Tanha
Ariel Soares Teles
Pratik Teli
Xiuxia Tian
Geetam Singh Tomar
Yuan-Yu Tsai
Pushpendra Kumar Verma
Ravi Verma
Vandani Verma
Vibhor Kumar Vishnoi
Phu Vo Ngoc
Putra Wanda
Ding Wang
Fangwei Wang
Feng Wang
Guoqing Wang
Li Wang
Libin Wang
Linfan Wang
Qingping Wang
Xiaogang Wang
Xingbo Wang
Xu Wang
Ying Wang
Jianghong Wei
Zhe Wei
Axin Wu
Na-I Wu
Chengbo Xu
Degang Xu
Lei Xu
Chengbo Xu
Yashveer Yadav
Wei Yajuan
Jun Yan
Changsong Yang
Li Yang
Rui Yang
Wenjie Yang
Yifei Yao
Jun Ye
Pinghao Ye
Fangfang Yin
Huang Yiwang
Lin You
Huifang Yu
Lei Yu
Hang Yue
Taskeen Zaidi
Noor Zaman Zaman
Jianjun Zhang
Sherali Zeadally
Jianping Zeng
Fangguo Zhang
Futai Zhang
Jianhong Zhang
Jie Xiu Zhang
Qiu-Yu Zhang
Shanshan Zhang
Yanshuo Zhang
Yinghui Zhang
Zonghua Zhang
Hongzhuan Zhao
Mingju Zhao
Yuntao Zhao
Zhiping Zhou
Ye Zhu
Yingwu Zhu
Frank Zhu
Aaron Zimba

Guide for Authors
International Journal of Network Security

IJNS will be committed to the timely publication of very high-quality, peer-reviewed, original papers that advance the state-of-the art and applications of network security. Topics will include, but not be limited to, the following: Biometric Security, Communications and Networks Security, Cryptography, Database Security, Electronic Commerce Security, Multimedia Security, System Security, etc.

1. Submission Procedure

Authors are strongly encouraged to submit their papers electronically by using online manuscript submission at http://ijns.jalaxy.com.tw/.

2. General

Articles must be written in good English. Submission of an article implies that the work described has not been published previously, that it is not under consideration for publication elsewhere. It will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

2.1 Length Limitation:

All papers should be concisely written and be no longer than 30 double-spaced pages (12-point font, approximately 26 lines/page) including figures.

2.2 Title page

The title page should contain the article title, author(s) names and affiliations, address, an abstract not exceeding 100 words, and a list of three to five keywords.

2.3 Corresponding author

Clearly indicate who is willing to handle correspondence at all stages of refereeing and publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address.

2.4 References

References should be listed alphabetically, in the same way as follows:
For a paper in a journal: M. S. Hwang, C. C. Chang, and K. F. Hwang, ``An ElGamal-like cryptosystem for enciphering large messages," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 2, pp. 445--446, 2002.

For a book: Dorothy E. R. Denning, *Cryptography and Data Security*. Massachusetts: Addison-Wesley, 1982.

For a paper in a proceeding: M. S. Hwang, C. C. Lee, and Y. L. Tang, ``Two simple batch verifying multiple digital signatures," in *The Third International Conference on Information and Communication Security (ICICS2001)*, pp. 13--16, Xian, China, 2001.

In text, references should be indicated by [number].

Subscription Information

Individual subscriptions to IJNS are available at the annual rate of US$ 200.00 or NT 7,000 (Taiwan). The rate is US$1000.00 or NT 30,000 (Taiwan) for institutional subscriptions. Price includes surface postage, packing and handling charges worldwide. Please make your payment payable to "Jalaxy Technique Co., LTD." For detailed information, please refer to http://ijns.jalaxy.com.tw or Email to ijns.publishing@gmail.com.